Developing Response Metrics for the Economic Census

Justin Ward[1] and Diane K. Willimack[1]

U.S. Census Bureau, 4600 Silver Hill Rd, Suitland, MD 20746

Abstract

In the past, the U.S. Census Bureau used the check-in rate to measure the performance of data collection operations for the Economic Census. The check-in rate is the proportion of mailout that has returned a form and serves as a proxy indicator of response, but does not necessarily indicate receipt of usable data. We are developing unit and item response metrics to provide more meaningful information about data quality.

Since we calculate the unit response rate after completion of all data collection and processing operations, we developed a proxy metric for the unit response rate that can be calculated in real time and, along with the check-in rate, used to monitor and manage collection activities. We use 2012 Economic Census data to examine the behavior of these new response metrics.

The unit response rate and its proxy will always be lower than the check-in rate, which includes units returning forms without usable data. During collection, the proxy performs similarly to the check-in rate. After collection, the proxy metric and the unit response rate are similar with their differences mainly explained by limitations of information available during collection.

Key Words: Economic Census, response rates, data quality

1. Introduction

The U.S. Census Bureau conducts a mandatory economic census every five years covering reference years that end in 2 or 7. For example, preparations for the 2017 Economic Census are now in their final stages, with data for reference year 2017 being collected during 2018. Detailed financial and product data are collected at the establishment level, and statistics are produced for 18 of the 20 economic sectors covered by the North American Industrial Classificiation System (NAICS), with geographic detail down to the ZIP code level, where possible, along with state and national estimates. Policymakers, industry analysts, economic development commissions, academic researchers, and businesses themselves use Economic Census data. Primary uses include estimation of U.S. Gross Domestic Product, input-output and productivity measures, and a variety of other purposes such as gauging industry growth and competitiveness. The Economic Census is also used to update the Census Bureau's Business Register (BR), from which sample frames are extracted for industry-level annual, quarterly and monthly economic surveys.

In the past, the Economic Census has used "check-in rate" to monitor performance of data collection operations. The check-in rate, or return rate, is the ratio of the number of returned forms to the number of forms mailed out. However, since a returned form does not necessarily indicate useable data were received, "check-in rate" did not meet the Statistical

---

[1] Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Quality Standards of the Office of Management and Budget (OMB) (2006, see https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/statpolicy/standards _stat_surveys.pdf) or the Census Bureau (2010, see http://www.census.gov/quality/standards/index.html).

Most ongoing economic survey programs at the Census Bureau compute two types of response rates: a unit response rate and weighted item response rates. (See Census Bureau Statistical Quality Standards, Appendix D3-B: Requirements for Calculating and Reporting Response Rates: Economic Surveys and Censuses, https://www.census.gov/about/policies/quality/standards/appendixd3b.html). The unit response rate (URR) is the ratio of the total unweighted number of "responding units" to the total number of reporting units eligible for data collection. The URR is an indicator of the performance of data collection for obtaining usable responses. Weighted item-level response rates include Quantity and Total Quantity Response Rates (QRR and TQRR). These rates measure the proportion of each estimate obtained directly or indirectly from the survey unit. The QRR measures the weighted proportion of an estimate obtained directly from respondents; the TQRR expands the rate to include data from acceptable secondary sources of sufficient quality, such as administrative records. The QRR and TQRR are indicators of the quality of an estimate at the item-level.

In 2011, a team of mathematical statisticians and subject matter experts was chartered to develop response metrics for the Economic Census that were analogous to those provided by other economic programs to provide information on data quality for data users and stakeholders, as well as to meet OMB and Census Bureau quality standards. In addition, the team also developed metrics indicating data quality to augment "check-in rate" as a tool for monitoring and managing real-time data collection processes. The team completed its work in 2015, for implementation during the 2017 Economic Census. Previous papers by Lineback et al. (2012) and Fink and Lineback (2015) describe earlier stages of this work, including its motivation and features of the economic census design impacting the development of these metrics. This paper documents completion of this work, provides final definitions of various response metrics, and illustrates these metrics with computations based on 2012 Economic Census data.

We begin this paper with an overview of the Economic Census, its past data collection procedures, and data collection plans for the 2017 Economic Census that have implications for calculating and using the new response metrics. We then define the key new response metrics for the 2017 Economic Census. We illustrate the calculations of these new metrics using 2012 Economic Census data and discuss their interpretation. We end with some suggestions for how these metrics might be used in conjunction with one another, along with the familiar "check-in rate," to evaluate data quality during and after data collection and processing, and consider next steps in our investigation.

2. Overview of Economic Census

The Economic Census is a mandatory, self-administered survey of U.S. business establishments in the eight major business trade areas: manufacturing, construction, mining, retail, services, wholesale, finance-insurance-real estate (FIRE), and utilities-transportation. Generally collected at the establishment level, past economic censuses featured multi-mode data collection strategies. Self-administered paper or electronic forms were tailored by industry, resulting in more than 600 different form versions for the 2012 Economic Census. Paper forms were distributed and returned by mail. Electronic reporting

options for the 2012 Economic Census consisted of a downloadable software, primarily for large multi-unit companies (MUs), as well as a Web-based instrument, marketed primarily to small single unit businesses.

Although response to the Economic Census is required by law, the Census Bureau employs an intensive contact strategy in order to maintain high response rates. The core contact strategy consists of initial mail-out and up to four nonresponse follow-up mailings. In addition, a number of other approaches are employed for reducing nonresponse. For example, businesses are permitted to request due date extensions. Additionally, the largest MUs in each industry are assigned "Account Managers," who are Census Bureau subject area specialists tasked with proactively making contact with respondents and building rapport to aid response from these critical cases. The 2012 Economic Census also featured targeted telephone follow-up based on propensity models using historical response/nonresponse characteristics and paradata. In past censuses, the final follow-up letter to selected delinquent, but critical, cases consisted of a stern notice from the General Counsel of the Department of Commerce emphasizing the mandatory authority under which the Economic Census is collected.

Plans for the 2017 Economic Census call exclusively for Web data collection. Multi-unit companies and single unit establishments will receive an initial mail invitation, providing a URL and their own unique online authentication information, and they will enter the electronic instrument by way of a secured portal. While many previously used nonresponse follow-up strategies will be implemented in the 2017 Economic Census, an intensive research program based on randomized experiments helped to identify and evaluate various contact strategies for their cost-effectiveness, including implementation of adaptive design techniques. Assuring the effectiveness of data-driven, adaptive nonresponse follow-up strategies motivated the development of improved response metrics associated with data quality, as described in this paper.

3. Check-in Rate

The Check-in Rate traditionally used by the Economic Census is defined as follows:

$$Check\text{–}in\ Rate = \frac{Number\ of\ Estabs.\ that\ Returned\ a\ Form}{Number\ of\ Establishments\ "Mailed"\ a\ Form} \times 100$$

It simply measures the proportion of forms mailed out that were returned, including both electronic and paper form responses. The check-in rate has been used in the past as an indicator of response in the Economic Census. However, it does not meet the OMB and Census Bureau standards and does not provide information on the quality of the response. An establishment could return a form with no useable information and be considered valid for the check-in rate.

4. Unit Response Rate

One characteristic of business surveys is their highly skewed target populations, where there are many, many small businesses and few large businesses, as measured by their employment or revenues. Data users are often interested in estimates of totals, e.g., total

revenues, and typically a relatively small number of businesses make up the majority of a total. To avoid the over-representation of small cases, which are common to business surveys, the URR in a business survey setting is often computed without using design weights. The URR formula employed by economic surveys and censuses at the Census Bureau is

$$URR = \frac{R}{E + U} \times 100,$$

where
R is the count of reporting units selected for the sample that were eligible for data collection and classified as a response,
E is the count of reporting units selected for the sample that were eligible for data collection, and
U is the count of reporting units selected for the sample for which eligibility could not be determined.

In other words, the URR represents the proportion of reporting units eligible for data collection that provided a valid response. To define the URR for the Economic Census, we need to define a reporting unit, a unit eligible for data collection, and valid response criteria. A reporting unit is an entity from which data are collected.

In general, for the Economic Census the reporting unit is the establishment. An establishment is defined as a business or industrial unit at a single, physical location that produces or distributes goods or that performs services and for which separate records are kept. A reporting unit will be counted in the sector in which it is ultimately published.

Only reporting units eligible for data collection or of unknown eligibility are included in calculations for the URR. Generally, eligibility status will be consistent with the rules for other business survey programs conducted by the Census Bureau. For instance, reporting units, including births and deaths, active at any point during the reference period will be eligible, and any duplicate reporting unit or reporting unit added in error will be considered ineligible.

To be considered a valid response, a reporting unit must provide sufficient information in its response. For the Economic Census, an establishment that returns a valid response for total receipts will have provided sufficient information. If an establishment meets this criteria and is a reporting unit eligible for data collection, or is of unknown eligibility, then the establishment will be included in the URR calculation.

5. Proxy Unit Response Rate

For the Economic Census, the eligibility of a record will not be known until collection and processing are completed. Thus, the URR cannot be calculated in real-time during collection. To adapt collection strategies in real-time, a new metric that can act as an indicator, or "proxy", for URR needed to be developed.

The Proxy URR uses only real-time data available during collection. The Proxy URR uses the following formula derived from the URR formula,

$$Proxy\ URR = \frac{Number\ of\ Estabs.\ that\ Returned\ a\ Form\ with\ Sufficient\ Information}{Number\ of\ Establishments\ "Mailed"\ a\ Form} \times 100,$$

Due to incomplete information during real-time collection there is a slight difference in the denominators of the URR and the Proxy URR. The URR's definition of a valid response cannot be used for the Proxy URR as the validity isn't completely determined until after processing; likewise regarding eligibility. For the Proxy URR, an establishment will be considered to have provided sufficient information if it has returned a form with a receipt total greater than or equal to 0 or a tax-exempt revenue greater than or equal to 0. If an establishment meets this criterion and has been deemed "checked-in", then it will be included in the calculation for the Proxy URR.

6. Comparison of Check-in, URR, and Proxy URR

Before analyzing the values of the check-in rate, URR, and Proxy URR from the 2012 Economic Census, the differences between their criteria must be understood. Both the check-in rate and the Proxy URR are real-time calculations and have the complete mailed universe as their denominator. URR will use a subset of this universe once eligibility status is determined.

For an establishment to be considered "checked-in", it only needs to return the Economic Census form. The Proxy URR and URR have an additional requirement that the establishment include sufficient information on a returned form to be considered a reporting unit.

Table 1 shows that the overall check-in rate for the 2012 Economic Census was about five percentage points higher than the URR, which are 83.71% and 78.72%, respectively. The ineligible establishments that returned a form and the establishments that did not provide sufficient information account for this difference. The URR is about one percentage point higher than the Proxy URR, 77.78%, for the 2012 Economic Census. The difference here is more complicated as the status of an establishment's eligibility is not determined until post-processing. Also, a unit's response can change during post-collection processing.

Table 1 also shows these unit response metrics for industrial sectors at the 2-digit NAICS level, using data from the 2012 Economic Census. These response metrics were calculated based on final data after the completion of post-collection processing, drawn from 2016 databases for the 2012 reference period. Thus, industry-level Proxy URR figures appearing in Table 1 may be considered to be somewhat artificial. While their denominators, the mailed universe, remain constant, response status contributing to the numerator is based on data at the close of data processing, rather than data at the close of data collection. In addition, there are numerous other factors related to post-collection processing that contribute to differences between the URR and the Proxy URR.

Table 1: Unit Response Metrics from the 2012 Economic Census[1], By Industrial Sector
(2-digit NAICS-level)

Sector	Check-in Rate	Proxy URR	URR
All	**83.71%**	**77.78%**	**78.72%**
Mining	81.39%	75.44%	68.83%
Utilities	89.39%	84.63%	79.18%
Construction[2]	78.02%	74.26%	N/A
Manufacturing	79.10%	75.79%	75.02%
Wholesale Trade	81.43%	74.38%	73.67%
Retail Trade	87.88%	82.63%	87.52%
Transportation and Warehouse	79.65%	74.13%	68.83%
Information	81.76%	74.85%	74.76%
Finance and Insurance	87.67%	81.50%	80.01%
Real Estate and Rental and Leasing	85.10%	79.24%	78.43%
Professional, Scientific and Technical Services	84.07%	78.85%	77.36%
Management of Companies and Enterprises	84.84%	77.32%	74.73%
Administrative and Support, Waste Management and Remediation Services	81.10%	73.34%	71.83%
Educational Services	79.43%	73.24%	70.42%
Health Care and Social Assistance	85.76%	78.79%	78.49%
Arts, Entertainment and Recreation	80.86%	76.96%	70.37%
Accommodation and Food Services	81.39%	74.07%	77.25%
Other Services (except Public Administration)	83.20%	78.34%	74.88%

[1]Data taken from 2016 versions of the Business Register and associated databases, with reference period of 2012.
[2]The 2012 Economic Census used a sample survey for the Construction Sector, and was processed differently than all other sectors shown. The URR formula created for implementation in the 2017 Economic Census does not account for these differences. Thus, a URR cannot be produced for the Construction Sector using 2012 Economic Census data.
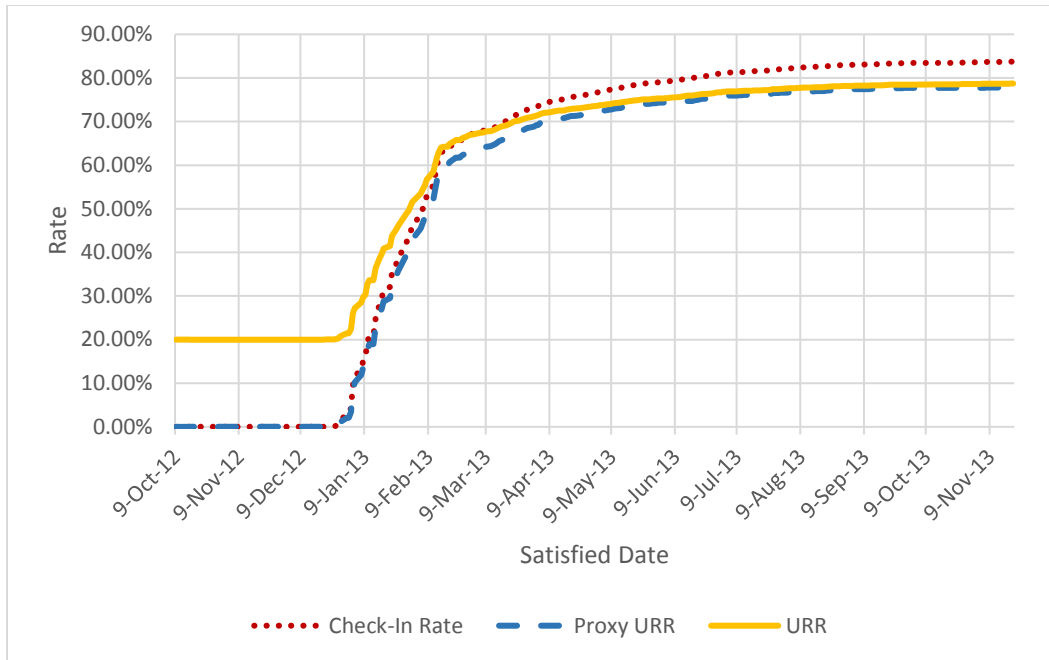
Figure 1: 2012 Economic Census Response Metrics by Satisfied Date

Figure 1 shows how the check-in rate, the Proxy URR and the URR vary over time during data collection. The x-axis, called "Satisfied Date," corresponds to when the record was added to the Business Register. Since we are trying to compare the URR with the real-time metrics, check-in rate and Proxy URR, for this chart we make the assumption that we have perfect information on the eligibility status and response information of an establishment as soon as it is received. This means that the data used to calculate the check-in rate and Proxy URR are from the end of 2012 Economic Census processing. Based on the data available from the Business Register, some records that contribute to the URR did not have a value for Satisfied Date. These records cause the URR to start at 20% at the beginning of the timeline in order for the rate at the end of the timeline to be accurate. We see that the Proxy URR and the URR are lower than the check-in rate, but mirror the check-in rate's changes over time.

7. Item Response Rates

As noted earlier, due to highly skewed target populations in business surveys, a relatively small number of tabulation units account for the majority of estimates of total values for key variables. It follows that one must weigh the cost-error tradeoffs of obtaining reported data from the large number of establishments making up a small percentage of the estimate of a total value.

An alternative is the use of administrative data in place of respondent data for survey units whose values are expected to contribute little to the estimated total, where administrative data are obtained from an administrative source with the same reference period as the survey. Although another alternative is to take a sample of survey units, and use weighted estimates to represent the population of interest, the Census Bureau's access to administrative data is sufficient, and the cost of additional data collection is not warranted. The item response rates employed by economic survey programs at the Census Bureau, the TQRR and QRR, reflect the unique properties of business surveys. A tabulation unit is a

unit that is used in estimation. For most business surveys, the tabulation unit is the same as the reporting unit.

The TQRR is defined as

$$TQRR = \frac{\sum_{i=1}^{N}(q_i + r_i)w_i t_i}{\sum_{i=1}^{N} w_i\, t_i} \times 100,$$

where
w_i is the design weight of tabulation unit i,
r_i is the indicator variable for reported data for tabulation unit i,
q_i is the indicator variable of data from an alternative source deemed to be of sufficient quality for tabulation unit i,
t_i is the data value for unit i, and
N is the total number of eligible tabulation units.

TQRR is the proportion of the weighted total of a data item directly reported by a respondent plus data obtained from an alternative source deemed to be of sufficient quality. Note that the numerator and denominator include a weighting adjustment factor, w_i, which equals one if the unit was selected with certainty. Note also that $q_i + r_i \leq 1$, because q_i and r_i are mutually exclusive. A TQRR is typically calculated for each of a survey's key data items. Three key data items were identified for the Economic Census: receipts, payroll, and number of employees.

The above formula makes the distinction between reported data and alternative source data. Alternative source data typically come from one of three sources: a survey or census with the same reference period; administrative records; or some other validated source, such as company annual reports, trade association statistics, or Securities and Exchange Commission filings. Administrative data is the primary source of alternative source data for the Economic Census. Administrative data on receipts, payroll, and number of employees are maintained on the BR and updated regularly from sources such as federal tax records.

A derivative of the TQRR that is also used is the QRR, which is defined as

$$QRR = \frac{\sum_{i=1}^{N} w_i r_i t_i}{\sum_{i=1}^{N} w_i t_i} \times 100.$$

The QRR represents the proportion of the weighted total of a data item from data directly reported by respondents.

For the Economic Census, the tabulation unit is the establishment. At the end of the survey processing cycle and prior to estimation, reporting unit estimates are applied to the establishment level, and an establishment is tabulated in the industry in which it is ultimately published.

At the estimation stage, a given data item is considered reported if the corresponding reporting unit provided a valid response for the item of interest. For example, if a parent record provided a valid response for payroll, then the corresponding children will have a status flag indicating that payroll was reported.

Two other metrics will be calculated for future economic censuses: administrative data rate (ADR) and imputation rate (IR). The ADR varies from the TQRR in that it only accounts for the data obtained from administrative records, excluding data obtained from other alternative sources and data received by direct response. The formula for ADR is defined as

$$ADR = \frac{\sum_{i=1}^{N} w_i d_i t_i}{\sum_{i=1}^{N} w_i t_i} \times 100$$

where

d_i is the indicator variable for administrative data for tabulation unit i and item t.

The imputation rate is the percentage of the estimated total that is not obtained from directly reported or alternative source data or more simply,

$$IR = 100 - TQRR.$$

For the 2012 Economic Census, all alternative source data are considered administrative data because the indicators for data from other alternative sources were not in place at the time. This causes the ADR to simply be the difference between the TQRR and QRR. Table 2 shows that TQRR for all three key items are greater than 85%. Larger portions of the TQRRs for key items payroll and employment are from administrative data than for receipts, because these two items are more easily found in administrative data such as federal tax records.

Table 2: Item-level Response Metrics for the 2012 Economic Census

Key Item Variable	TQRR	QRR	ADR[1]	IR
Receipts	93.46%	87.54%	5.92%	6.54%
Payroll	87.71%	76.45%	11.26%	12.29%
Employment	86.23%	71.39%	14.84%	13.77%

[1]ADR is the difference between TQRR and QRR, because flags indicating data from alternative sources other than administrative data were not available for the 2012 Economic Census.

8. Proxy TQRR and Proxy QRR

Similar to URR, the item-level response metrics cannot be calculated in real-time during data collection. Many trade areas rely on knowledge of their current coverage of these key variables, most commonly payroll, during collection to help inform decisions whether to adjust collection strategies. Thus, proxies for both the TQRR and the QRR for the key item payroll were developed to be used in real-time. The formulas for Proxy TQRR and Proxy QRR are the same as their counterparts but with the variables defined using only real-time information. Most variables use a corresponding real-time value found on the Business Register, but a few needed alternate definitions. The response indicator variable, r_i, from the previously defined TQRR and QRR formulas, uses the same sufficient information condition used for Proxy URR to identify valid responses. Valid responses are returned forms with receipt totals greater than or equal to 0 or tax-exempt revenues greater than or

equal to 0. The universe for these two proxy metrics is not all tabulation units, but all mailed establishments in addition to a subset of establishments that were predetermined to be non-mail units. These units have possible administrative data to use for the calculations. The indicator for alternative source data, q_i, now represents this subset of establishments.

Table 3 shows that the 2012 Economic Census payroll TQRR and its proxy have a difference of about 6.58 percentage points while the QRR and its proxy have a difference of 2.23 percentage points.

Table 3: Comparison of Item-level Response Metrics and Their Proxy Values for the 2012 Economic Census

Key Item Variable	TQRR	Proxy TQRR	QRR	Proxy QRR
Payroll	87.71%	81.13%	76.45%	74.22%

9. Conclusion

The Economic Census produces the most detailed benchmark data available about the U.S. economy. It is not like other survey programs with respect to design or magnitude of information collected. In 2012, the Economic Census produced and used a check-in rate to measure the status of the data collection process. However, this check-in rate overestimates the response rate and does not meet OMB and Census Bureau standards. The URR developed in this paper more accurately calculates the true response rate and follows both OMB and Census Bureau standards. The drawback of the URR is that it cannot be calculated in real-time during data collection to help make decisions whether to adjust collection strategies. The Proxy URR acts as an indicator of the URR and can be calculated in real-time to be used to monitor and manage data collection processes and aid decisions about adjusting collection procedures.

In addition to the unit response metrics, we developed item-level response metrics for three key items: receipts, payroll, and employment. These item response metrics are TQRR, QRR, ADR, and IR. TQRR calculates the proportion of a key item's estimated total obtained from response data or from alternative sources considered to be of sufficient quality, in particular, administrative data. The QRR and ADR breaks down the TQRR into the proportion of the estimate received from response data and specifically from administrative data, respectively. The IR calculates the percentage of the total not obtained from these sources. These four metrics provide better indicators of the quality of the data and resulting statistical estimates than what had been available in previous years.

With increasing costs for the Economic Census and other surveys, it is necessary to use new metrics such as those presented here to gain a better understanding of the data we collect and the data collection process. With this understanding, we can make more informed decisions about collection strategies and adapt those methods in real-time while maintaining high quality data.

10. Future work

The new metrics presented in this paper can lay the foundation for new collection strategies. In previous Economic Census, the collection mode has been a mix of paper and

electric collection modes. For the 2017 Economic Census, we need to understand how these metrics behave with a move to a completely electronic collection mode.

These new metrics, and more specifically their proxies, provide the tools for real-time metrics of data quality. With these tools, we can adjust collection strategies and adapt to the landscape of a changing collection environment. Future work will be needed to fully understand how to use tools, such as the Proxy URR, Proxy TQRR, and Proxy QRR, to adapt collection strategies and understand the impact of those possible decisions.

References

Fink, E.B. and Lineback, J.F. (2015) "An Update on Developing Response Metrics for the Economic Census," presentation at the 2015 FedCASIC Workshop, U.S. Census Bureau, Suitland, MD, March 4-5.

Lineback, J.F., Oliver, B., and Willimack, D.K. (2012) "Developing Response Metrics for the Economic Census," *Proceedings of the 2012 Federal Committee on Statistical Methodology Research Conference*, U.S. Office of Management and Budget, Washington D.C.

U.S. Census Bureau (2010). Statistical Quality Standards. http://www.census.gov/quality/standards/index.html.

U.S. Office of Management and Budget (2006) Standards and Guidelines for Statistical Surveys. https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf