

What Can Web Paradata Tell Us About 2012 Economic Census Business Reporting?

Elizabeth M. Hoeffel, Joanna Fane Lineback, and Gianna Dusch

U.S. Census Bureau¹, Washington, DC 20233

Abstract

Survey paradata are increasingly used to understand respondent behavior and survey outcomes. In this paper, we analyze 2012 Economic Census electronic paradata to examine Web-based reporting patterns of single-establishment businesses. These paradata include the timing and frequency of user actions as well as systems usage information. Also incorporated into our analysis are respondent business characteristics. These data are available from the U.S. Census Bureau's Business Register. We use our findings to explore ideas for using these paradata to influence data collection activities.

Keywords and phrases: Paradata, Economic Census, Business Survey, Electronic Data Collection, Web Surveys

1. Introduction

"Paradata" refers to the process information that comes as a by-product of data collection (Couper, 1998).

In this paper, we analyze the electronic reporting behaviors of 2012 Economic Census single-establishment businesses. For the analysis we use paradata from the Web collection instrument and business characteristics data from the U.S. Census Bureau's Business Register, which contains information such as business sector. Here, we discuss the motivation for this work.

The reality of business and establishment survey programs is that there are shrinking budgets and declining response rates. In addition to reducing costs, electronic collection has the added benefit of providing information about the data collection process. These paradata are obtained without placing additional burden on respondents.

Kreuter (2013) notes that the use of paradata has recently grown because of an increase in computer-aided data collection, paradata availability, and interest among survey sponsors in process quality.

Paradata have been shown to be useful for various purposes, particularly in household surveys of individuals. Although other areas at the Census Bureau have conducted large-scale paradata analyses, such as American Community Survey research by Horwitz et al. (2013) and internal public-sector surveys analyses, this is a first look at paradata for the economic census. The primary purpose of this research was to explore what paradata are available and what they might tell us about business respondent behaviors and reporting instrument effectiveness.

Census Bureau business and establishment survey data are primarily collected using two self-administered modes: mail-out/mail-back paper and electronic. Electronic collection is through one of two custom-built Census Bureau software products, depending on the size of the business. For multi-establishment businesses and large single-establishment businesses, data are collected through the downloadable software Surveyor. For other single-establishment businesses, data are collected through the Web application Centurion. There are other collection modes, such as telephone, for a small portion of cases.

The Census Bureau has been using electronic reporting in some form to collect data from businesses for decades. Previously, data quality was often determined by examining edit failures. The economic census program plans to use paradata in the future to better understand data collection and make decisions during the survey life cycle. Paradata

¹ *This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, or technical issues are those of the authors and not necessarily those of the U.S. Census Bureau.*

provide the opportunity to examine respondent behavior, instrument performance, evidence-based adaptive design decisions, and a more in depth look at data quality. At the time of this writing, November 2013, the 2012 Economic Census is still in the field, and data collection is expected to end later this month. Therefore, we are in the process of collecting and analyzing paradata from our electronic collection instruments.

These paradata include the timing and frequency of user actions such as log-in/log-out and submission-related activities. Additionally, we are able to combine this information with business characteristics such as business sector and type of survey form.

This paper is organized as follows: Section 2 provides background information on the economic census. Section 3 discusses analysis methods, instrument navigation for a typical respondent, and limitations. Section 4 provides preliminary results. Section 5 offers a discussion of these results. Finally, Section 6 discusses future research.

2. Background

This section provides a high-level overview of the economic census, focusing on the data collection process.

The Census Bureau conducts the economic census every five years, covering reference years that end in 2 or 7. The economic census is a 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. Data collected on the economic census include employment labor costs and output, as well as assets, expenditures, inventory, and other industry-specific items. Data are generally collected at the establishment level. Approximately 4.7 million establishments received the 2007 Economic Census, and approximately 4 million establishments received the 2012 Economic Census.

Because economic census forms are tailored by industry, there are more than 500 different questionnaire versions. Paper forms are distributed and returned by mail. Various options to report electronically have been available since the 1987 Economic Census, primarily for large multi-establishment companies. While Surveyor has been available since 2002 to all establishments eligible for the economic census, the 2012 Economic Census is the first to offer direct internet, or Web-based, reporting to single-establishment businesses.

Each single-establishment business receives one of the following three types of survey forms: a classification form, a “long” form, or a “short” form. The classification form collects industry classification information. Establishments for which we have insufficient identifying information receive a classification form. The long and short forms both collect information on key economic items such as employment, payroll, and employment. The long form, however, collects this information at a more detailed level. For example, where the long form asks for annual payroll by job, the short form asks for total annual payroll. Single-establishment businesses that receive the long form also receive a supplemental form, which collects information about the ownership and control of the establishment. Some long form recipients also receive a second supplemental form, which collects information on items such as foreign ownership, research and development activity, royalties, and manufacturing activities. Most reporting units receive a long form. Only small reporting units in the manufacturing and mining sectors may receive a short form.

Although response to the economic census is required by law, the Census Bureau employs an intensive contact strategy in order to maintain high return rates. Contact initially begins in the autumn of the reference year, when large multi-establishment companies are first provided with economic census forms and notified about electronic reporting procedures. The purpose of this early contact is to facilitate businesses’ planning for economic census response as the reference year ends. Mailout to the rest of the units eligible for the economic census occurs mid-December of the reference year, with a due date of February 12 the following year. There are as many as four follow-up mailings to non-respondents, the first of which begins approximately one week after the due date. Intervals between subsequent mailouts vary, but are roughly 30 days. The final mailing for the 2012 Economic Census occurred in July 2013.

For more information on the economic census, refer to the 2007 Economic Census User's Guide located at http://www.census.gov/econ/census07/www/user_guide.html.

3. Analysis

3.1 Centurion Paradata

For this research, we analyzed data from single-establishment respondents who have logged into the Web-based Centurion instrument as of July 28, 2013. The research for this paper is exploratory in nature, and a large part of this project has been understanding the available data and how we might analyze them.

There are 13 user actions captured by the Centurion software. The list below describes the actions used for the research in this paper. For each, we know the time (date, hour, minute, and second) of the event.

- A log-in failed because of an incorrect password.
- An account was locked because of 5 password failures.
- A log-in failed because the account was locked.
- There was a log-in success.
- The main menu was visited.
- A user clicked on a start button to begin a form.
- A user clicked on a resume button to continue a form.
- A user clicked on a review button to review a form.
- The review screen was visited.
- A user clicked on a link to a question from the review screen.
- There was a submission.
- There was a log-out.
- The account was disabled.

In addition to the actions above, the internet protocol (IP) address and user agent string are available. Pape and Barron (2013) explored the user agent string, and, for this research, we also analyzed the user agent string, which shows the device, operating system, and browser used.

Additional information is captured, but they are not considered at this stage of the research. These items are listed below:

- Questionnaire versions that were started, resumed, and reviewed.
- Number of errors and warnings present on the review screen.
- Items visited directly from the review screen.

3.2 Instrument Navigation

Below are the steps for a typical user with 2 forms (i.e., the long form and the ownership or control supplement).

Before logging into the electronic reporting instrument, a user must first log-in to the Business Help Site (BHS) using the same log-in id and password provided for the economic census. As the name suggests, the BHS provides survey completion instructions and assistance for business survey respondents along with access to data publications. From here, a user logs into the 2012 Economic Census using the assigned id and password and is taken to the main menu (see Figure 1).

From the main menu, the user pushes the start button and goes through the questions. For each question, there is a button to return to the main menu. If this happens before getting to the review screen, the button would say "resume." After getting to the review screen, the button would say "review." From the main menu, the user is taken to the survey questions.

Figure 1. Main Menu Screen- 1st Visit

2012 Economic Census: Dashboard - Windows Internet Explorer

https://ecapp-centserv.centurion-qa.ssd.census.gov/econ2012/dashboard

File Edit View Favorites Tools Help

2012 Economic Census: Dashboard

2012 Economic Census

The foundation for measuring business activity across the U.S. economy

United States Census Bureau Economic Statistics

Logout Instructions Contact Us Help

Main Menu

Please complete all forms listed below. You will have an opportunity to review your responses at the end of each form.

Company Address

ECON SINGLE

4700 SILVER HILL ROAD
SUITLAND, MD 207640000

- Below are your forms for the 2012 Economic Census. You will be able to submit your data to the Census Bureau once all your forms are completed without errors.
- Note: You can only submit your data once a day.

Submit Forms

| Action | Status | Form | OMB | Due Date | Problem Pages | View/Print PDF |
|-----------------------|-------------|--------------------------------|--|---------------|----------------------------------|----------------|
| Start | Not Started | MC-31106(L): Confectionery | No.: 0607-0938 Expiration: 12/31/2013 | Feb. 12, 2013 | 0 With Errors 0 With Warnings | |
| Start | Not Started | NC-99530: Ownership or Control | No.: 0607-0938 Expiration: 2013-12-31 | Feb. 12, 2013 | 0 With Errors 0 With Warnings | |

Done

Trusted sites

100%

Once the user gets through all of the questions for the first form, she is taken to the review screen (see Figure 2). This screen shows any errors or warnings. The three red boxes in Figure 2 show that there are three errors. Clicking on the link to the left of the screen returns the user to that question. From this screen, the user can go back to the main menu by clicking the “main menu” button at the bottom of the screen.

Figure 2. Review Screen

If there is another form to complete, the user goes through the same process again, clicking the start button for that form (see figure 3).

Figure 3. Main Menu Screen- 2nd Visit

Once the user has fixed any errors or warnings and is done with the final form, she clicks on the main menu button again from the review screen.

This takes the user back to the main menu screen, in this case, for the third time (see figure 4). Once all forms are complete, the “Submit Forms” button is activated, turning from grey to black, and the user clicks the button to submit the forms. Also, if the user wants to log out at any point, she can click the “Logout” button on any screen.

Figure 4. Main Menu Screen- 3rd Visit

2012 Economic Census
The foundation for measuring business activity across the U.S. economy

United States Census Bureau
Economic Statistics

Logout Instructions Contact Us Help

Main Menu
Please complete all forms listed below. You will have an opportunity to review your responses at the end of each form.

Company Address
ECON SINGLE
4700 SILVER HILL ROAD
SUITLAND, MD 207640000

- Below are your forms for the 2012 Economic Census. You will be able to submit your data to the Census Bureau once all your forms are completed without errors.
- Note: You can only submit your data once a day.

Submit Forms

| Action | Status | Form | OMB | Due Date | Problem Pages | View/Print PDF |
|--------|-----------------------------------|--------------------------------|--|---------------|----------------------------------|----------------|
| Review | Completed 10/29/2013, 17:03:54 | MC-31106(L): Confectionery | No.: 0607-0938 Expiration: 12/31/2013 | Feb. 12, 2013 | 0 With Errors 0 With Warnings | |
| Review | Completed 10/29/2013, 17:03:18 | NC-99530: Ownership or Control | No.: 0607-0938 Expiration: 2013-12-31 | Feb. 12, 2013 | 0 With Errors 0 With Warnings | |

To submit one form a user must, at a minimum, complete the steps below:

1. Log-in successfully
2. Visit the main menu
3. Click the start button
4. Visit the review screen
5. Return to the main menu screen
6. Submit the data

Note that to submit one form the user must visit the main menu at least twice, to submit two forms she must visit the main menu at least three times, and to submit three forms (the maximum number of forms a reporting unit would receive) she must visit the main menu at least four times.

3.3 Analysis Methods

Our analysis of the 2012 Economic Census web paradata is broken into three major components: frequency of user actions, timing between events, and software usage counts. First, we provide frequencies of the 13 user actions listed in Section 3.1. The overall picture may alert us to areas of concern with the data collection instrument. We break the actions into two categories: those related to the log-in events and those related to post-log-in events.

Next, we consider the time between the initial log-in and the last recorded user event. We analyzed these data by type of survey form, by number of forms started, and by sector. Because the number of respondents that logged in but did not submit forms was higher than expected, additional analysis was done to compare submitted and unsubmitted forms by type of survey form, number of forms started, and sector.

Finally, we examine the information obtained from user agents including operating system, browser, and device used to complete the questionnaire by the frequency of usage.

3.4 Limitations

There are limitations to this analysis. First, as of this writing, data are still being collected for the 2012 Economic Census, so the results reported here do not represent the final outcomes for all Centurion respondents. In other words, our results will change as we receive more data. Thus, the analysis presented here does not indicate final distributions of respondent behaviors. Nevertheless, it does demonstrate that the economic census program has the capability to collect and analyze paradata in real-time, a key requirement for using paradata effectively in adaptive design decisions. For instance, as these paradata are being updated weekly, we are able to monitor results on a weekly basis.

A second limitation is that only “server-side” paradata are being collected. Server-side paradata are captured from the Census Bureau’s server, and is somewhat limited in scope. Ideally, these paradata would contain respondents’ actions at any point in the survey, sometimes called client-side paradata. Thus, we cannot analyze user actions such as answer changes or movements within a specific screen.

Finally, we only analyzed data from the respondents who successfully logged into the online instrument. In general, only single-establishment businesses are able to use the online instrument. Additionally, we do not analyze paradata from the Surveyor software offered for multi-establishment business reporting. While there are paradata captured from Surveyor respondents, it is more limited in scope than the Centurion paradata. We also do not consider paper respondents. Therefore, a large, important percentage of the population is not covered in this initial analysis.

4. Results

This section contains preliminary results based on 738,492 single-establishment businesses that were able to log in to the Centurion software as of July 28, 2013. The results are presented in the following order: frequency of user actions, timing between events, and software usage counts.

4.1 Frequency of User Actions

The frequencies of the user actions are shown in Tables 1 and 2.

Table 1 includes actions that may indicate trouble logging into the system. For those users who entered Centurion, 7.14% entered an incorrect password at least once. In other words, over 50,000 users had trouble logging into Centurion. This is even after successfully logging into the BHS, which, as stated previously, requires the same password to log in. We suspect this number would be larger if we were able to account for the users who were never able to successfully log into the system. Additionally, if a user had five unsuccessful log-in attempts, the account was locked for 15 minutes. This occurred for 0.12% of all users. Another cause of a failed log-in was trying to log-in while the account was locked. This occurred for 0.04% of all users. Finally, one account was disabled due to a security concern.

Table 1. Rate of Log-in Event by Frequency

| Number of Actions | Event | | | |
|-------------------|-------------------------------|----------------|-------------------------------|------------------|
| | Failed Log-in: Wrong Password | Account Locked | Failed Log-in: Account Locked | Account Disabled |
| 0 | 92.86% | 99.88% | 99.96% | 100% |
| 1 | 5.50% | 0.08% | 0.02% | 0% |
| 2 | 0.90% | 0.02% | 0.01% | 0% |
| 3+ | 0.74% | 0.02% | 0.01% | 0% |

Table 2 shows the frequency of remaining user actions starting from successful log-in. Since this analysis only considered the users who successfully logged in, there were no cases included in this analysis that did not have at least one successful log-in. Of the users, 78.32% logged in once (implying that if a user submitted her forms, then the forms were completed in one sitting), 14.88% logged in twice, 4.23% logged in three times, and 2.57% logged in four or more times.

Table 2: Rate of User Action by Frequency

| Number of Actions | User Action | | | | | | | | |
|-------------------|-------------|-------------------|---------------|---------------|----------------|---------------------|----------------------------|---------------------------|------------|
| | Logged in | Main Menu Visited | Forms Started | Forms Resumed | Forms Reviewed | Review Form Visited | Review Page: Reviewed Form | Number of Times Submitted | Logged out |
| 0 | N/A | 0.02% | 0.76% | 80.04% | 79.87% | 6.09% | 84.44% | 12.39% | 46.82% |
| 1 | 78.32% | 4.38% | 28.88% | 13.64% | 12.99% | 20.50% | 8.11% | 86.69% | 46.18% |
| 2 | 14.88% | 11.26% | 63.77% | 3.95% | 3.69% | 47.85% | 3.14% | 0.87% | 5.37% |
| 3 | 4.23% | 27.83% | 6.44% | 1.38% | 1.57% | 13.08% | 1.57% | 0.04% | 1.15% |
| 4 | 1.44% | 28.88% | 0.14% | 0.54% | 0.77% | 4.80% | 0.88% | 0.00% | 0.32% |
| 5 | 0.58% | 13.28% | 0.01% | 0.23% | 0.41% | 2.39% | 0.55% | 0.00% | 0.09% |
| 6 | 0.26% | 6.20% | 0.00% | 0.11% | 0.25% | 1.56% | 0.37% | 0.00% | 0.04% |
| 7 | 0.13% | 3.18% | 0.00% | 0.06% | 0.15% | 1.02% | 0.24% | 0.00% | 0.02% |
| 8 | 0.07% | 1.78% | 0.00% | 0.03% | 0.09% | 0.67% | 0.16% | 0.00% | 0.01% |
| 9 | 0.04% | 1.08% | 0.00% | 0.02% | 0.06% | 0.48% | 0.11% | 0.00% | 0.00% |
| 10+ | 0.07% | 2.11% | 0.00% | 0.03% | 0.15% | 1.56% | 0.44% | 0.00% | 0.00% |

Additionally, 4.40% visited the main menu less than the minimum number of visits required (two visits) to submit their data. In order to submit, a user had to visit the main menu between two and four times, depending on the number of forms received. Most users visited the main menu two, three, or four times, as expected. About 27.63% of cases visited the main menu more than four times.

On the main menu page, a user can select any of the forms listed to start. There is a start button for each form. Less than 1% of users did not click on the start button at all (0.76%), 28.88% of users clicked on the start button once, 63.77% of users clicked on the start button twice, and 6.44% of users clicked the start button three times.

Once a user starts a form, the start button changes to a resume button for that form. Most users (80.04%) did not use the resume button, 13.64% used it once, and 6.35% used it two or more times. Similarly, the resume button changes

to a review button after the review screen was accessed. The results show 79.87% of users did not use the review button, while 12.99% used it once, and 7.14% used it two or more times.

When a user finishes completing all the questions for a given form, a review screen is displayed. The review screen shows if any questions have errors or warnings. A small percentage of users, 6.09%, did not reach the review screen, while 20.5% visited the review screen once, 47.85% visited it twice, and 25.56% visited it three or more times. Once on the review screen, a user can click on links to each question. Of all users, including those that did not get to the review screen, 15.56% clicked on at least one link on the review screen.

Finally, while 87.60% of users submitted their responses one or more times, about 12.39% logged in but never submitted a form. Just under half (46.82%) of users did not log out of the online instrument, while the remaining 53.18% logged out at least once.

4.2 Timing Between Events

This section addresses the timing between events. Of the 738,492 users that entered the online instrument as of July 28, 2013, shown in Table 2, 5,606 (0.76%) did not click the start button on the main menu while the remaining 99.14% started at least one form. (Users are only able to submit if all forms are complete.) Table 3 shows that among users who submitted a form, most (68.81%) started and submitted two forms, which we expected. However, among users who did not submit a form, 36.57% started two forms, while most (61.46%) started only one form.

Table 3. Percentage of Forms Started by Submission Status

| Number of Forms Started | Submitted (n=646,976) | Unsubmitted (n=85,910) |
|-------------------------|-----------------------|------------------------|
| 1 | 25.00% | 61.46% |
| 2 | 68.81% | 36.57% |
| 3 | 6.19% | 1.97% |

Table 4 shows the median time from the first completed user action to the last completed user action by submission status crossed by number of forms started. (Note that users are not necessarily continuously logged in between the first and last action.) Again, users are unable to submit until all of their forms are complete. As expected, the more forms started, the longer the amount of time spent in the instrument. The median time to start and submit one or more forms was 19 minutes and 12 seconds, while the median time for users that did not submit forms was 10 minutes and 37 seconds. Although those users that ultimately submitted their forms overall spent more time logged into the instrument, the median time for those that submitted three forms was less than the median time for those that started three forms and did not submit.

Table 4. Median Time from Initial Log-in to Last User Action by Submission Status and Number of Forms Started*

| Number of Forms Started | Submitted (n=646,976) | Unsubmitted (n=85,910) |
|-------------------------|-----------------------|------------------------|
| Overall** (n=738,492) | 19'12" | 10'37" |
| 1 Form (n=214,569) | 6'25" | 5'17" |
| 2 Forms (n=476,576) | 25'07" | 21'16" |
| 3 Forms (n=41,741) | 32'18" | 41'16" |

* Times are formatted as mm'ss".

**Includes users that did not start any forms.

Table 5 shows the median time from the initial log in to the last user action by form type. As mentioned in Section 2, a classification, or class, form collects industry classification information. A long form collects detailed information on key items such as payroll, receipts, and employment. A short form collects less detailed information, and it is only used in the manufacturing and mining sectors. As seen in the table, users spent much less time on class forms compared to short and long forms, which was expected.

Table 5. Median Time from Initial Log-in to Last User Action by Submission Status and Form Type*

| | Submitted (n=646,779) | Unsubmitted (n=91,244) |
|-----------------|-----------------------|------------------------|
| Class Form | 6'13" | 3'39" |
| Long/Short Form | 25'34" | 11'55" |

* Times are formatted as mm'ss".

Table 6 shows the median times by business sector, from the first action to the final action by submission status. The construction sector has long median times for all but one of the categories shown in the table below, while the real estate and rental and leasing sector has shorter median times. For users who started one or two forms, in general, median times between actions are longer for users that ultimately submitted. However, most users that started three forms and did not submit spent longer in the instrument than those that ultimately submitted three forms. This is consistent with the results in Table 4 above, which shows the overall median time for those that started and submitted three forms is less than those that started three forms and did not submit.

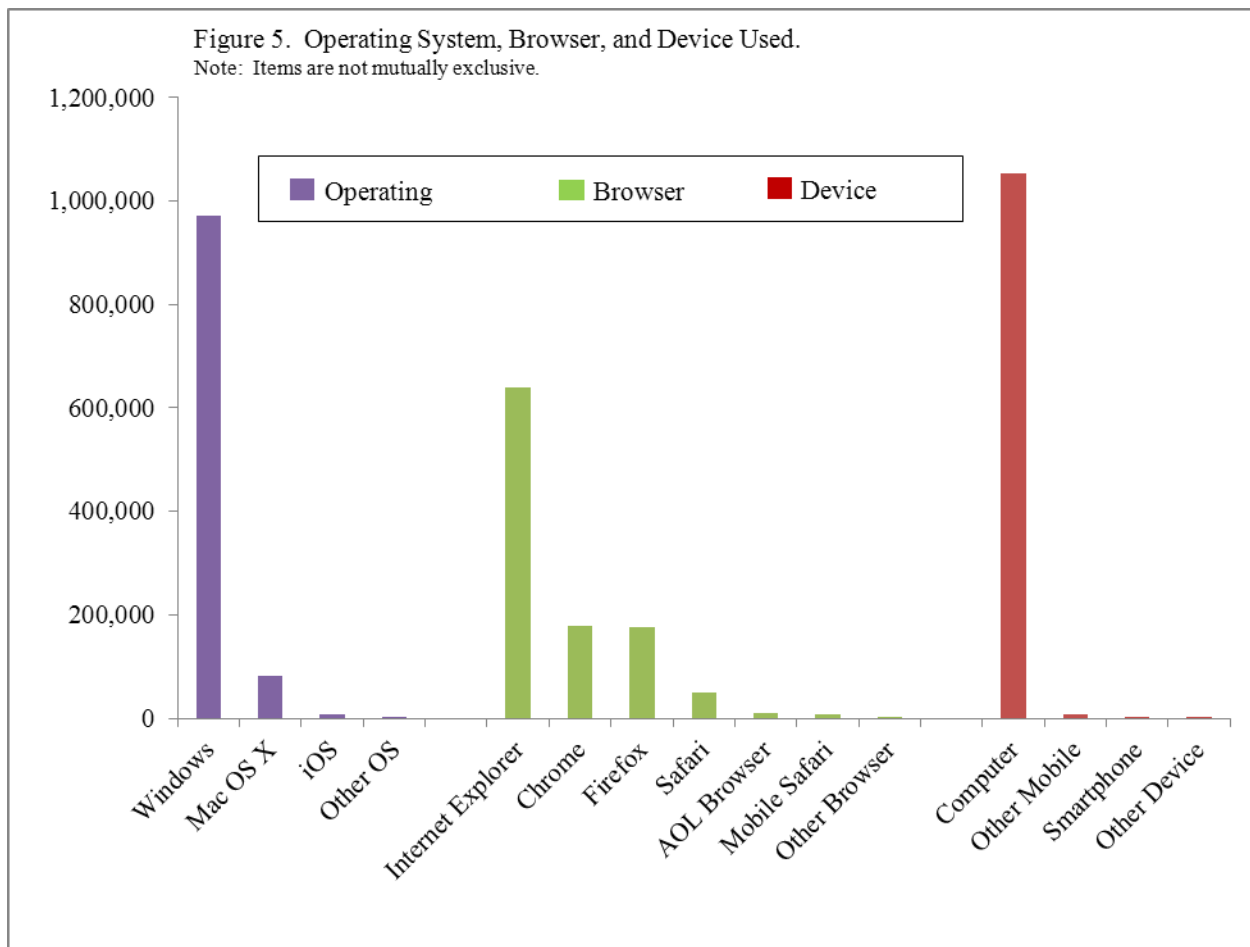
Table 6. Median Elapsed Time from Initial Log-in to Last User Action for Each Sector by Submission Status and Number of Forms Started*

| Sector | Started 1 Form | | Started 2 Forms | | Started 3 Forms | |
|--|--------------------------|---------------------------|--------------------------|---------------------------|-------------------------|--------------------------|
| | Submitted (n=161,666) | Unsubmitted (n=52,701) | Submitted (n=445,090) | Unsubmitted (n=31,380) | Submitted (n=40,041) | Unsubmitted (n=1,684) |
| Mining | 17'10" | 10'46" | 42'48" | 30'00" | 1°37'47" | 2°46'58" |
| Utilities | N/A | 12'06" | 20'34" | 15'37" | 35'08" | 21'17" |
| Construction | 50'51" | 07'14" | 50'03" | 44'14" | 42'40" | 50'58" |
| Manufacturing | 21'21" | 11'24" | 46'47" | 49'36" | 54'15" | 1°20'32" |
| Wholesale Trade | N/A | 07'03" | 32'08" | 27'34" | 38'20" | 36'10" |
| Retail Trade | 08'10" | 05'51" | 24'52" | 20'04" | 26'56" | 31'19" |
| Transportation and Warehousing | 07'09" | 04'58" | 18'38" | 15'43" | 23'14" | 27'01" |
| Information | 06'10" | 04'22" | 20'10" | 16'29" | 29'17" | 54'20" |
| Finance and Insurance | 04'56" | 04'07" | 19'15" | 15'36" | 28'24" | 25'05" |
| Real Estate and Rental and Leasing | 05'43" | 04'09" | 18'38" | 14'40" | 20'48" | 28'32" |
| Professional, Scientific, and Technical Services | 05'27" | 03'58" | 22'43" | 18'25" | 29'35" | 35'16" |
| Management of Companies and Enterprises | 06'26" | 04'01" | 16'06" | 14'31" | 30'28" | 148°40'40" |
| Administrative and Support and Waste Management and Remediation Services | 06'15" | 04'48" | 21'24" | 17'18" | 24'30" | 30'47" |
| Educational Services | 06'13" | 04'07" | 23'56" | 19'50" | 30'52" | 21'31" |
| Health Care and Social Assistance | 06'08" | 04'47" | 25'00" | 22'25" | 31'01" | 56'10" |
| Arts, Entertainment, and Recreation | 05'40" | 04'06" | 20'45" | 17'11" | 25'07" | 28'52" |
| Accommodation and Food Services | 06'04" | 04'27" | 19'30" | 18'01" | 26'24" | 28'37" |
| Other Services (except Public Administration) | 07'09" | 05'25" | 22'56" | 19'21" | 27'35" | 28'01" |

*Times are formatted as hh°mm'ss".

4.3 Software Usage Counts

In addition to frequency of user actions and timing between events, we have information about operating system, browser, and electronic reporting devices. Figure 5 shows the frequency of usage. Note that a user can use more than one of each operating system, browser and electronic reporting device. Microsoft's Windows is by far the most frequently used operating system, followed by the Macintosh computer and Apple's iOS. Internet Explorer is the most frequently used browser, followed by Chrome, Firefox, Safari, AOL, and Mobile Safari, which is the mobile devices using iOS. Desktop computers are the most frequently used reporting device, followed by other mobile devices that are not smartphones (for example, tablets), then smartphones, and, finally, all other devices combined, such as a Wii.



5. Discussion

Server-side paradata from the 2012 Economic Census Centurion software was used for our analyses. Data were available for all single-unit establishment users who successfully logged into the software. For this stage of research, we did not analyze data from potential Surveyor or paper respondents.

The analysis was broken into three major components: frequency of user actions, timing between events for submitted and unsubmitted forms, and software and hardware usage counts. Regarding the frequency of user actions, it took many users more than one attempt to log into Centurion. Additionally, almost half of users who logged into the system never logged out, and 12% of users who logged in never submitted a form.

For time spent using the online instrument, we analyzed the median time from the first action to the last action. Form submission requires a user to visit all required screens and click the submit button. As expected, the fewer the number of forms, the less time users spent in the system, with classification forms taking less time than the long/short forms. When comparing submitted and unsubmitted forms for users who started one or two forms, the time in the instrument was longer for those who submitted than those who did not submit. However, those who started and submitted three forms had a smaller median time in the system than those who did not submit. This may be attributable to the smaller sample sizes in the three-form category. Also, the time between first and last actions varies within sector, which is helpful for understanding how much burden we are placing on respondents. Additional burden is explainable in sectors such as manufacturing, in which the economic census is also filling the role of the Annual Survey of Manufactures, collecting additional details. In other sectors, it is less clear why it is taking longer to complete forms. This merits further investigation.

Finally, we examined information about software usage. While users overwhelmingly used Microsoft and Internet Explorer, some users used other types of software. While they largely used computers to report electronically, a small number used Smartphones or other devices. However, this should be monitored over time since this could change in the future.

These results lead us to ask the following questions:

- How do the failed logins for the economic census compare to BHS logins?
- What was the navigation pattern for users that did not submit forms?
- How many users switched devices?

Further analysis will attempt to address these questions.

6. Future Research

This work only scratches the surface of the research questions we would like to answer with the existing paradata. The next step is to look more into the findings from debriefings conducted with respondents shortly after they completed the economic census. For example, there is evidence from these debriefings that respondents had a hard time understanding the difference between error and warning messages, so we plan to analyze the number of error and warning messages users had on the review screen to supplement these findings. Respondents also mentioned having problems submitting the data, so we would like to analyze this further as well.

Of all the users that logged into the online instrument, 12.39% did not submit a form. Users cannot submit answers when they have errors, so perhaps this contributed to unsubmitted forms. In addition, 5,606 users logged in but did not start the form. A better understanding of the underlying causes of these events might improve the adoption of web reporting. Learning more about users that attempted to log in and were not able to would also be valuable.

Determining the likelihood of submitting a form based on respondents' operating systems, browsers, and devices respondents used is also something worth investigating. Perhaps users of certain devices are less apt to submit. Another thing to look into is whether users switch operating systems, browsers, or devices after signing in. Also, knowing the amount of time spent on each operating system, browser, and device, could be valuable in evaluating respondent burden.

For further insight, it would be helpful to merge the data for this research with auxiliary data from other sources. As an example, we are interested in how many forms were sent compared to how many forms were started. Additionally, as we consistently receive paradata, we would like to integrate use of this information in an adaptive design-type framework.

Finally, as client-side paradata become available in the future, more detailed research can be done. This would aid in identifying design flaws or measurement issues to improve the instrument and reduce respondent burden. For example, it would be useful to identify typical and atypical respondent navigation throughout the site, specific answers that respondents change, and when and where respondents access help.

References

Couper, M.P. (1998). Measuring Survey Quality in a CASIC Environment, *Proceedings of the Survey Research Methods Section, ASA*, pages 41-49.

Horwitz, R., Tancreto, J.G., Zelenak, M.F, Davis, M. (2013). Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument.
https://collab.ecm.census.gov/teamsites/paradata/Paradata_use_at_Census/2013_Horwitz_01.pdf

Kreuter, F. (2013). *Improving Surveys with Paradata*. Hoboken, NJ: Wiley

Pape, T. and Barron, S. (2013). A Glimpse Inside the Mind of a Respondent: Using Paradata to Improve Online Surveys. Paper presented at American Association for Public Opinion Research, Boston, MA.