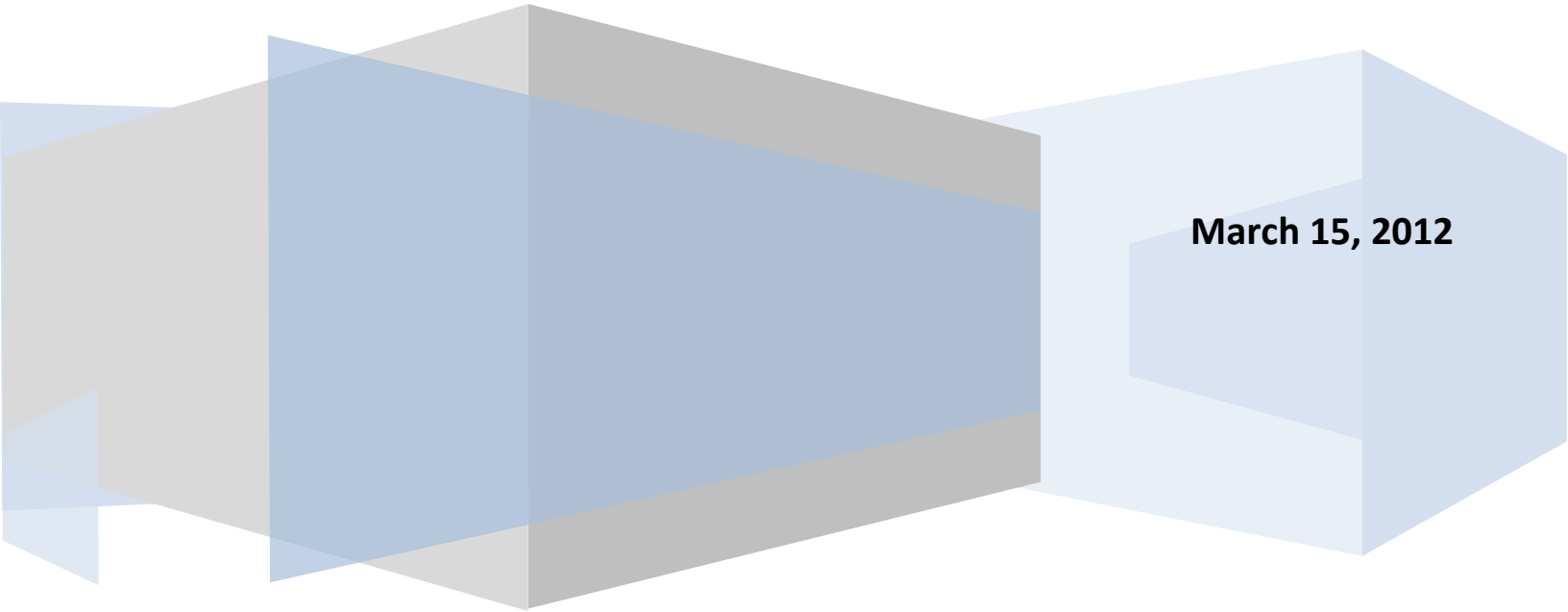
Censuses and Surveys of Governments: A Workshop on the Research and Methodology behind the Estimates

Using Paradata to Improve Questionnaire Design and Operations

Aneesah Williams



March 15, 2012

Using Paradata to Improve Questionnaire Design and Operations

Aneesah Williams
Governments Division
U.S. Census Bureau

Introduction

Paradata are data about the process by which survey data are collected. Due to the emergence of computer assisted surveys, paradata are now able to be captured through key-strokes of respondents, time stamps of movement, and navigation patterns throughout the questionnaire. For surveys administered by a field representative, there are also paradata captured through interviewer observations, vocal properties of interviewer and respondents, and interviewing strategies (Kreuter, 2010). We can extend the definition of paradata to include any evaluation that we learn from and use to revise our process. In the Governments (GOVS) Division, we plan to use several sources of paradata to improve our questionnaires and to improve our data processing. This paper will give a brief description of the various types of paradata and the benefits to be gained from applying the information received from these paradata into our survey processing.

The types of paradata that GOVS will obtain are web-based paradata, nonresponse follow-up (NRFU) dashboards, information from new, standardized editing and imputation flags, information from edit evaluations and nonresponse bias studies, and representativity indicators. Information from the first two forms of paradata are currently being collected, with the remainder to be obtained in conjunction with the 2012 Census of Governments (CoG). The 2012 CoG covers three major subject fields: government organization, public employment, and government finances. The Government Units Survey (GUS), which is part of the Organization Component, collects basic information about the functions and activities of governments and provides the universe of local governments for the remaining components of the 2012 CoG. The other two components will include censuses of Public Employment & Payroll, Government Finances, and Public Pensions.

Paradata discussed in this paper will allow GOVS to gather important information about the way our data are collected, to analyze and identify problem areas, and to make modifications that will enable us to improve our CoG survey operations.

Web-based Paradata

The Government Units Survey (GUS) is currently being collected using the Centurion web instrument. This software allows us to collect paradata as respondents are completing their

---

Disclaimer: This report is released to inform interested parties of research and to encourage discussion of work in progress. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

questionnaires. It provides time stamps for each click made during the collection process. The information we receive includes, but is not limited to, the unit identifier, the type of action taken, the time each action is taken, the screen on which the action was taken, the web address if a link was selected, each field (question) name, and the value input in each field (question).

Because of the amount of information produced, it was advisable to have a few questions in mind before studying the paradata. Some of the things we are interested in finding out are how often a response is changed (per question), what are the most frequent break off points in the questionnaire, how much time is being spent on each question, and whether the respondents are changing/deleting their contact information. Such information may pinpoint questions that need to be revised or clarified.

New Editing and Imputation Flags

In GOVS, a universal editing and imputation flag system is being implemented to provide information about our data. These flags characterize the response of the government unit and each data item. The collection unit source code (CUSC) identifies the initial source of the data (i.e., paper form response, web submitted response, etc.) for a government unit. The item edit/imputation code (EIC) identifies reported, edited, and imputed data at the item level (e.g., respondent reported, keying error correction, cell mean impute, etc.). The item source code (ISC) identifies the source of the data at the item level. It is often, but not always, the same as the CUSC. The item source reference code is 250 character field that identifies any characteristics the analyst deems noteworthy.

Previously, each survey used its own set of flags and even those with similar meanings were assigned differently across the various surveys. With the standardization of the flags, however, we will be able to compute response rates and other metrics more accurately, to identify points in the collection process that are ineffective, to determine which questions are imputed most often, and to pinpoint questions that cause the most edit failures. This information will help GOVS to improve our operations and produce better quality data.

The standardized flags will allow GOVS to capture information about the data collection process that were not previously attainable. For instance, suppose a respondent needs to gather financial information from several departments in order to complete the CoG: Government Finance component. This respondent may initially submit partial data on the revenues of the government using the web collection instrument, but complete the debt sections of the questionnaire via fax some time later. Now, the CUSC and ISC flags can be used in combination, to determine the frequency of surveys being submitted using more than one media, as in the example.

The new EIC flags will allow statisticians to determine how often a variable is a legitimate blank, as opposed to a nonresponse. We will have the ability to identify variables that are not applicable to a given unit or type of government. Information obtained from these paradata will enable us to edit our data more efficiently. We can avoid imputing values in variables where

they do not apply within a government unit. We will be able to determine such editing problems as respondents reporting a value in the wrong unit of measure (i.e., reporting in dollars instead of thousands). This type of information can assist in developing questions, writing questionnaire instructions, and improving the editing procedures.

Edit Evaluations

SMB is currently conducting an edit evaluation to investigate the effects of the current edits in place using 2011 Annual Survey of Public Employment & Payroll data. This evaluation can help identify areas where the edits are not performing well. We will be able to determine which edits are failing too many cases, as well as which edits are not identifying enough units with erroneous data. Improvements in the edit parameters and edit bounds will lead to more efficient processing of the data and less burden for the survey analysts.

By studying the frequency of edit failures, we will be able to identify potentially faulty questions. Such questions will stand out most if the imputation rate is also high. An edit evaluation is planned for the Annual Survey of State and Local Government Finances and the Annual Survey of Public-Employee Retirement Systems as well. These evaluations will be repeated continually through the 2012 CoG, providing valuable paradata to improve our editing operations.

Nonresponse Bias Analyses

A nonresponse bias analysis will be conducted on each of the 2011 annual surveys as well as each of the 2012 CoG components. These analyses will examine the potential for bias in estimates that can be attributed to nonresponse. This will help to determine how much bias is introduced in the surveys when the respondents differ from the non-respondents. Such information can be used to make adjustments to the weighting of units in intercensal years or to redesign the sample for one or more of the 2012 CoG components.

Representativity Indicators

Recently, we have considered using representativity indicators (R-indicators) to describe the quality of our surveys. In the world of survey methodology, most agree that response rates by themselves are not sufficient indicators of nonresponse bias. Thus, R-indicators are now being used to evaluate how similar a response data set is to the sample data set (Representativity Indicators for Survey Quality). If there is much variation in response probabilities, then the response data set is weakly representative. On the other hand, if all response probabilities are equal, the data set is considered strongly representative.

The R-indicators are often used in conjunction with an auxiliary variable in the survey. R-indicators can be used to ensure that a representative sample is obtained and to gather detailed information about auxiliary variables, which can then be used in adjustment weighting. We are

continuing to research other ways that this indicator can be used to obtain a more representative sample, thus reducing bias and improving our surveys overall.

Dashboards

Dashboards are being developed for each component of the 2012 CoG to guide nonresponse follow-up (NRFU). They will be used as an aide for survey analysts to determine where NRFU is needed. These dashboards monitor the check-in rates during the collection cycle of each survey. This includes unit check-in rates at national and state-by-type levels, as well as at the function level for special districts.

Another important feature of the dashboard is that it provides the number of respondents by imputation cell. This information is tracked to ensure that there are ample respondents for the imputation process. This helps to identify units that need to be followed up. The dashboards will also be used as a tool to ensure that the Census Bureau statistical quality standards are met (Craig, 2012).

Summary

Each of the forms of paradata being collected for the 2012 CoG will be useful for improving the quality of our surveys. They will also enable us to improve our operations by making them more efficient and yielding higher quality data. Paradata such as the dashboards and the standardized flags will allow us to see where our data collection processes need to be more effective. The web-based paradata and edit evaluations can point us to items on the questionnaires that may need revision. The nonresponse bias analyses and R-indicators will lead us to construct a better sample. Using these paradata will also ensure that our data products from the 2012 CoG meet the Census Bureau's statistical quality standards. The paradata collected during this census year will allow GOVS to set a baseline, improve the quality of our data and the efficiency of our processes, and ensure that we continue to meet the missions of the U.S. Census Bureau.

References

Craig, T. L. (2012). *The Implementation of Dashboards in Governments Division Surveys.*

Kreuter, F. a.-C. (2010). *Paradata.* Berlin: Working Paper Series of the Council for Social and Economic Data.

Representativity Indicators for Survey Quality. (n.d.). Retrieved January 30, 2012, from http://www.risq-project.eu.indicators.html