USCENSUSBUREAU
Helping You Make Informed Decisions

ECONOMIC EDITING REDUCTION RESEARCH REPORT

Team: L. Kaili Diamond
Beth Newman
Brian Dumbacher
Geoffrey Hill
Jeff Dalzell
Justin Nguyen (Lead)
Carma Hogue (Champion)

Contributors: William Gainor
Venita Holland
Demetrius Lambeth

Report version 1.1
3/22/2017

EXECUTIVE SUMMARY

As a part of the Quality Assessment Staff's efforts to build quality into all aspects of the Survey Life Cycle as well as to increase efficiency of processes and the timeliness of data releases, a team was chartered to examine current survey editing to determine if editing and data review time could be reduced. The team lead is Justin Nguyen, the Economic Directorate's quality auditor. The team was to examine (1) the effect of analyst data changes on the final estimates, (2) the efficiency of the current edits to identify influential observations that affect estimate quality, (3) the availability of flags, audit trails, and edit evaluations for use in engineering a more "adaptive" edit processing system, and (4) the data review practices outside of the edit system. Sometimes editing practices add bias to estimates, and although it may be suspected, attempts to measure this type of bias are difficult at best. The team has not looked at this aspect yet.

The team started their research with the Annual Capital Expenditures Survey (ACES) and will continue to look at the Quarterly Services Survey and other surveys as time and resources permit. It is thought that what was found in ACES can apply to other surveys, but further research is needed to determine what adjustments will need to be made. Making the edit process more adaptive (enabling managers to shift analyst resources to more volatile or problematic cases) will ensure greater efficiency and timeliness. Monitoring changes in estimates and standard errors throughout the production cycle can be helpful for modeling when to stop editing data for certain North American Industry Classification System (NAICS) codes. Standardizing editing flags and procedures can enable analysts to shift from survey to survey without a long learning curve. Built-in monitoring of edit failures and audit trails could enable managers to understand more efficiently how the edit systems are working. Regular examination of paradata can also aid in questionnaire design that can allow data to be reported more accurately. The team will be examining as much of this as possible.

The results of the ACES research have shown that some simple changes may result in earlier release of estimates. What is yet to be determined is if the quality of the resulting estimates may actually improve with less micro-editing. The research has also shown the potential of using percent difference criteria and visualizations to model when editing can stop for certain NAICS codes. Such a model would be a key part of an adaptive editing process. The team is making the following recommendations:

Recommendation 1: Update the questionnaire design to reduce data slide occurrences.

Data slides account for both the greatest impact on estimates and the most time spent by analysts. These errors, which result from respondents reporting in dollars instead of the requested thousands of dollars, can be reduced or eliminated by changing the questionnaire and changing collection from thousands of dollars to dollars. Addressing this source of measurement error has been mentioned as far back as the 2002 ACES Quality Profile (Company Statistics Division, 2002, p. 29).

Recommendation 2: Follow specific editing guidelines determined by the known impact of editing practices.

Mathematical statisticians and survey analysts should work together to develop an editing hierarchy decided by the impact of the edits on estimates. As an example for ACES, data slides

have the most impact on the estimates, so all data slides should be given top priority and resolved first. Other details must be worked out after the slides are addressed.

Recommendation 3: Identify NAICS codes that need further review and set an edit priority order to these NAICS codes.

Analysts should identify NAICS codes that need further review and set a priority order for these NAICS. Table 5 gives one example of NAICS code priority order methodology, which uses three percent differences as criteria to flag for further review as well as a measure of discrepancy to set priority order. This type of table could be used to redistribute resources to NAICS that are likely to need further edits, have historically had the largest influences on the final estimate, and have the largest prior to current year changes.

Recommendation 4: Make edit flags more descriptive and detailed.

In order to conduct real-time edit evaluations for adaptive design approaches to editing, the edit flags must be more descriptive and preferably standardized.

Recommendation 5: In order to conduct similar analyses for other surveys, save the survey's estimation results files (ERFs) and control files frequently, perhaps twice a week as is done in ACES.

ACES is following a best practice, and others should emulate what is done with ACES.

Recommendation 6: Provide systematic analyst feedback to help define and address specific questionnaire shortcomings.

Feedback from analysts about their interactions with respondents may serve as an informal evaluation of the questionnaire and the data collected. In conjunction with the paradata, an examination of questions that continuously fail edits should be undertaken to see if the questionnaire wording or structure can be changed to reduce the number of edit failures.

Recommendation 7: Research applications of Big Data techniques such as machine learning to aid editing.

Data visualizations and machine learning methods based on predictors from audit trails and historical data have potential to identify influential observations such as data slides.

Future Research

The team will continue to expand their research on ACES and apply their methods on other surveys such as the Quarterly Services Survey (QSS) and the Services Annual Survey (SAS). Applying the same methods to previous years of ACES data and obtaining more data on the relationships between factors such as estimate stability and stopping points are additional steps in the modeling process. The EconEdit team would also like to create a hierarchical system of editing that prioritizes edits based on their impact on estimates and to investigate the possibility of automating certain edit types.

TABLE OF CONTENTS

Section 1: Introduction 1

Section 2: Research Background 2

Section 3: Research Scope 4

Section 4: Research Methodology 5

Section 5: Experiment: Examining Quantities Over Time 6

Section 6: Experiment: Editing in the Absence of Edit Failures 13

Section 7: Experiment: Impact of Editing 15

Section 8: Experiment: Modeling Stopping Points 17

Section 9: Summary of Results 23

Section 10: Recommendations 24

Section 11: Future Research 26

References 27

Appendix A: Editing Terms and Definitions 28

Appendix B: Research-Specific Variables and Flags 30

Appendix C: Additional Tables and Plots 33

Section 1. Introduction

Beginning with "An Evaluation of Edit and Imputation Procedures used in the 1982 Economic Censuses in Business Division" by Brian Greenberg and Thomas Petkunas (1986), it has been demonstrated that editing and imputation processes are often inefficient with regards to their effects on estimates. The impact on final estimates is not generally the basis for changes made to reported data, automated or by an analyst. One way to reduce the investment in editing would be to determine the most impactful edits and resolve them first.

The Annual Capital Expenditures Survey (ACES) is a mature program that has investigated this issue in the past. Since ACES does not use automated imputation for either missing or suspect data, it would seem that there is a reduced set of issues to examine. An automated weighting adjustment handles tabulating unit nonresponse, but analysts conduct all interactions with item nonresponse or suspect data. The quality of the ACES estimates over time is then not a function of changing imputation models, but clearly a function of the reported data, the interaction of the analyst with the dataset, and the overall effectiveness of the nonresponse weighting adjustment.

ACES is an annual survey that collects information about capital expenditures made in the previous calendar year. Data collection for a new sample begins with a mailout each March, three months after the end of the calendar year. The data collection period extends about 8 months, into October. Data editing may continue past this point for several more weeks. The release of the final data product is usually in February, which is 11 months after initial mailout, and 14 months after the survey reference year. A reduction in the time spent on data collection and editing could move dissemination into the same year as mailout.

In the past, ACES staff have discussed earlier release dates for publications. Changes to the automated estimation routines late in the ACES 2007 cycle began the archival of estimates and corresponding standard errors. ACES 2008 was the first full cycle of ACES to have this information. This archiving of estimates has continued from ACES 2008 into the current cycle, ACES 2015. However, in the earlier years archiving did not begin until late summer and was not always on a set schedule. By the ACES 2014 cycle, estimates and standard errors were archived twice a week (Tuesdays and Thursdays) beginning with survey week 7.

Some sources of suspect data are inconsistencies in the distribution of total company investments among various industries and investment types. Making the microdata internally consistent is laudable but may not have appreciable effects on the estimates. Internally consistent microdata is important if the survey deems the microdata to be its most important product. While the outcome of the research into ACES may only be generally indicative of the cost benefits of abbreviated data collection, data editing, or both collection and editing, ACES has simplicities and archival data that benefit this research.

The following outlines the organization of the rest of this report. Section 2 provides background on previous editing reduction research in the Economic Directorate, and Section 3 describes the scope of this research. This report focuses on ACES. Section 4 describes ACES data sources, methods, and the four experiments conducted. Sections 5 through 8 provide details and results for the four experiments. Section 9 summarizes results and Section 10 provides recommendations. Lastly, Section 11 gives ideas for future research.

Section 2. Research Background

For decades, the U.S. Census Bureau has known the problem of over-editing data. In 1994, Carol King and Michael Kornbau issued a report "Inventory of Economic Area Statistical Practices." In it, they noted, "Based on the answers to the inventory questions on editing and imputation, our major conclusion is the Economic Area surveys have a large investment – both in terms of software and personnel – in editing and imputation procedures but, with few exceptions, have limited knowledge of how different types of editing and imputation affect the resulting estimates." They went on to recommend that programmers, methodologists and subject matter experts need to gain a better understanding of the efficacy of editing and imputation procedures.

In 2005, the Business Process Improvement Sub-team on Editing Efficiency conducted a study of editing and imputation practices. This study investigated how many projects measured the effects of their editing and imputation practices on published estimates. Of the 51 projects that responded, only 8 compared estimates calculated from unedited data to what was published. Moreover, only 3 respondents indicated a plan to research flags to determine which items had the most changes in order to assess edits, edit parameters, or data collection instruments. Recommendations from this report included "programs should calculate edit-failure rates" and "programs should measure the effect of the editing process on the resultant data." At the time, 31 respondents indicated they did not calculate edit failure rates. All of this flowed naturally from the original Business Process Improvement team's findings on the number of programs spending tens of thousands of hours editing data. Also of note from the Business Process Improvement Sub-team on Editing Efficiency was the finding that editing accounted for an estimated 25 to 40 percent of overall survey costs.

In terms of methodology for improving the processing of Economic Census data, the 2005 Business Process Improvement Sub-team report suggested the Census Bureau adopt more Hidiroglou-Bertholot edits (Hidiroglou and Bertholot, 1986). Fifty percent of the respondent programs had units in consecutive cycles with an identifier to allow for matches of reported data between cycles. Hidiroglou-Bertholot (HB) edits were preferred as a possible way to standardize edits across systems. The sub-team proposed this as a possible way to shorten the time used for editing data as well, since 30 percent of the respondent programs used a top-down approach of "largest" edit-failing units until the end of the survey cycle.

In 2015, Beth Newman, a mathematical statistician working on ACES, began research on identifying possible stopping points for ACES data collection and editing. Investigation focused on three sector-level estimates of total capital expenditures for employer companies, tracking both the estimate and coverage rates from the initial mailout to publication using ACES 2013 data. Newman's findings suggested that data editing and late responses during the last nine weeks before publication did not influence the total capital expenditures estimate for the three sector-level estimates as much as the previous collection weeks. Her findings note that, even though capital expenditures totals remain relatively stable throughout the final nine weeks, the categorization of employer capital expenditures by type (e.g., structures/equipment, new/used) may change significantly over this time. Newman recommended research on additional survey items, as well as exploration of other quality indicators and nonresponse adjustment techniques for multi-unit companies.

The efficiency of editing techniques is of particular importance to methodologists in ESMD. In 2015, Chin-Fang Weng, a mathematical statistician working on public sector surveys, authored "New Methodologies for Identifying Response Error," which looked at the impact of editing using an HB edit methodology, a bivariate index (BI) editing methodology, and a robust regression (RR)

editing methodology. The BI editing methodology for identifying potential errors is particularly interesting because it is based on two criteria: relative difference, which relates to the probability of a true error, and relative importance, which relates to the impact that the error has on estimates. Simulation results showed that both the BI and RR methods simultaneously improve data quality and identify fewer edits.

On an added note, Appendix A contains important editing terms and definitions.

Section 3. Research Scope

It is common for survey methodologists to question editing practices. This is largely due to the substantial resources that editing consumes. However, there are other concerns with editing practices and timelines including the delay of data releases and the possibility that over-editing can actually have a negative impact on data quality (Granquist, 1997).

Therefore, the task of answering questions about the Economic Directorate's current editing processes and schedules was given to the Economic Editing Reduction Research (EconEdit) team. The first task was to determine whether certain economic surveys could stop the editing process earlier than is currently standard practice. If so, are there clear stopping points after which the quality of the estimates remains high? If clear stopping points are not currently present, are there ways to determine these stopping points in the future?

While the team plans to look at a number of surveys, the team chose ACES as the initial survey of focus. It would have been labor-intensive and inefficient to investigate all item edits for all North American Industry Classification System (NAICS) codes in ACES. Therefore, the EconEdit team began by taking a high-level approach that took all NAICS and items into account, and then focused on the NAICS codes that had the most influence on total capital expenditure estimates, as well as items determined to be most important or most influenced by analyst edits. In general, the team set out to decide whether the editing process for ACES could stop earlier, without detriment to the estimates, thus saving valuable resources and moving estimate dissemination into the same calendar year as the initial mailout.

Section 4. Research Methodology

4.1 Data

The EconEdit team focused on the ACES 2014 cycle and used various data sources in its research. The team obtained audit trails for the survey's control, item, and roster data. The control data contain unit-level information such as inactivity status, check-in dates, nonresponse follow-up status, and sample weights for each of the approximately 52,000 respondents. The roster data contain values for the 600-level variables (see figure B2), whereas the item data contain values for the other variables on the ACES questionnaire.

Each record in the audit trails contains information about a single update to a data value. Updates correspond to actions such as the loading of a reported value or the editing of a value by an analyst. The variables on the item and roster audit trails include unit ID, item code, the old and new values, the old and new reporting flags, the name of the user who updated the value, the name of the program or routine that updated the value, and the program date and time stamp. Using the date and time stamp, it is possible to determine a unit's values for a given date and time.

The EconEdit team also used estimation results files (ERFs) and control files saved twice a week throughout the ACES 2014 cycle. Methodologists in ESMD working on ACES saved these files for every Tuesday and Thursday. The ERFs contain estimates and standard errors by NAICS code for the variables in the roster table. These estimates and standard errors are calculated using production code in the Standard Economic Processing System (StEPS), and weighting adjustments account for nonresponse. The ERFs, control files, and audit trails complement each other in very informative ways. For example, one can observe changes in the estimates from the ERFs over time and then go through the audit trails to investigate edits that could have contributed to these changes.

4.2 Experiments

As mentioned previously, the main goal of this research was to determine whether the editing process for ACES could stop earlier. There were many things to look at and approaches to take in making this determination, so the EconEdit team decided to break the problem down into smaller pieces, or experiments. Each experiment focuses on a different aspect of the over-editing and stopping point challenge. Table 1 summarizes these experiments. Sections 5 through 8 describe details and results.

Table 1: Summary of Experiments

Experiment	Purpose
Examining Quantities Over Time	To examine raw sums, estimates, standard errors, and the number of edit failures over time. (Section 5)
Editing in the Absence of Edit Failures	To understand the nature of changes made to data that did not correspond to addressing an edit failure. (Section 6)
Impact of Editing	To quantify the impact of editing on estimates by NAICS code and by edit type. (Section 7)
Modeling Stopping Points	To model when to stop editing certain NAICS codes and switch resources to other NAICS codes (Section 8).

Section 5. Examining Quantities Over Time

Examining raw sums, estimates, standard errors, and the number of edit failures over time and understanding how changes in these quantities relate to editing procedures were important first steps in the analysis.

5.1 Raw Sums

Using only the data in the item and control audit trails, the EconEdit team examined how raw sums changed over the ACES 2014 cycle. Figure 1 plots raw sums (normalized to the unit interval) over time for items 221 (total capital expenditures for structures), 222 (total capital expenditures for equipment), and 224 (total capital expenditures). Receiving and correcting data reported in the wrong units caused the plots to exhibit large spikes and drop-offs. The most common case for ACES involved values reported in dollars instead of thousands of dollars. Data slides are the edits that correct such data. For 221 and 222, the raw sums stabilize at the end of September, whereas the raw sum for 224 stabilizes at the end of October.

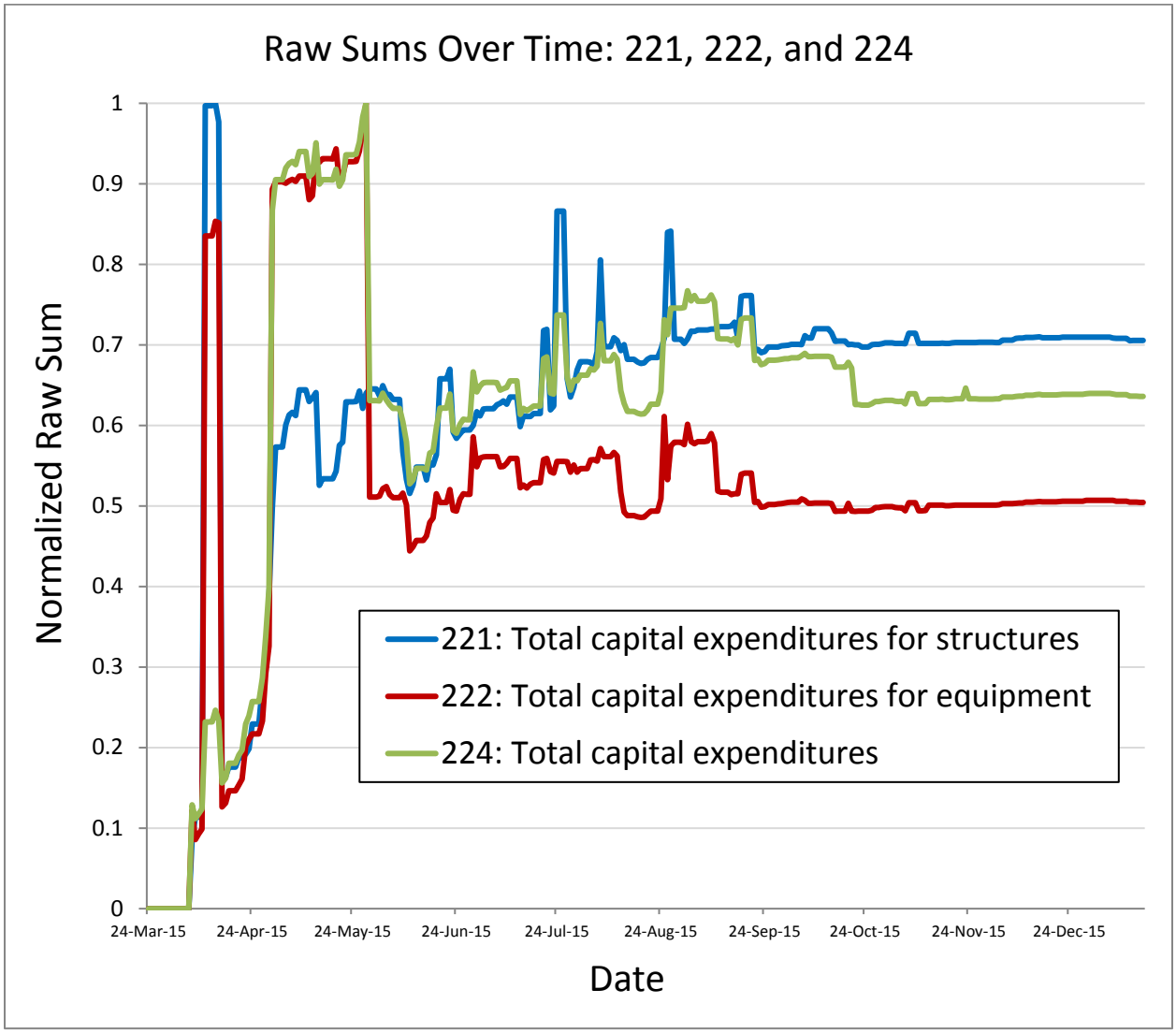


Figure 1: Plots of raw sums over the ACES 2014 cycle for items 221, 222, and 224.

5.2 Estimates

Due to the saving of ERFs for ACES throughout data collection and data editing, the EconEdit team was able to observe how estimates calculated using production code and nonresponse weighting adjustments behaved over time. Many ACES stakeholders are interested in capital expenditures estimates for companies with employees at detailed NAICS code levels. As a start to future investigations, this research analyzed six selected NAICS codes out of the 132 four-digit NAICS codes available on the ERFs. Influential sectors were chosen based on their contributions to the total capital expenditures estimate. The six select industries are the following:

- 2211 (electric power generation, transmission, and distribution)
- 3240 (petroleum and coal products manufacturing)
- 3361 (motor vehicle, body, trailer, and parts manufacturing)
- 5171 (wired telecommunications)
- 5222 (depository credit intermediation)
- 5310 (real estate)

Tracking estimates over the data collection period provided a high-level view of estimate stability and insight into possible stopping times. Previous sector level research by Newman (2015) looked at total capital expenditures estimates (item 610) over the data collection period. Newman noted that while the total capital expenditures estimates appeared stable, analysts made many edits to "new" and "used" and between "structures" and "equipment" near the end of data collection. These changes possibly made the estimates more unstable. Using this information, analysis focused on new (612) and used (613) structures, and new (615) and used (616) equipment. The team created graphs of estimates and their 90 percent confidence intervals over the data collection period, broken down by NAICS code and item.

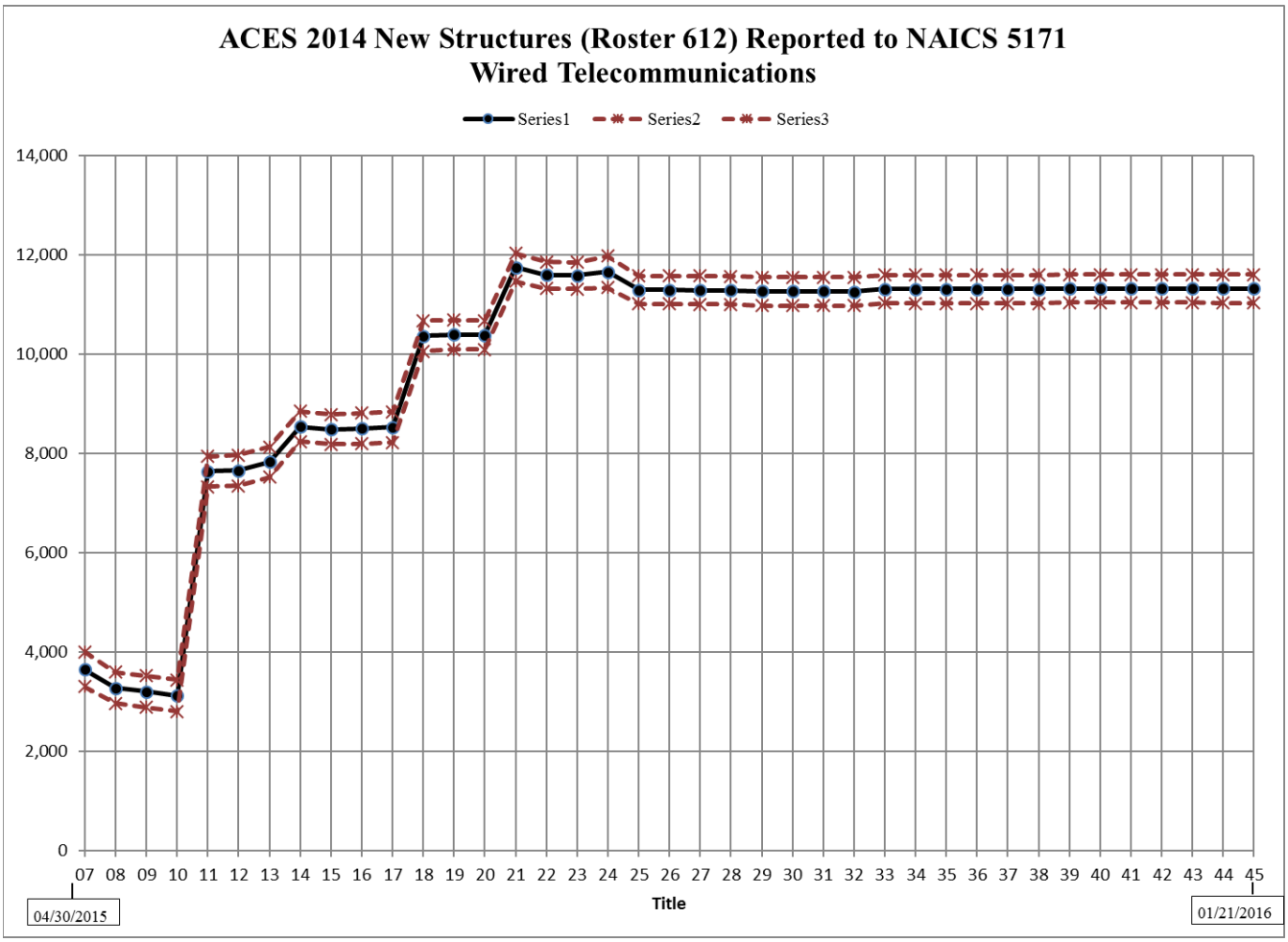


Figure 2: Plot of estimates and 90 percent confidence interval limits over time for new structures (612) reported to NAICS code 5171.

Figure 2 is an example plot for new structures (612) reported to NAICS code 5171 (wired telecommunications). Note that the confidence intervals reflect sampling error only. This graph does not account for sources of non-sampling error such as measurement error. These graphs helped select the four dates chosen for closer examination. These points in time were collection weeks 33 (October 29, 2015), 37 (November 26, 2015), 41 (December 24, 2015), and 45 (January 21, 2016). Week 45 is the final week and corresponds to the published estimate. The estimates and confidence intervals for these four dates were then isolated. Figure 3 plots the same estimates and 90 percent confidence interval limits as in Figure 2 but for these four possible stopping points.

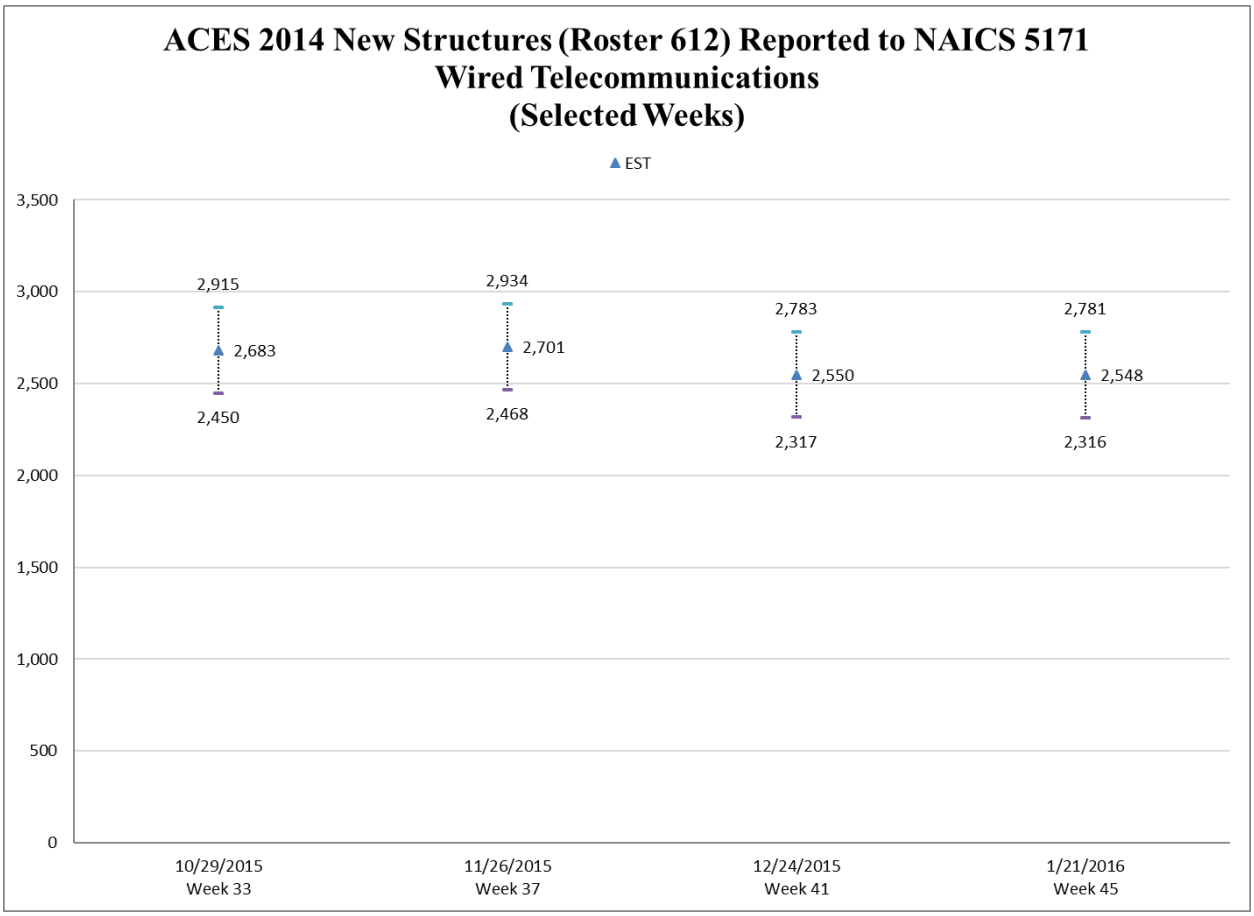


Figure 3: Plot of estimates and 90 percent confidence limits for four possible stopping points for new structures (612) reported to NAICS 5171.

Another way of looking at the stability of the estimates is to examine how far away they are, in terms of percent difference, from the final estimates. Figure 4 plots these percent differences for the six selected four-digit NAICS codes that are identified previously as well as for all NAICS codes. It is interesting to observe that for all of the four-digit NAICS codes except 5222, the percent differences are close to 0 percent in early October. For 3240 and all NAICS codes, the plots exhibit bumps in late October and early November during which the percent differences increase or decrease but then return close to their original values.

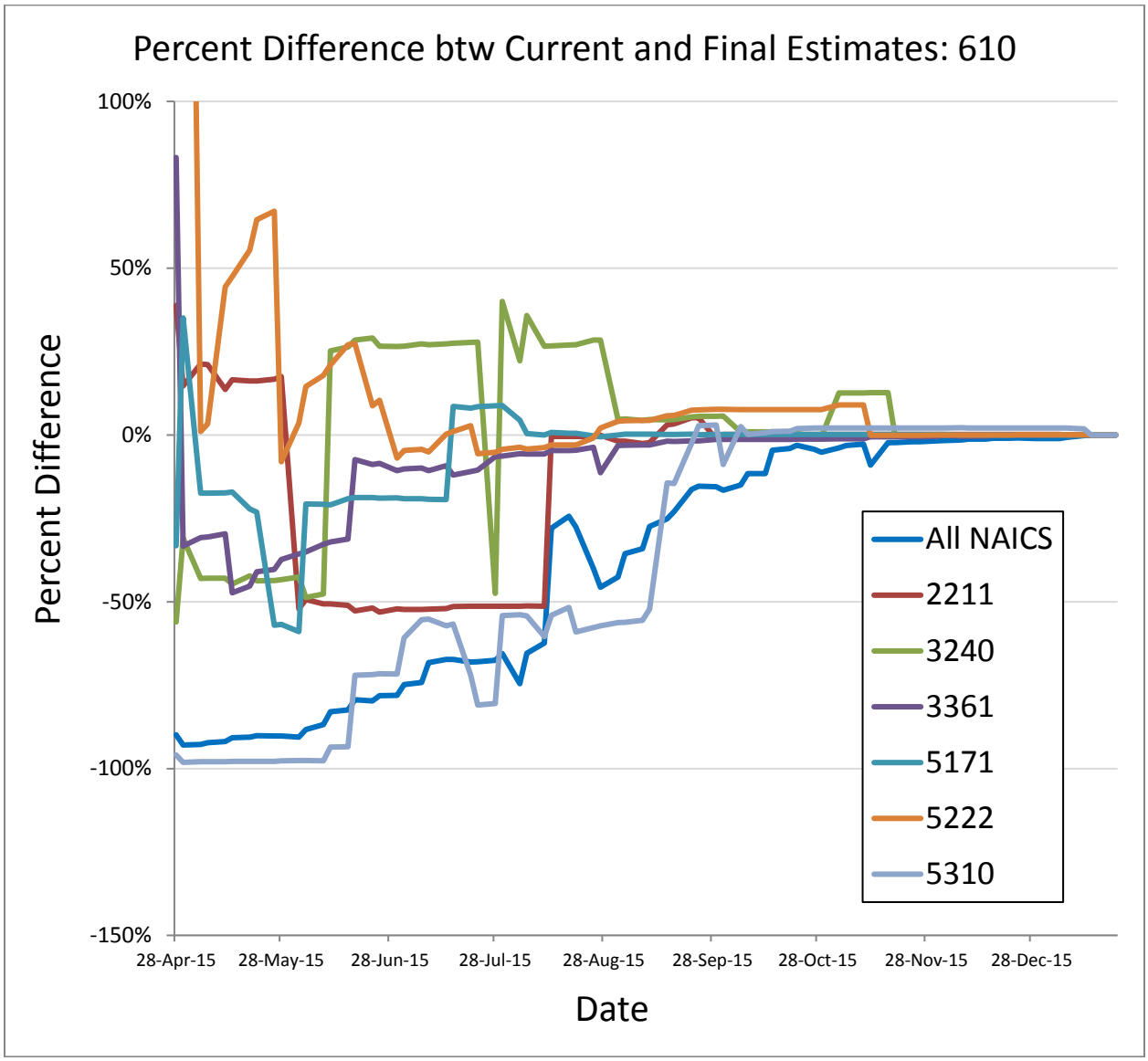


Figure 4: Plot of percent differences between current and final estimates over the ACES 2014 cycle for item 610 (total capital expenditures).

For each time point, NAICS code, and roster variable on the ERFs, the EconEdit team observed whether the 90 percent confidence interval (CI) contains the final estimate. Figure 5 plots the proportion of CIs containing the final estimates over time for Roster item 610 (total capital expenditures).

It is interesting to see that the proportions for the roster variables exhibit similar behavior throughout the ACES 2014 cycle[1]. The proportion for all 132 NAICS codes equals about 30-40 percent at the end of April, 70 percent at the end of September, 80 percent at the end of October, and 90 percent in early December.

[1] Graphs of other roster items excluded due to disclosure avoidance laws.

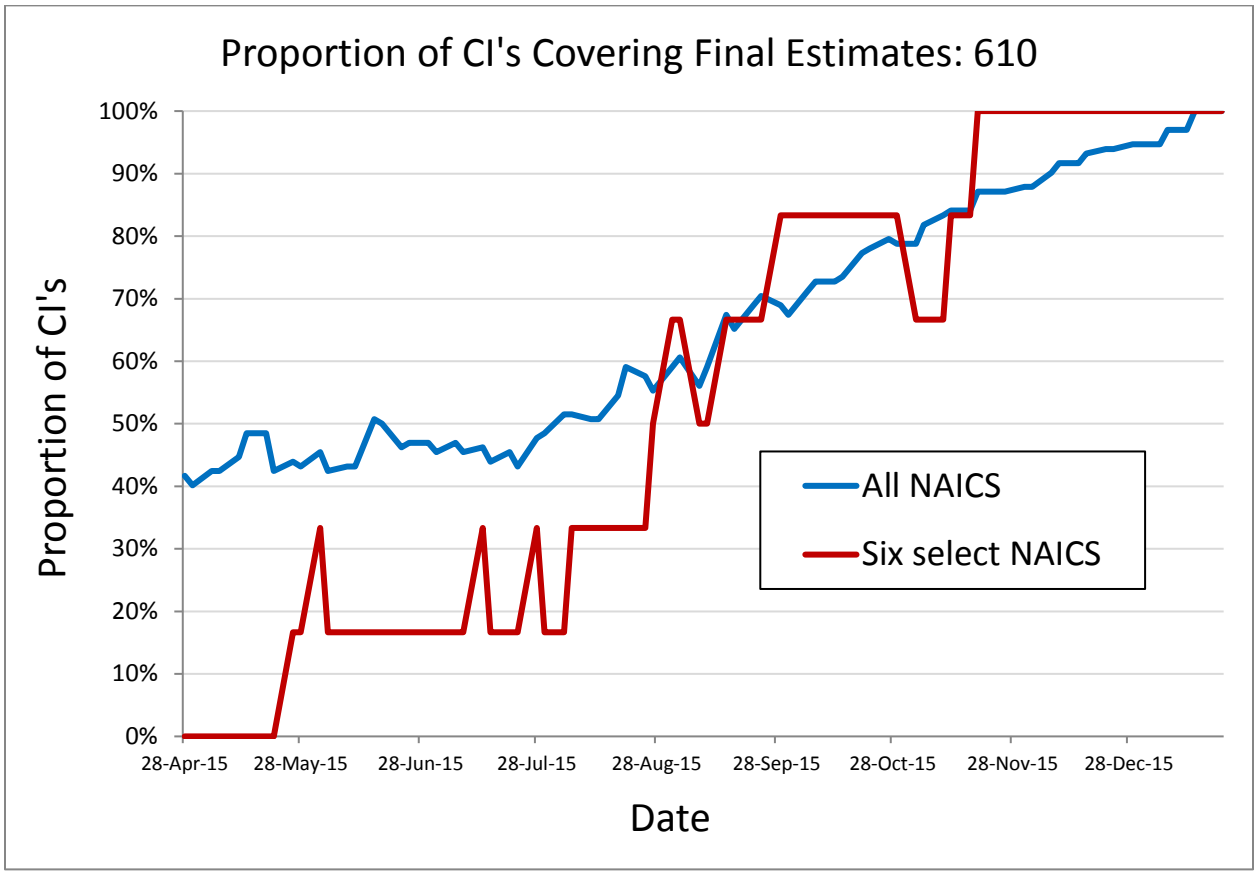


Figure 5: Plot of proportion of 90 percent confidence intervals covering final estimates over the ACES 2014 cycle for item 610 (total capital expenditures).

The proportion of CI's covering the final estimate for item 610 (total capital expenditures) was also examined for ACES survey years 2010, 2011, and 2013 (see figures C1-C3 in Appendix C). The team used the available ERFs to create these graphs. While the general trends appear to be similar, there is not enough information for these years to draw any concrete conclusions at this time. The exclusion of survey year 2012 was due to the government shutdown in October of 2013, which interrupted the editing process.

5.3 Number of Edit Failures

The flags on the audit trails only distinguish between reported data, automatic edits, and analyst edits. The EconEdit team applied the item edit criteria in Appendix B and created its own edit failure flags to investigate the changes made to data values. There are 29 failure flags in total. The exclusion of flags 23-27 from this part of the analysis was because they have to do with roster values, which were missing for a large majority of records on the item audit trail.

Figures 6 and 7 plot the number of edit failures by flag. Showing the workload of the analysts and the varying importance of the flags were the motivation for these figures. Differences in scale created the need for two figures. Figure 6 shows that the number of edit failures tended to decrease over the ACES 2014 cycle. However, for some flags, the number does not quite reach 0. Figure 7 shows some flags that exhibit different behavior. Flag 1 deals with allocating the "other" capital expenditures reported in the 300-level items between structures and equipment in the 200-level table. See Appendix B for a screenshot of this part of the ACES questionnaire. The number of failures for flag 1 increases to a peak in May and then rather rapidly decreases. Flags 28 and 29

deal with whether the company ceased operation or changed ownership. These flags do not have to be resolved, thus calling into question the need for these flags.

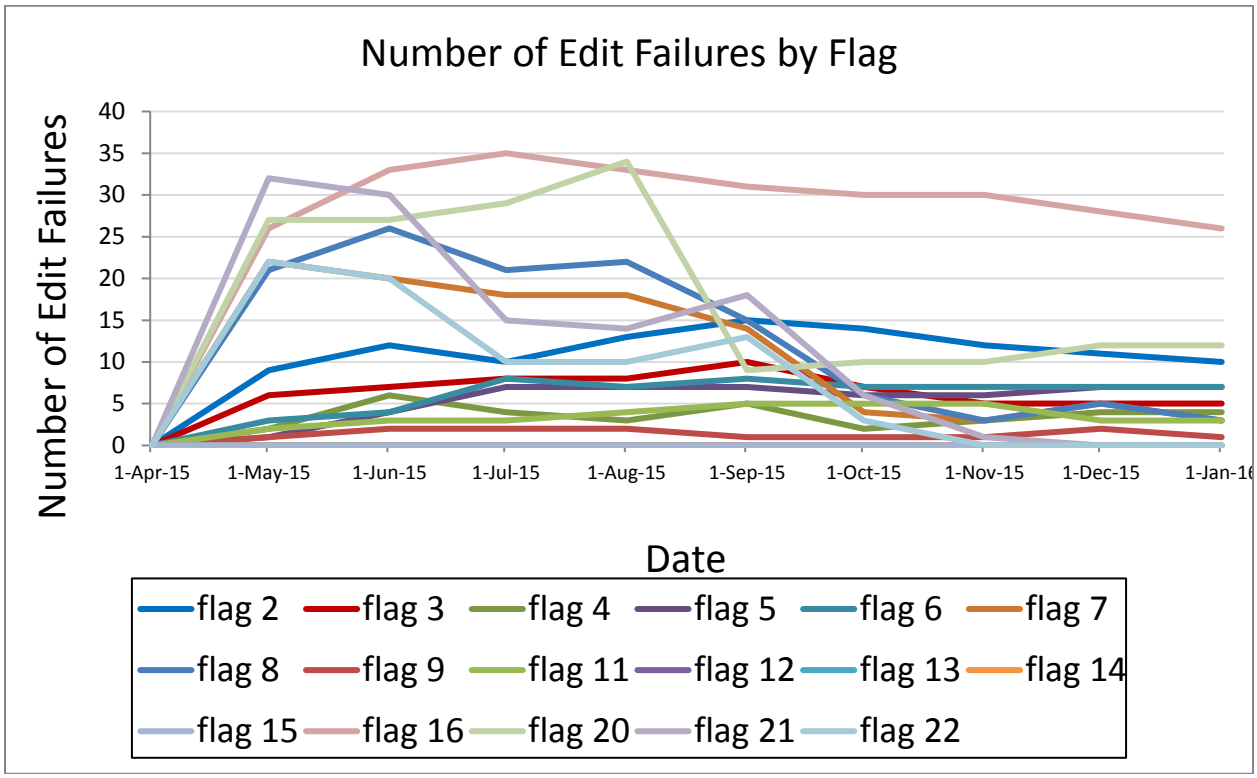


Figure 6: Number of edit failures over time by select flag.

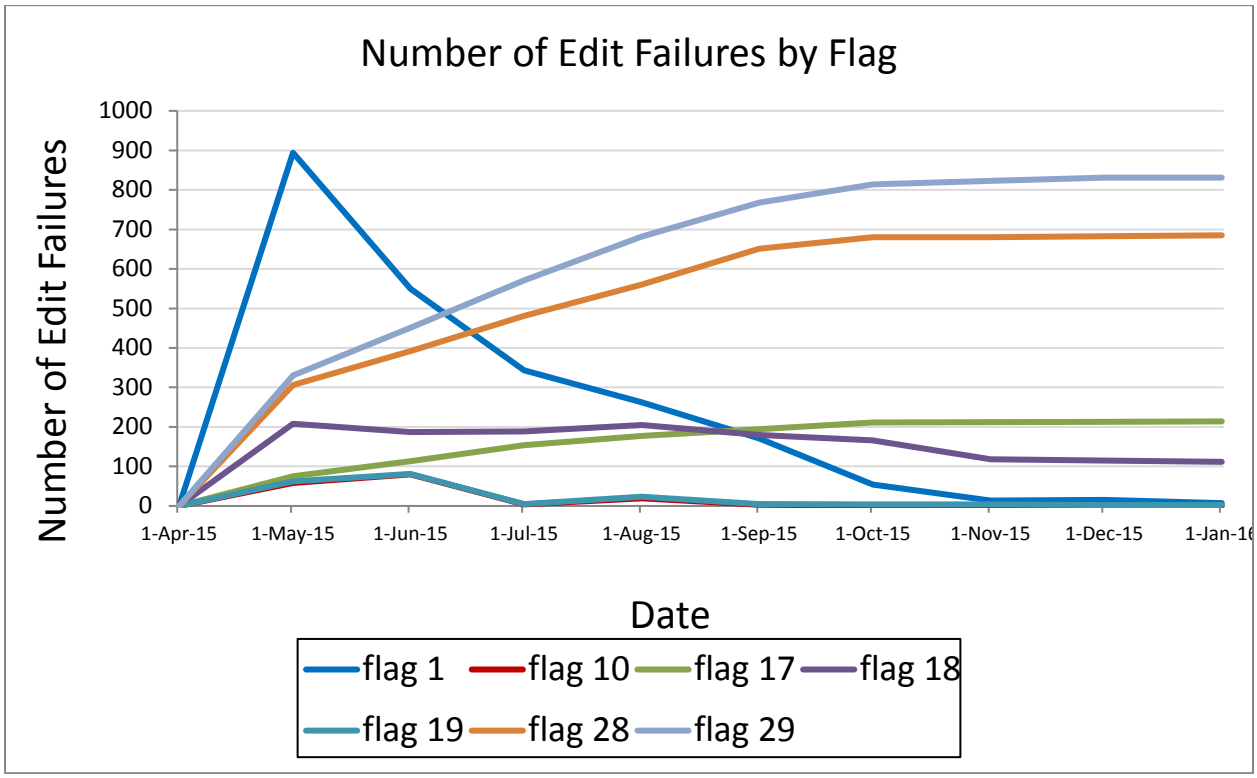


Figure 7: Number of edit failures over time by select flag.

Section 6. Editing in the Absence of Edit Failures

The act of making microdata as accurate as possible is laudable, but it may involve editing the same records multiple times or making very small changes to values. This extra effort may not affect final estimates appreciably. See Granquist and Kovar's 1997 paper "Editing of Survey Data: How Much is Enough?" for a good discussion of over-editing. In this experiment, the EconEdit team used the item and control audit trails to examine edits made in the absence of edit failures, i.e. edits made to units whose item values passed all of the edit criteria. The team refers to edits that adjust the data but do not correct any edit failures as "non-failure" edits. The EconEdit team examined the nature and impact of these edits.

6.1 Non-Failure Edits

To determine the non-failure edits, the EconEdit team applied the edit criteria in Appendix B and excluded edit failures 23-27, as done in Section 5. The EconEdit team identified 9,083 non-failure edits, classifying 3,110 of which as data slides. Because there is no data slide flag, the EconEdit team developed its own criteria to identify old and new values that differ by powers of ten. See Appendix B for these criteria. Figure 8 plots the number of non-failure edits by item and data slide. The 900-level and roster items had very few non-failure edits. The team omitted these items from Figure 8 to save space. It was readily apparent that a large number of non-failure edits were made to items 101 (gross depreciable assets) and 162 (gross domestic sales). ACES does not publish estimates for these two items but does provide these estimates to the Bureau of Economic Analysis (BEA). Analysts edited these two items throughout the data collection cycle.

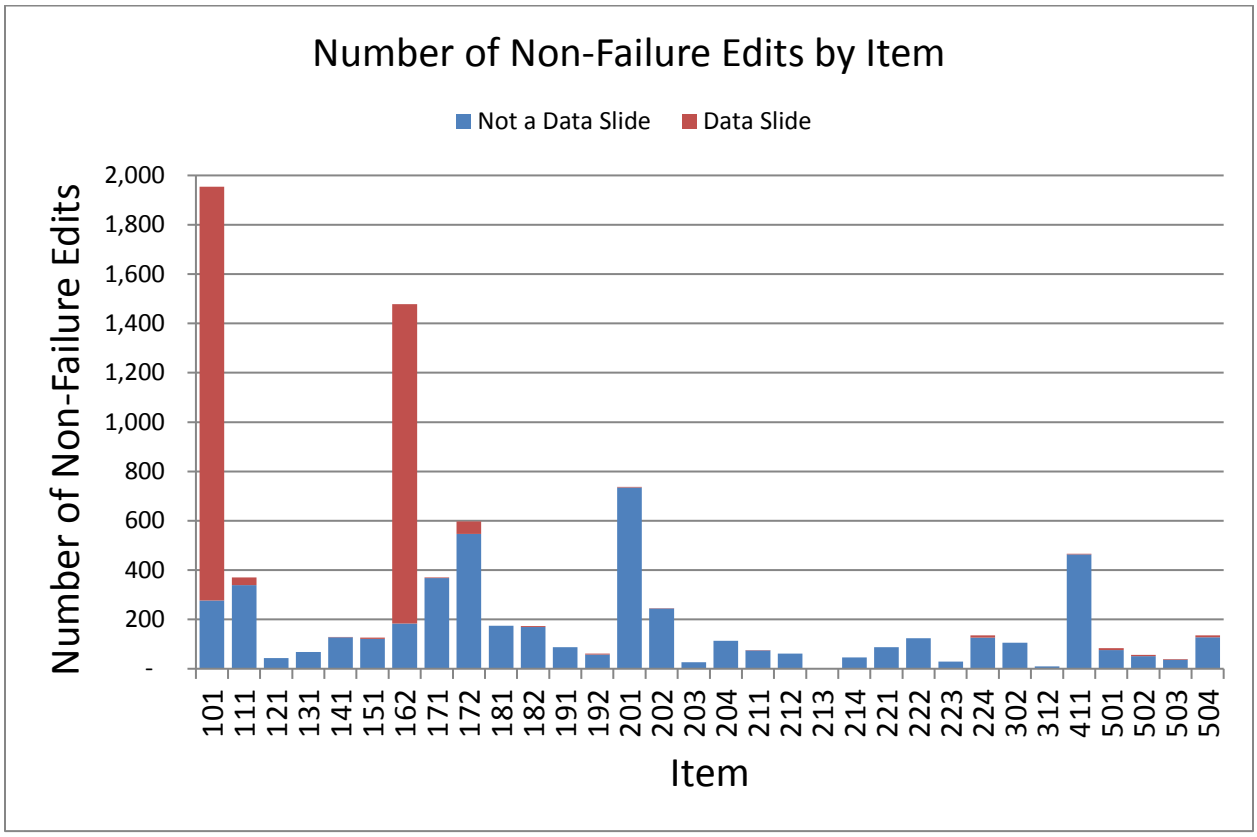


Figure 8: Number of non-failure edits by item and data slide.

6.2 Changes in Value

Next, to quantify the impact of these non-failure edits on the data, the absolute differences between the old and new values were calculated. Because of the large scale of the numbers involved, the team applied a natural log transformation to the absolute differences:

$$\boldsymbol{log_dif = ln|newval - oldval|}$$

This part of the analysis excluded a small number of edits that had the same old and new values. Figure 9 plots the distribution of these logged differences by whether or not a data slide caused the difference. The largest differences tended to be associated with data slides. This was further evidence that data slide adjustments had the most impact on the data. A large majority of the overall change in data values for the non-failure edits was attributable to data slides as opposed to other types of edits.

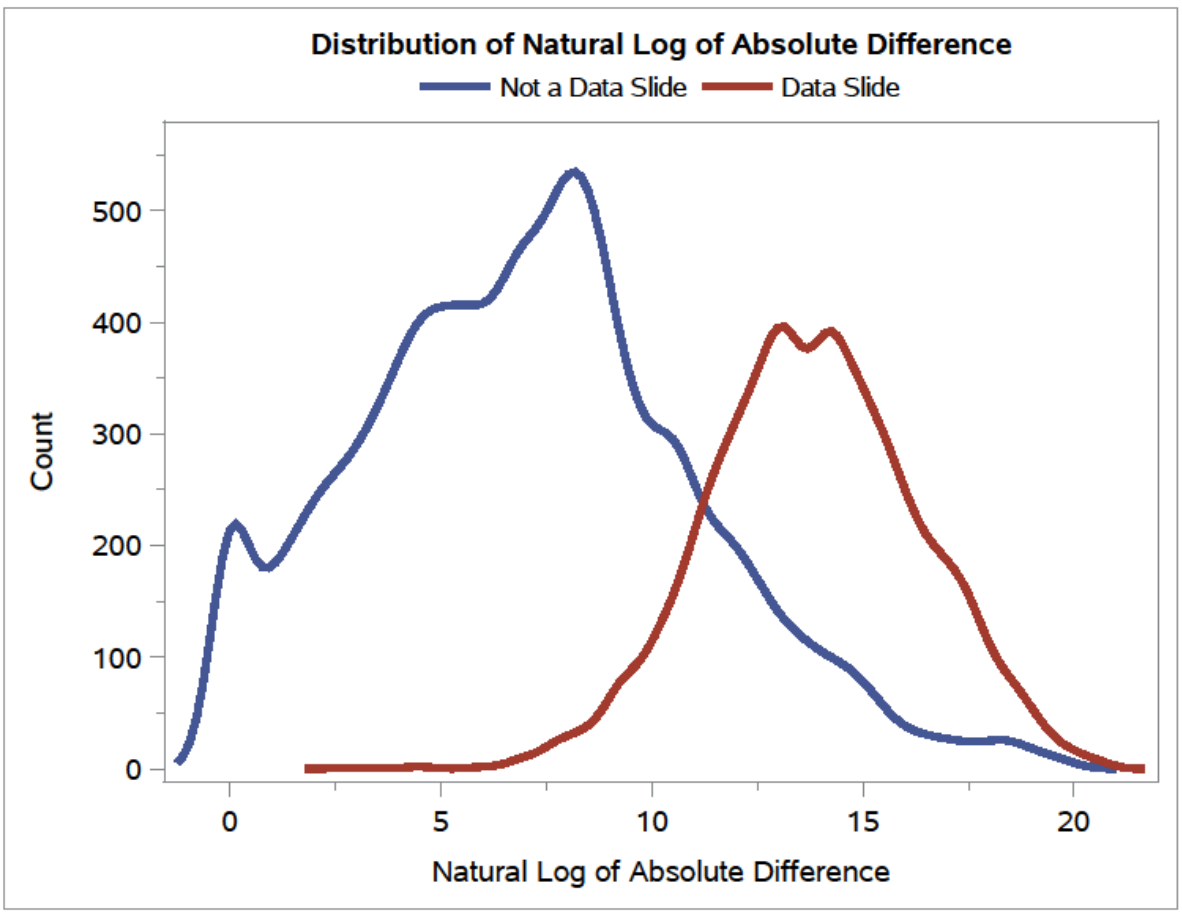


Figure 9: Distribution of natural log of absolute difference between old and new values for non-failure edits broken down by data slide status.

Section 7. Impact of Editing

In the third experiment, the EconEdit team examined the impact of editing on estimates. The previous two experiments have shown that the data slide edit type had the largest impact on ACES data and estimates. However, the team did not know the impact of other types of analyst edits or what items accounted for the most analyst edits.

7.1 Impact on Estimates

To answer these questions, the team removed all data slides from the dataset. Then the dataset was separated into each type of item value (reported, automatically edited, and analyst edited). For each ID, the team kept the last edit for each type and dropped the previous edits of the same type. Then these three datasets were merged to create the final dataset which contained one observation per ID with a value for each type (reported, automatically edited, and analyst edited) where applicable. The absolute difference in the final values of each item follows:

- |reported value – automatically edited value|
- |reported value – analyst edited value|
- |automatically edited value – analyst edited value|

The team summed these absolute differences by stratum and item. However, since certain items will inherently have larger values these raw values on their own had little meaning. Therefore, we needed a ranking metric that also accounted for the relative size of the item values. The chosen metric was:

$$\frac{|\boldsymbol{newval} - \boldsymbol{oldval}|}{\boldsymbol{oldval}} \times \boldsymbol{ln}\,|\boldsymbol{newval} - \boldsymbol{oldval}|$$

This was calculated using the value after analyst edits as the new value (newval) and the last value assigned to the item (reported or automatically edited) as the old value (oldval). The team summed these values across sample strata for each item. We then ordered by the absolute value of this metric, creating a ranking system that took into account the relative size of the items. Table 2 shows that the "other" and structures and equipment items are greatly impacted by analyst edits. This makes sense as the analysts often try to evaluate what the respondent put into the "other" columns and redistribute these totals to the proper items.

Table 2: Items Most Impacted by Analyst Edits, 2014 ACES

Rank	Item	Item Description
1	0211	Used Structures
2	0411	Capital Leases
3	0214	Total Used Structures and Equipment
4	0223	Other Total Capital Expenditures
5	0302	Item Value for Other Capital Expenditure
6	0203	Other New Structures and Equipment
7	0213	Other Used Structures and Equipment
8	0312	Item value for Other Capital Expenditure
9	0221	Total (New and Used) Structures
10	0212	Used Equipment

7.2 Time Spent on Editing

Another key component of this research was to try to determine where the analysts were spending most of their time editing. While the team cannot know the actual time spent per edit, they instead counted the number of analyst edits per item and then ranked them in descending order. Table 3 shows where analysts make the most edits. As should be expected, there is some overlap between the items that are most impacted by analyst edits and the items that are edited most by analysts.

Table 3: Items with the Most Analyst Edits, 2014 ACES

Number of Corrections	Item	Item Description
5,417	0222	Total Equipment
5,176	0202	New Equipment
4,437	0223	Other Total Capital Expenditures
4,213	0302	Item Value for Other Capital Expenditure
4,208	0141	Gross Depreciable Assets
4,112	0203	Other New Structures and Equipment
3,865	0204	Total New Structures and Equipment
3,797	0224	Total Capital Expenditures
3,565	0172	Value of Sales in the Largest Industry Reported by the Company
3,202	0111	Total Capital Expenditures
2,989	0201	New Structures
2,969	0162	Gross Domestic Sales, Operating Receipts, and Revenue
2,510	0221	Total (New and Used) Structures
2,486	0101	Gross Depreciable Assets (Structures and Equipment) at Beginning of Year
2,262	0151	Accumulated Depreciation and Amortization at Year End
2,026	0214	Total Used Structures and Equipment

Three of the six most edited items are ‘other’ items. This shows further evidence that analysts were spending a large proportion of their time redistributing the values from the “other” categories.

Section 8. Modeling Stopping Points

In this last experiment, the EconEdit team researched macro-level models for determining when to stop editing certain NAICS codes and shift editing resources to other NAICS codes. Such a model would be a key part of an "adaptive" editing process. This experiment focuses on when estimates tend to stabilize around reasonable values and not when data collection can stop. Data collection is involved to the extent that newly obtained data may affect estimates. The models considered in this experiment took the form of regression models and percent difference criteria to flag changes in estimates. In addition, the team explored methods that assigned an editing priority to NAICS codes.

8.1 Methods for Flagging Estimates

Models that flag abrupt changes in estimates can be useful for determining which NAICS codes data need review. First, the EconEdit team considered a weighted robust regression model for each four-digit NAICS code in which modeled the estimate for total capital expenditures as a linear function of the date (expressed numerically as the number of days since January 1, 1960):

$$estimate_t = \beta_0 + \beta_1 date_t, \quad t = 1, \dots, T,$$

where T is the number of dates under consideration. To consider estimate uncertainty, the regression weight was set equal to the inverse square of the estimate's standard error. These models were fit using a moving window of estimates from the ERFs using SAS PROC ROBUSTREG, which has options for identifying outliers. The logic to this approach is that if the most recent estimate is not flagged as an outlier, then resources can be shifted to other NAICS codes. PROC ROBUSTREG flags data points as outliers if their residuals are greater than three estimated model standard deviations. Estimates identified as outliers in this manner corresponded reasonably well to large changes in estimates observed after a period of stability. Figure 10 shows an example.

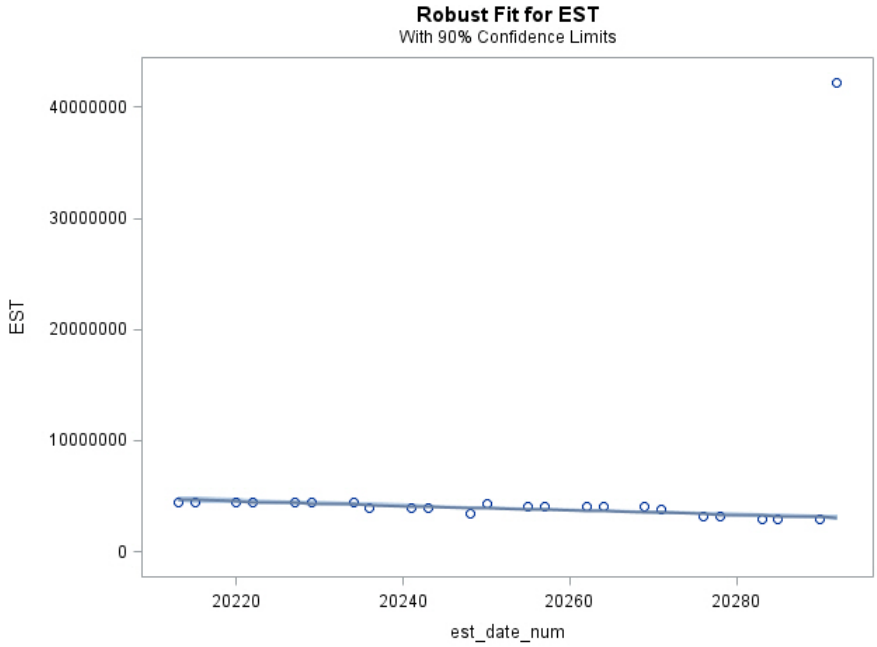


Figure 10: 2014 robust regression for total capital expenditures for a certain NAICS code using ERFs from May 1, 2015 through July 23, 2015. The estimate on the far right corresponds to July 23 and received the appropriate outlier flag. This would indicate that resources be shifted to this NAICS code.

With this robust regression approach, the number of data points,T, used in model fitting was relatively small and limited by how often the ERFs were created. Along with some of the other model parameters in PROC ROBUSTREG, T seems arbitrary. A more non-parametric approach to identifying outliers may be more appropriate. To this end, the EconEdit team considered models based on percent differences in the estimates and standard errors. Meeting any of the following three criteria flagged that estimate for review:

1. The ERF-to-ERF percent difference in the estimate is greater than 10 percent
2. The ERF-to-ERF percent difference in the standard error is greater than 10 percent
3. The percent difference between the estimate and previous year's final estimate is less than -66 percent or greater than 200 percent

The thresholds used in these criteria are arbitrary as well but easier to interpret. Figure 11, which resembles a control chart, plots 2014 estimates over time for NAICS code 2131 (support activities for mining) with estimates flagged according to these criteria. Appendix C provides similar plots for years 2010 through 2013 (plots C11-C14). For these plots, note the differences in scale.

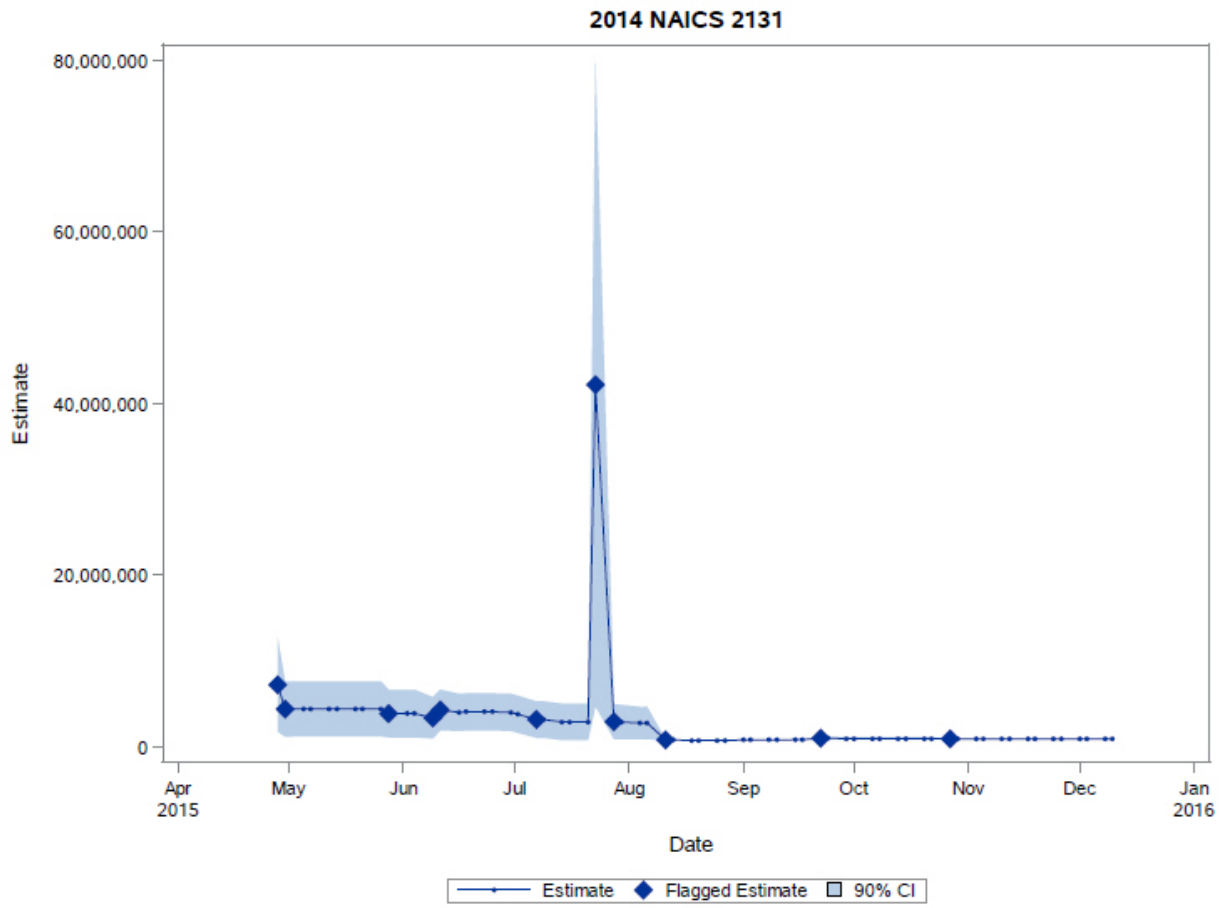


Figure 11: 2014 times series plot of estimates and 90 percent confidence band for NAICS 2131 with estimates flagged according to the percent difference criteria. These criteria suggest several periods of stability beginning in mid-August during which time a shift in resources to other NAICS codes could have occurred.

This percent difference method does a good job identifying spikes and plateaus in the time series plots. The EconEdit team also considered flagging estimates based on percent differences in the coefficients of variation (CVs), but since the CV is a function of the estimate and standard error, the team did not find that this added any new information. Additionally, it is possible for both the estimate and standard error to change drastically while the CV stays pretty much the same.

It is important the EconEdit team considered criteria besides estimate stability, namely the percent difference between the estimate and the previous year's final estimate, because estimates may be stabilizing around an unreasonable value. This is apparent in Figure 12, which plots 2014 estimates over time for NAICS code 3130 (manufacturing textile mills). The estimate and standard error increase greatly in early July and remain at unreasonably large values until analysts addressed the underlying data issue a month later in mid-August. Because of the third percent difference criterion that compares the estimates to the previous year's final estimate, the estimates during this month-long period are flagged appropriately.

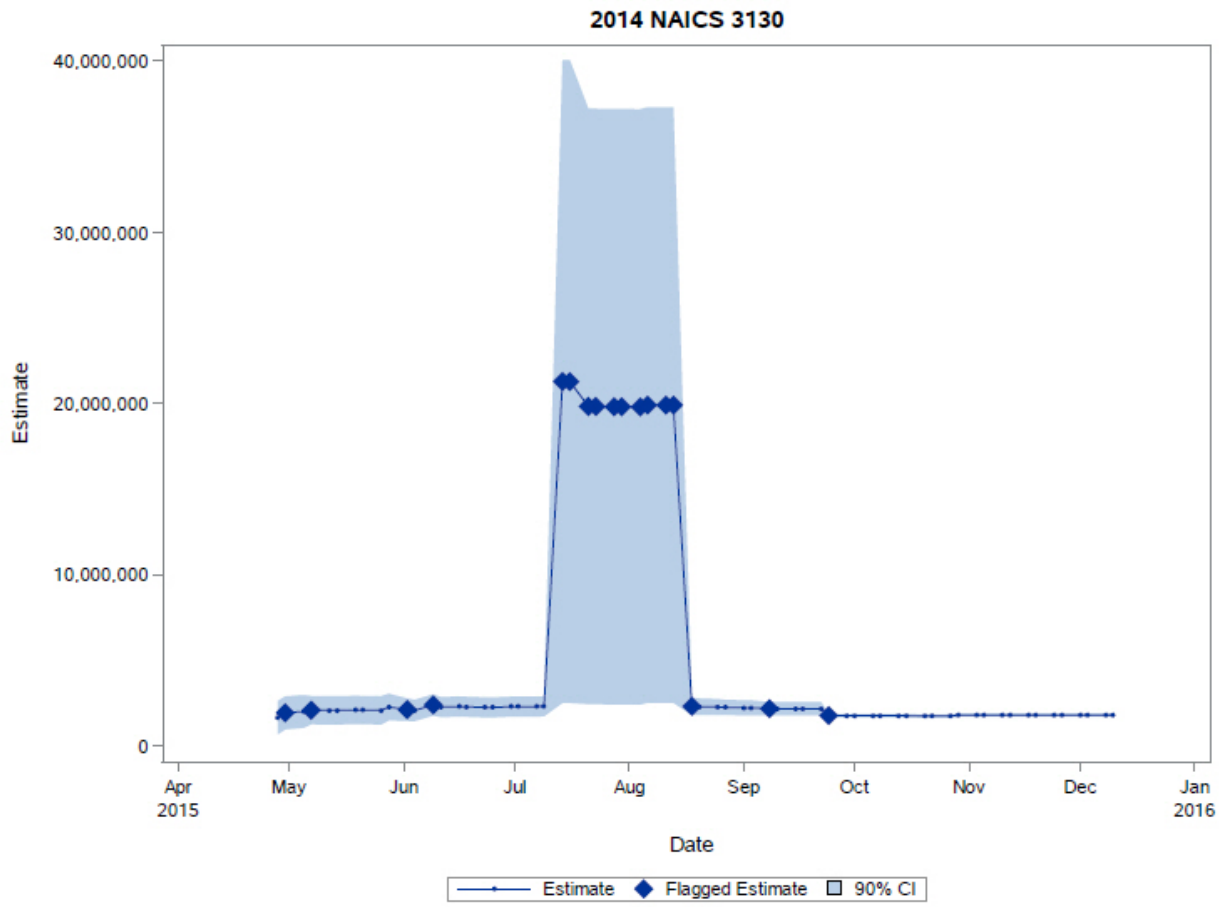


Figure 12: 2014 time series plot of estimates and 90 percent confidence band for NAICS code 3130 with estimates flagged according to the percent difference criteria. The estimate and standard error increase greatly in early July and stay flagged until the underlying data issue is addressed.

The team applied percent difference criteria to the estimates for each ERF in the ACES 2014 cycle. Figure 13 page plots the number of flagged estimates over time. As expected, the number of flagged estimates decreases to zero.

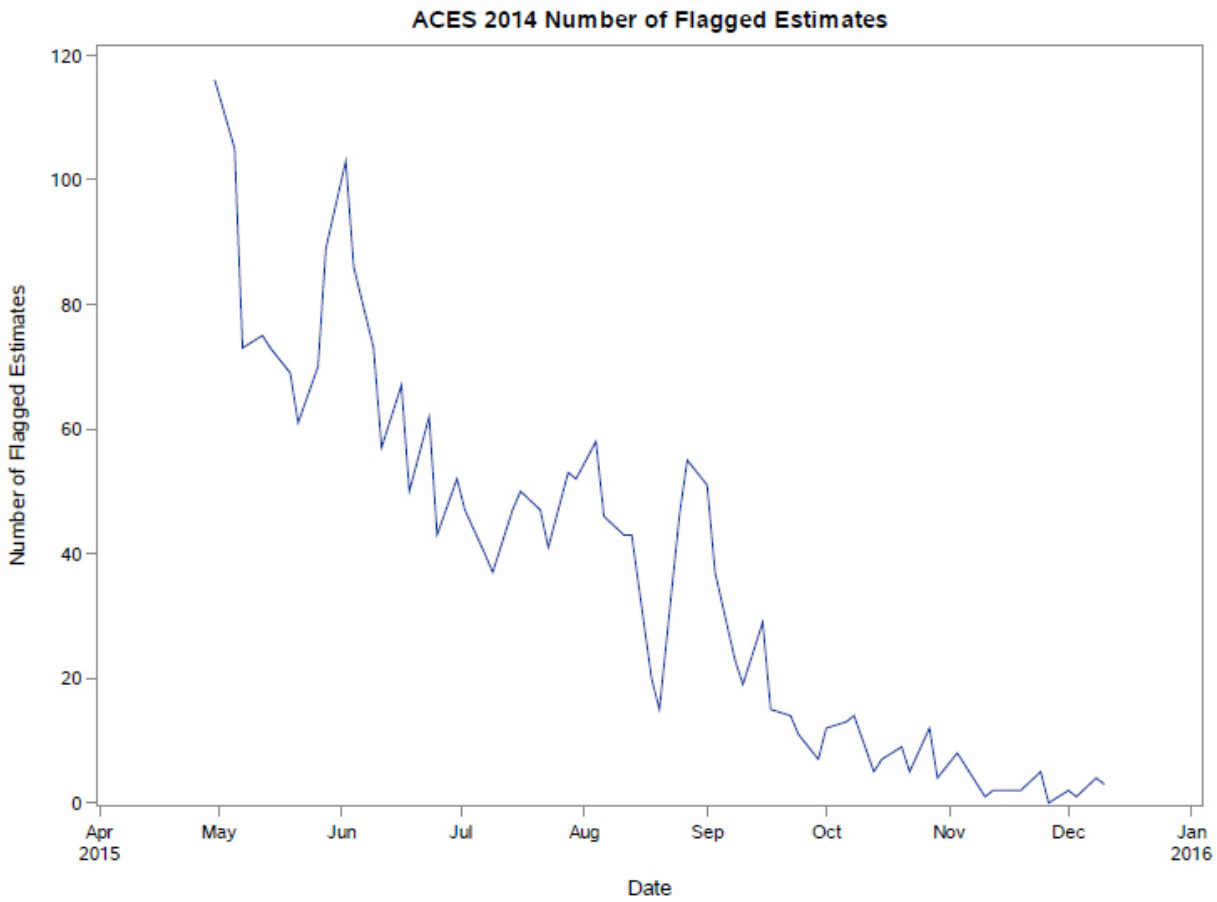


Figure 13: Aces 2014 time series plot of number of flagged estimates according to the three percent difference criteria. As expected, the number of flagged estimates decreases to zero.

8.2 Priority of NAICS Codes

When the flagging of multiple NAICS codes for data review at the same time, and it is important to prioritize which NAICS codes need to be reviewed. Continuing to look at ACES 2014 cycle, the EconEdit team flagged estimates in the ERFs according to the percent difference criteria from Section 8.1 and sorted the flagged four-digit NAICS code according to importance and discrepancy from the final 2013 estimates. A NAICS code was considered key if its percent of the overall total capital expenditures for 2013 was greater than 1 percent. Table 4 lists these 28 NAICS and their percentages. For example, according to the last ERF from 2013, NAICS code 2110 made up 11.8 percent of overall total capital expenditures.

Table 4: Key NAICS Codes According to Percent of Overall 2013 Total Capital Expenditures

NAICS	Percent of 2013 Total	NAICS	Percent of 2013 Total
2110	11.80	4599	1.58
2211	6.58	3110	1.51
5222	6.26	5221	1.51
6221	3.84	2212	1.30
5171	3.60	5241	1.29
5310	3.22	4520	1.28
5321	3.20	3344	1.22
5172	2.44	3251	1.20
3361	1.93	4840	1.18
2130	1.81	4810	1.17
7220	1.74	4820	1.12
4230	1.68	3240	1.11
6110	1.64	4240	1.03
5324	1.63	2380	1.03

The measure of discrepancy between the current year (CY) estimate and prior year (PY) final estimate is based on the ratio of the two and is given by the following formula:

$$\text{measure} = \max\left(\frac{estimate_CY}{estimate_final_PY}, \frac{estimate_final_PY}{estimate_CY}\right).$$

This measure assigns equal value to an estimate that is either larger or smaller than the final 2013 estimate by the same factor. As an example of assigning priority, Table 5 presents the priority ordering of the flagged NAICS for October 27, 2015, which is close to week 33. The criteria refer to the three percent difference criteria in Section 8.1 and explain why the NAICS was flagged.

Table 5: Priority Ordering of NAICS for October 27, 2015

Rank	NAICS	Key NAICS	Discrepancy Measure	Criterion 1	Criterion 2	Criterion 3
1	7220	1	1.54	1	1	0
2	4599	1	1.29	0	1	0
3	2380	1	1.26	0	1	0
4	3313	0	6.13	1	1	1
5	5242	0	2.35	1	1	0
6	3350	0	2.27	1	1	0
7	2131	0	1.73	0	1	0
8	5111	0	1.43	0	1	0
9	5615	0	1.39	1	1	0
10	5417	0	1.17	1	0	0
11	5418	0	1.11	0	1	0
12	4830	0	1.10	1	0	0

Section 9. Summary of Results

In examining estimates and other quantities over time, it became clear that around week 33 (October 29, 2015) many of the item estimates (at the four-digit NAICS level) stabilize and the associated confidence intervals contain the final estimate. Even when a data slide is introduced and the estimate inflates temporarily, the estimates tend to return to their previous value. This finding highlights two key points in data collection and editing for ACES. First, the nonresponse weighting adjustment seems to be working well. Second, in most cases the continuation of spending resources on edits past this point has diminishing value versus the cost.

However, there are cases where the estimate does not stabilize until near the end of the editing process. This would be cause for concern if the editing time were shortened without changing the analyst editing methodology. Perhaps a more efficient method would be to have the analysts assigned to NAICS whose estimates stabilize early in the process assist with more volatile industries. In a few cases, the estimates appear to stabilize early and then undergo a shift in later weeks. Many of these situations are likely due to a large company reporting late, a data slide that was not immediately corrected, or analysts waiting until the last day to make decisions about edits. Further investigation into what is causing estimates to shift in the final weeks should be conducted. Once the cause of these shifts is discovered, a change in methodology should be implemented to identify these changes earlier in the process.

The EconEdit team also examined "non-failure" edits. A sizeable number of non-failure edits were made to items 101 (gross depreciable assets) and 162 (gross domestic sales), which are not published but provided to BEA. Far fewer non-failure edits were made to 200-level items possibly to adjust data because of changes made to the roster table. This analysis of non-failure edits further demonstrated that data slides have a very large impact on ACES data. A large majority of the overall change in values in the data is attributable to data slides as opposed to other types of edits.

Through using the ranking metric $\frac{|\boldsymbol{newval}-\boldsymbol{oldval}|}{\boldsymbol{oldval}} \times \boldsymbol{ln}|\boldsymbol{newval}-\boldsymbol{oldval}|$, it became clear that the next most impactful edits after data slides in terms of change in value and in the number of edits has to do with the allocation of the "other" values to either structures or equipment. These allocations cost the analysts a significant amount of time. Another issue with these "other" categories is the way in which they are entered into StEPS. As it stands, analysts must go into the data and read the comments from the respondent and make decisions on how to allocate the values. Standard imputation practices may be difficult to implement under the current method of flags and notation. Perhaps the implementation of machine learning techniques would allow for the automation of imputation, thus saving even more analyst resources.

Section 10. Recommendations

While these recommendations are specific to ACES, they are most likely applicable to other surveys.

Recommendation 1: Update the questionnaire design to reduce data slide occurrences.

Data slides account for both the greatest impact on estimates and the most time spent by analysts. These errors are caused by respondents not reporting their values in the correct units. The most typical data slide involves respondents reporting in dollars instead of in thousands of dollars as requested. The number of data slide occurrences could be reduced by simply allowing respondents to report in dollars. This simple change to the questionnaire has resulted in considerable reductions in the number of data slides in other surveys.

Recommendation 2: Follow specific editing guidelines determined by the known impact of editing practices.

It has been shown that a very small proportion (5-15%) of the most impactful edits often contribute around 90% of the total change in the estimate and can get the estimate to within 1-2% of the final value (Granquist, 1997). For this reason, mathematical statisticians and analysts should work together to develop an editing hierarchy based on the impact of editing type. For example, it is clear that data slides have the greatest impact on ACES estimates. Therefore, analysts should focus on data slides first leaving other edits until data slides have been identified and corrected. This hierarchy can also help distribute editing resources more efficiently. Analysts who have corrected all data slides in their own NAICS industry or sector could be asked to help with other areas that are still working these cases. Working edits from greatest impact to least will help ensure that the most accurate estimate possible is created in the least amount of time.

Recommendation 3: Identify NAICS that need further review and set an edit priority order to these NAICS.

Analysts should identify NAICS that need further review and set a priority order for these NAICS. Table 5 gives one example of NAICS priority order methodology, which uses three percent differences as criteria to flag for further review as well as a measure of discrepancy to set priority order. This type of table could be used to redistribute resources to NAICS that are likely to need further edits, have historically had the largest influences on the final estimate, and have the largest prior to current year changes.

Recommendation 4: Make edit flags more descriptive and detailed.

One difficulty in this research was identifying exactly what edit an analyst had made. The current method of flagging edits only distinguishes between reported data, automatic edits, and analyst edits. To further understand what edits have the greatest impacts on data and estimates, new flags should be created and used to identify specifically the type of change that was made. Moreover, these new flags should include a flag indicating whether the change was made to correct a data slide. Data slides for ACES are corrected manually, and currently the only way of documenting them is through notes written by analysts. The EconEdit team had to develop its own criteria for determining data slides based on old and new values. See Appendix B for these criteria.

Recommendation 5: In order to conduct similar analyses for other surveys, save the survey's estimation results files (ERFs) and control files frequently, perhaps twice a week as is done with ACES.

Having access to estimates that have been calculated over time using production code that accounts for nonresponse allowed the EconEdit team to research when estimates and corresponding standard errors for ACES tend to stabilize. These are important factors to consider when determining when editing can stop. The EconEdit team echoes the sentiment of the 2005 Business Process Improvement Sub-team on Editing Efficiency that audit trails are very useful as well. The ERFs, control files, and audit trails are highly informative and complementary data sources for this type of editing research.

Recommendation 6: Provide systematic analyst feedback to help define and address specific questionnaire shortcomings.

Most economic surveys are self-administered, so survey or program staff analysts in the individual subject areas, rather than field representative interviewers, often have contact with respondents. Feedback from analysts about their interactions with respondents may serve as an informal evaluation of the questionnaire and the data collected. These interactions include "Help Desk" phone inquiries from respondents and follow-up phone calls to respondents by analysts investigating suspicious data flagged by edit failures. Analyst feedback is more useful when analysts systematically record comments from respondents in a log. The log enables qualitative evaluation of the relative severity of questionnaire problems, because strictly anecdotal feedback sometimes may be overstated.

Another way to obtain analyst feedback is to conduct focus groups with the analysts who review data and resolve edit failures. These focus groups can identify questions that may need to be redesigned or evaluated by other methods. Regardless of how respondent feedback is captured, analysts should provide feedback early in the questionnaire development cycle of recurring surveys to identify problematic questions.

Recommendation 7: Research applications of Big Data techniques such as machine learning to aid editing.

The audit trails and historical data contain many variables that can be used as predictors in machine learning to identify influential observations such as data slides. Big Data selective editing techniques based on visualizations known as tableplots (Tennekes *et al.*, 2013) should also be researched to possibly identify potential influential observations.

Section 11: Future Research

This section describes future research that the EconEdit team would like to pursue.

Apply similar methods to previous years of ACES and obtain data for use in modeling stopping points.

Determining how the stability of estimates is associated with other factors has helped the EconEdit team understand what quantities are important to monitor when trying to determine a stopping point. Applying the same methodology to previous years of ACES data and obtaining more data on the relationships between these quantities and stopping points are additional steps in the modeling process.

Apply similar methods to the Quarterly Services Survey (QSS) and the Services Annual Survey (SAS).

The EconEdit team has obtained QSS "snapshot" data for quarters 2014q3-2016q3. These data include weekly revenue estimates for each publication level starting at or around publication day 30 and running until about day 65. The QSS team is working to provide the EconEdit staff with audit trails for the same quarters. The EconEdit team will then use similar methodology to what was used with ACES to evaluate the editing process for QSS. SAS is a large survey with a variety of variables, so there is potential for studying Big Data selective editing techniques such as tableplots. QSS is a subsample of SAS. Therefore, the EconEdit team plans to use lessons learned in evaluating QSS to help navigate the size and complexity of SAS.

Determine the impact of micro edits to create a hierarchical editing system.

While it is clear that data slides have the largest impact and that moving values from the "other" items also has an influence, it is not known how much of an effect other edits have. These other micro edits may have no discernable impact on estimates at all. Knowing which edits have the most influence will allow for the creation of a hierarchical system of editing that would allow the most impactful edits to be handled first, thus stabilizing estimates much sooner in the process. However, to accomplish this research, edit flags need to be made more descriptive and detailed.

Investigate possibility of automating certain edit types.

It is believed that some of the edits currently made by analysts could be automated. Many other surveys in the Economic Directorate utilize imputation methods for specific variables. ACES might be able to automate such things such as the distribution of values in the "other" categories with imputation techniques or fixing adding-complex errors with raking. Further investigation into these possibilities and others should be pursued.

References

Company Statistics Division. (2002). Annual Capital Expenditures Survey Quality Profile 2002. Dated August 2002. EDMS #219142.

Editing Efficiency Team. (2005a). Business Process Improvement Sub-team on Data Editing: Supporting Documents. EDMS #35530.

Editing Efficiency Team (2005b). Recommendations of the Business Process Improvement Sub-team on Editing Efficiency. EDMS #35529.

Granquist, L. (1997). The New View on Editing. *International Statistics Review, 65*(3), 381-387.

Granquist, L. and Kovar, J.G. (1997). Editing of Survey Data: How Much is Enough? *Survey Measurement and Process Quality*, 415-435.

Greenberg, B. and Petkunas, T. (1986). An Evaluation of Edit and Imputation Procedures used in the 1982 Economic Censuses in Business Division. SRD Research Report Census/SRD/RR-86/04. EDMS #239525.

Hidiroglou, M.A. and Bertholot, J.M. (1986). Statistical Editing and Imputation for Periodic Business Surveys. *Survey Methodology*, 12, 73-83.

King, C. and Kornbau, M. (1993). Inventory of Economic-Area Statistical Practices Phase 1: Program Requirements and Survey Design. ESMD Report Series ESMD-9306. EDMS #182375. November 1993.

Newman, B. (2015). Annual Capital Expenditures Survey: Identifying Possible Stopping Points for Data Collection and Editing.

Puts, M., Daas, P., and de Waal, T. (2015). Finding Errors in Big Data. *Significance*, The Royal Statistical Society, June 2015.

Tennekes, M., de Jonge, E., and Daas, P.J.H. (2013). Visualizing and Inspecting Large Datasets with Tableplots. *Journal of Data Science, 11*, 43-58.

Weng, C.F. (2015). New Methodologies for Identifying Response Error. 2015 Federal Committee on Statistical Methodology Research Conference.

Appendix A. Editing Terms and Definitions

Table A1 contains commonly used editing terms and their definitions. Many definitions are taken from Puts *et al.* (2015).

Table A1: Editing Terms and Definitions

Term	Definition
Automatic editing	Editing in which data are adjusted by computers without human interaction.
Balance edit	An accounting-type edit that ensures values at a detailed level sum to values at a more aggregated level.
Bivariate indexing editing	Editing method for identifying potential errors based on two criteria: relative difference and relative importance. Relative difference, also known as suspicion, relates to the probability of a true error. Relative importance, also known as effect, relates to the impact that the error has on final estimates. See Cheng (2015) for more information.
Data slide	Edit that corrects data reported in the wrong units. The most common data slide for ACES involves dollar figures reported in dollars instead of thousands of dollars.
Hidiroglou-Berthelot editing	A widely used manual statistical editing method designed for periodic business surveys that is based on ratios and requires the selection of several parameters.
Interactive editing	A special case of manual editing in which the effects of adjusting data can be seen immediately.
Macro-editing	Editing that checks whether the dataset as a whole is plausible. In general, macro-editing involves comparing aggregates and checking the distribution of values. This type of editing is performed when most of the data have been collected.
Manual editing	Editing in which all data observations are corrected in detail. Data tend to be checked and adjusted in separate steps. This process is time-consuming and often involves cycles during which observations are examined and adjusted multiple times.
Robust regression editing	Editing method based on examining residuals from a robust regression of the current time period's data against the previous time period's data. See Cheng (2015) for more information.

Selective editing	Editing in which interactive editing is restricted to observations identified as having the most influential errors. Influence typically is quantified by score functions.
Stopping rule	Statistical rule for determining when the editing process can be stopped without affecting the quality of the final estimates appreciably.
Tableplot	A visualization that aids in selective editing of large datasets and in monitoring the effects of the editing process on the quality of the data. See Puts *et al.* (2015) for details and examples.
Unit flag	A flag that identifies edit rule failures at the unit level. The flag might indicate that: i. Edit rules could not correct failure and must be corrected by an analyst ii. Edit rules resolved the failure but must be reviewed by an analyst to either accept edit action or correct data iii. Edit rules could not correct failure and must be corrected by an analyst iv. Nonresponse for data item(s) that must be reviewed by analyst to either accept edit action or correct data v. Edit rules resolved the failure but does not need to be reviewed

Appendix B. Research-Specific Variables and Flags

Figure B1 is a screenshot of the data fields for the 200- and 300-level items on the ACES questionnaire.

ITEM 2 **CAPITAL EXPENDITURES**

Report the following domestic capital expenditures data for the entire company. *(Refer to page 4 of Instructions)*

Example: if figure is $1,179,125,628.00 report → Bil. 1 Mil. 179 Thou. 126

Row	CAPITAL EXPENDITURES *(Refer to Page 2 of Instructions)*	Structures (1) Bil. Mil. Thou.	Equipment (2) Bil. Mil. Thou.	Other *(Describe in Item 3)* (3) Bil. Mil. Thou.	Total *(Add columns 1+2+3)* (4) Bil. Mil. Thou.
20	Capital expenditures for NEW structures and equipment ***(Include major additions, alterations, and capitalized repairs to existing structures)***	item 201	item 202	item 203	item 204
21	Capital expenditures for USED structures and equipment	item 211	item 212	item 213	item 214
22	**TOTAL capital expenditures** ***(Add Rows 20 + 21)***	item 221	item 222	item 223	item 224
					Total should equal Item 1A, Row 11

ITEM 3 **List the items included in "Other."** Report in thousands of dollars. **Furniture and fixtures, computers, capitalized computer software, and motor vehicles** should be reported as equipment. **Leasehold improvements** should be considered new structures or new equipment based on what is being improved.

Row	(1) Description of Capital Expenditures	(2) Bil. Mil. Thou.
30	item 301	item 302
31	item 311	item 312

Figure B1: ACES Questionnaire 200- and 300-Level Items

Figure B2 is a screenshot of the data fields for the 600-level items on the ACES questionnaire. This table is known as the roster and breaks down total capital expenditures by NAICS code.

ITEM 6 **CAPITAL EXPENDITURES BY INDUSTRY**

Complete Item 6 for each industry in which the company had operations and made capital expenditures in 2015. *(Refer to page 5 of the Instructions.)* Please refer to the complete list of possible industry codes and descriptions beginning on page 6 of the Instructions, Definitions, and Codes List manual.

STRUCTURES + EQUIPMENT + OTHER = TOTAL

Industry Category Code	Structures (New (2), Used (3))	Equipment (New (5), Used (6))	Other (New (8), Used (9))	TOTAL CAPITAL EXPENDITURES (0)
RKEY1	Roster 612	Roster 615	Roster 618	Roster 610
RKEY1	Roster 612	Roster 615	Roster 618	Roster 610
*69 COLUMN TOTALS	Item 0D612	Item 0D615	Item 0D618	Item 0D610
	Should equal Item 2, Row 20, Column 1	*Should equal Item 2, Row 20, Column 2*	*Should equal Item 2, Row 20, Column 3*	***Should equal Item 2, Row 22, Column 4***
	Should equal Item 2, Row 21, Column 1	*Should equal Item 2, Row 21, Column 2*	*Should equal Item 2, Row 21, Column 3*	

Total Struct Roster derived 611

Total Equip Roster derived 614

Total Other Roster derived 617

Roster 613 / Roster 616 / Roster 619

***NOTE:** If you need additional space to report more industries, please use the enclosed continuation sheet. Row 69, "Column Totals", should equal the sum of the columns above plus the sum of the columns on the continuation sheet(s).

Figure B2: ACES Questionnaire 600-Level Items (Roster)

Below is the criteria that the EconEdit team used to determine data slides. The variable *ds* is the data slide flag, and the variable *ds_factor* is the data slide factor, which is a power of ten. The criterion involving the max function is intended to exclude cases with very small values and is most relevant to the case of a suspected data slide with *ds_factor* equal to ten. The most common data slide in ACES involves values reported in dollars instead of thousands of dollars. To correct this, the data value is reduced by a factor of 1,000.

```
if oldval = . then oldval_re = 0;
else oldval_re = oldval;
if newval = . then newval_re = 0;
else newval_re = newval;

if (newval_re ne 0) and (oldval_re ne 0) and
     (usrnme_re ne "stpsprod") then do;
     if ((ceil(newval_re/10^8) = oldval_re) or
          (ceil(oldval_re/10^8) = newval_re) or
          (floor(newval_re/10^8) = oldval_re) or
          (floor(oldval_re/10^8) = newval_re)) and
          max(newval_re/(10^8*oldval_re),
               oldval_re/(10^8*newval_re)) > 0.9
          then do;
          ds = 1;
          ds_factor = 10^8;
     end;
     …
     …
     *Similar if/else statements for other powers of ten;
     …
     …
     if ((ceil(newval_re/10) = oldval_re) or
          (ceil(oldval_re/10) = newval_re) or
          (floor(newval_re/10) = oldval_re) or
          (floor(oldval_re/10) = newval_re)) and
          max(newval_re/(10*oldval_re),
               oldval_re/(10*newval_re)) > 0.9
          then do;
          ds = 1;
          ds_factor = 10;
     end;
end;

if ds ne 1 then ds = 0;
```

Table B1 on the next page lists preconditions and criteria for the 29 item edit failures. The EconEdit team used these criteria to set its own edit failure flags.

Table B1: Item Edit Failure Criteria

Failure	Precondition	Criteria
1	STRATUM != 3	223 != 0 OR 213 != 0 OR 203 != 0 OR 302 != 0 OR 312 != 0
2	STRATUM != 3	141 != 101 + 111 + 121 - 131
3		204 != 201 + 202 + 203
4		214 != 211 + 212 + 213
5		221 != 201 + 211
6		222 != 202 + 212
7		224 != 204 + 214
8		224 != 221 + 222 + 223
9	STRATUM in (3, C, F)	224 != 201 + 202 + 211 + 212
10	STRATUM in (3, C, F)	411 > 201 + 202
11	STRATUM in (3, C, F)	504 != 501 + 502 + 503
12		101 < 0 OR 111 < 0 OR 121 < 0 OR 131 < 0 OR 141 < 0 OR 151 < 0 OR 162 < 0 OR 172 < 0 OR 182 < 0 OR 192 < 0 OR 201 < 0 OR 202 < 0 OR 203 < 0 OR 204 < 0 OR 211 < 0 OR 212 < 0 OR 213 < 0 OR 214 < 0 OR 221 < 0 OR 222 < 0 OR 223 < 0 OR 224 < 0 OR 302 < 0 OR 312 < 0 OR 411 < 0 OR 501 < 0 OR 502 < 0 OR 503 < 0 OR 504 < 0
13	STRATUM NOT in (3, C, F)	111 != 224
14	STRATUM NOT in (3, C, F) AND CKNDTE != .	111 = . AND 141 = .
15	STRATUM NOT in (3, C, F)	151 > 141
16	STRATUM NOT in (3, C, F)	162 < 172 + 182 + 192
17	TEMP > 2000 employees	224 = $0
18	STRATUM in (3, C, F)	224 > $150
19		411 > 204
20	STRATUM in (3, C, F)	411 > $15
21	STRATUM NOT in (3, C, F)	504 > 202
22	STRATUM NOT in (3, C, F)	504 > 222
23	STRATUM NOT in (3, C, F)	224 != 610
24	STRATUM NOT in (3, C, F)	201 != 612
25	STRATUM NOT in (3, C, F)	211 != 613
26	STRATUM NOT in (3, C, F)	202 != 615
27	TEMP > 2000 employees	212 != 616
28	STRATUM NOT in (3, C, F)	962 = 1
29	STRATUM NOT in (3, C, F)	971 = 1

Note: The notation != stands for “not equals” and . for “missing.”

Appendix C. Additional Tables and Plots

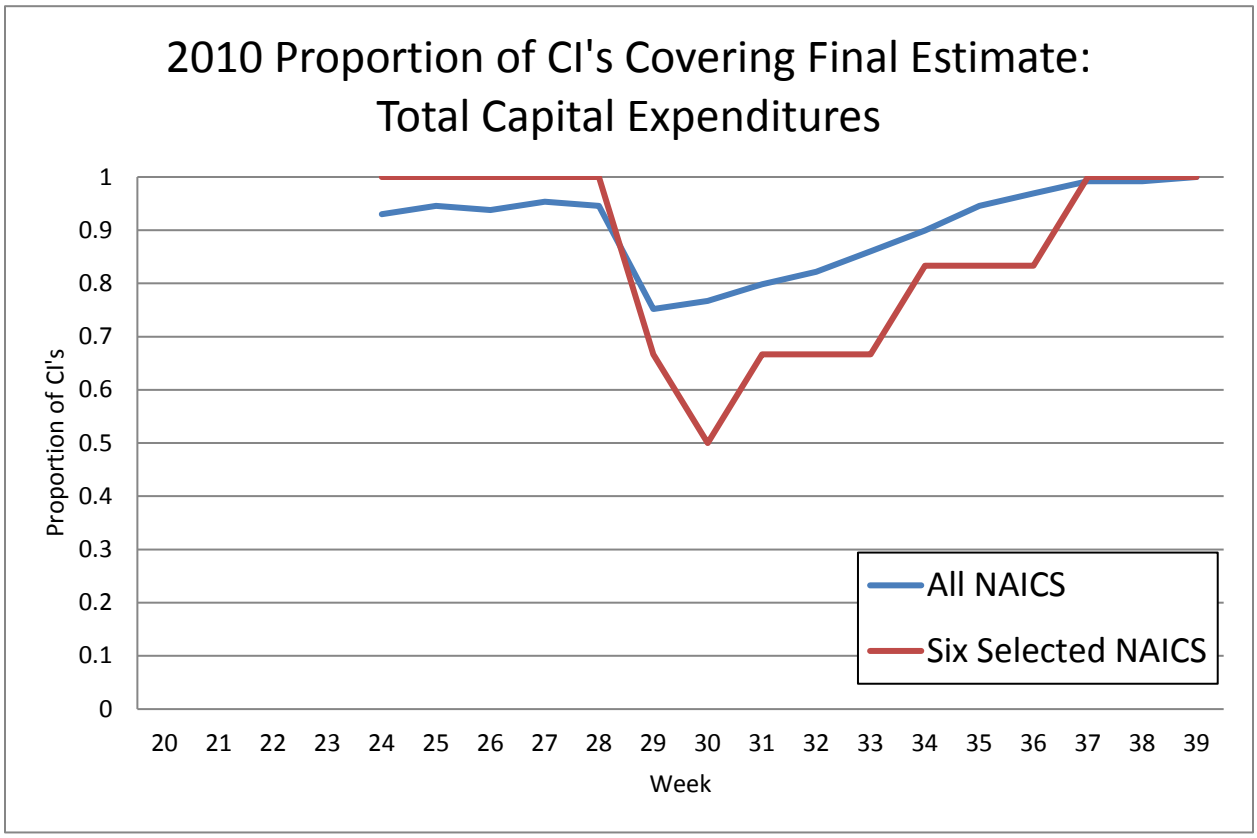


Figure C1: Plot of proportion of 90 percent confidence intervals covering final estimates over the ACES 2010 Cycle for item 610 (total capital expenditures).

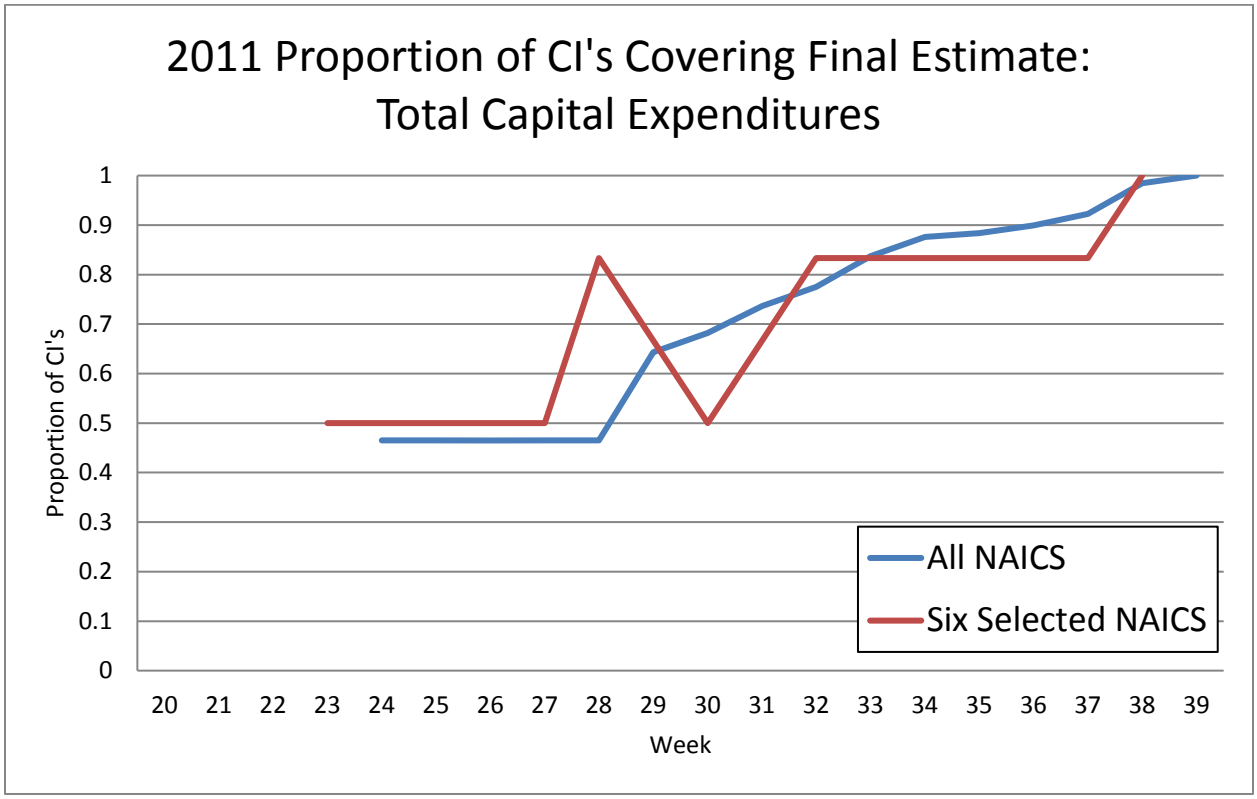


Figure C2: Plot of proportion of 90 percent confidence intervals covering final estimates over the ACES 2011 Cycle for item 610 (total capital expenditures).

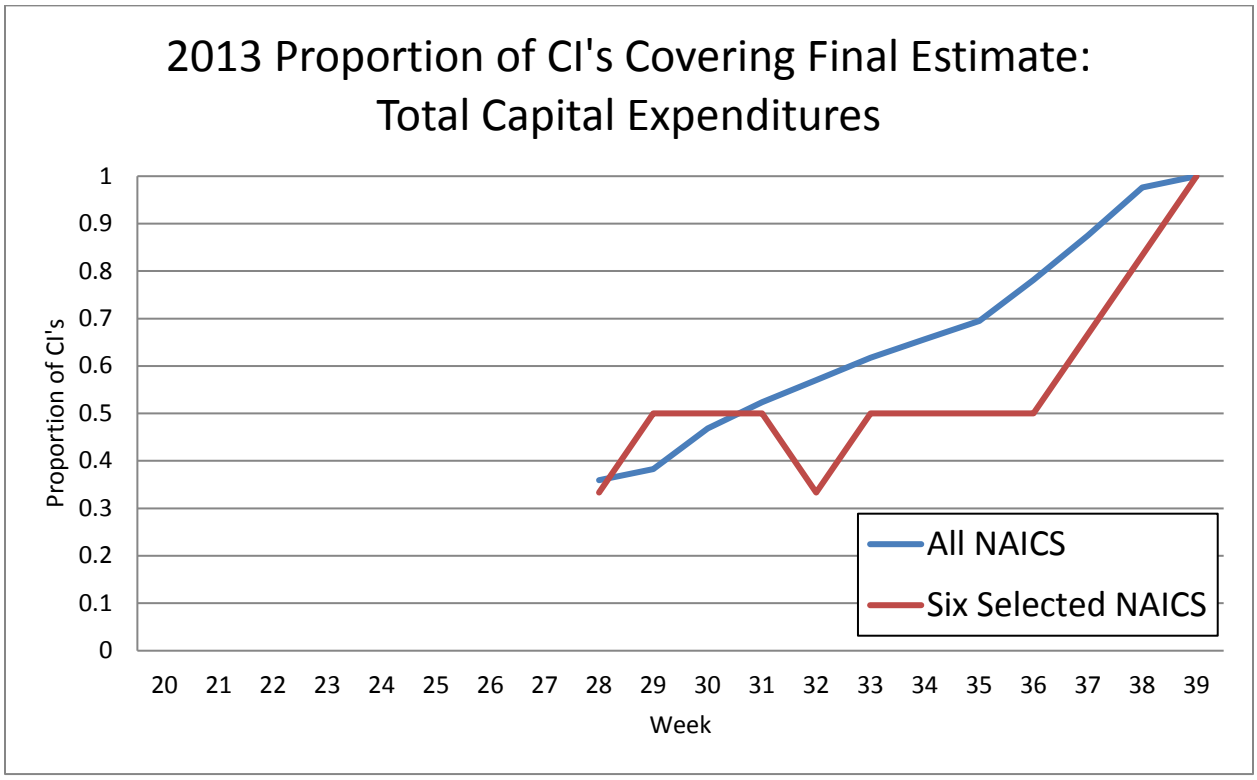
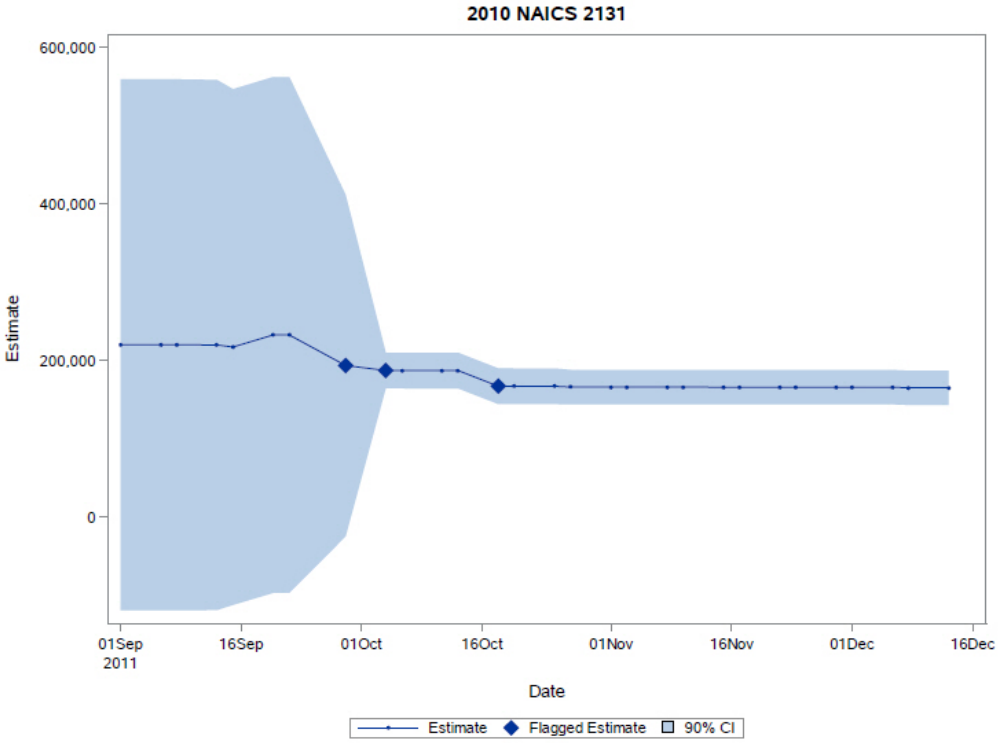


Figure C3: Plot of proportion of 90 percent confidence intervals covering final estimates over the ACES 2013 Cycle for item 610 (total capital expenditures).[2]



[2] ACES Survey Year 2012 has been omitted due to 2013 Government Shutdown impact on editing

Figure C4: 2010 plot of estimates over time and 90 percent confidence band for NAICS 2131 with estimates flagged according to percent difference criteria.

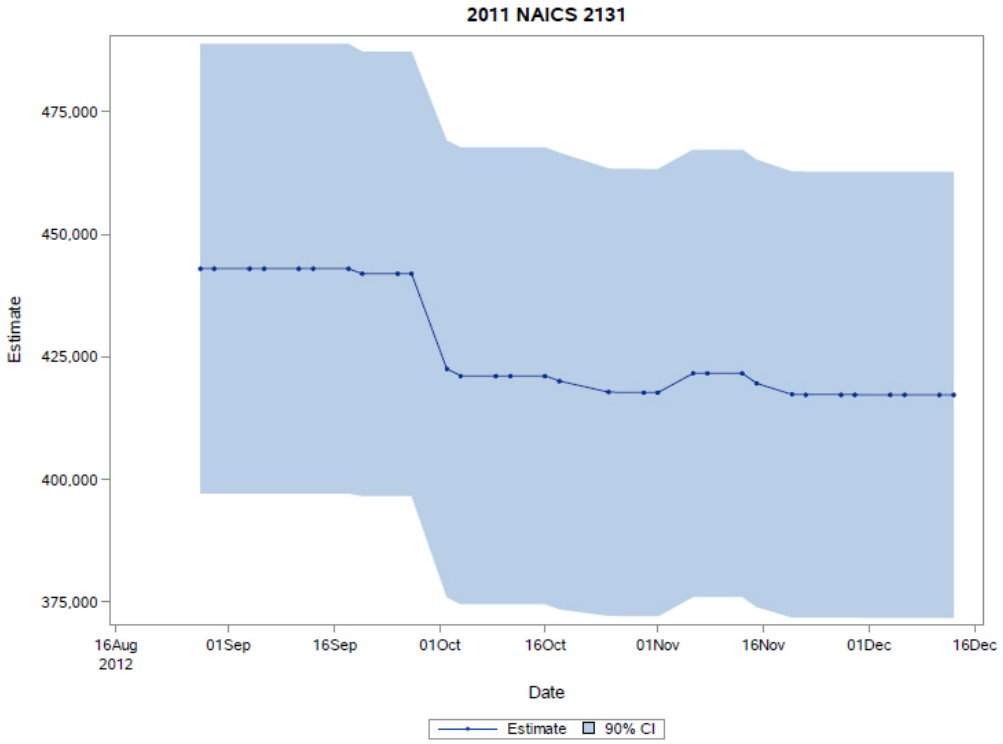


Figure C5: 2011 plot of estimates over time and 90 percent confidence band for NAICS 2131. According to the percent difference criteria, no estimates were flagged as needing review.

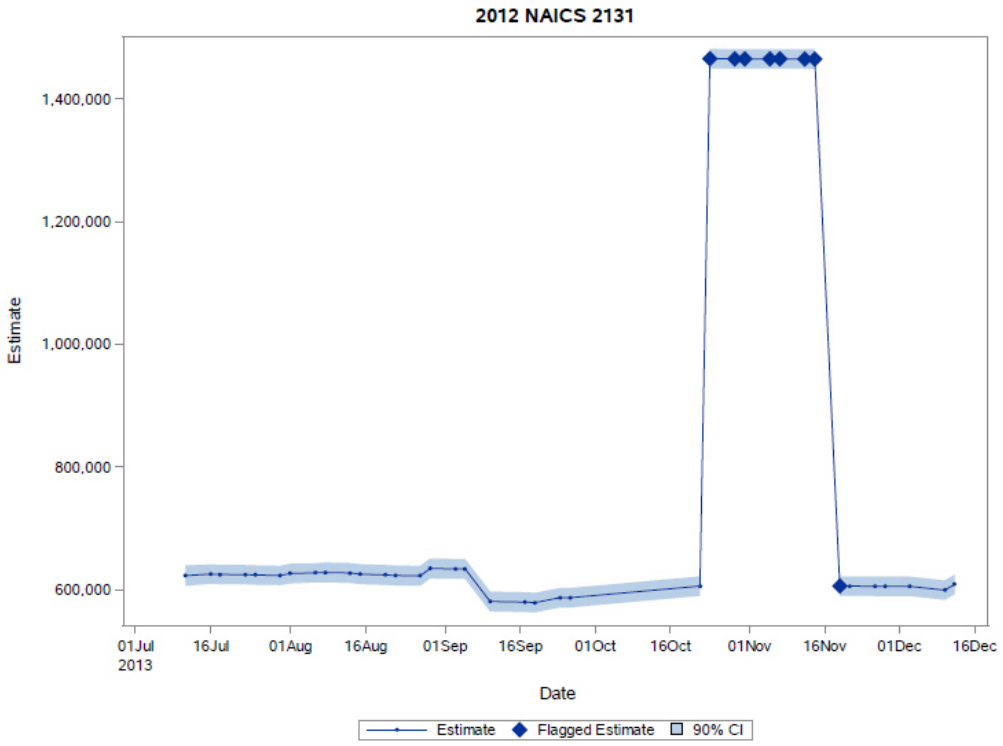


Figure C6: 2012 plot of estimates over time and 90 percent confidence band for NAICS 2131 with estimates flagged according to percent difference criteria.

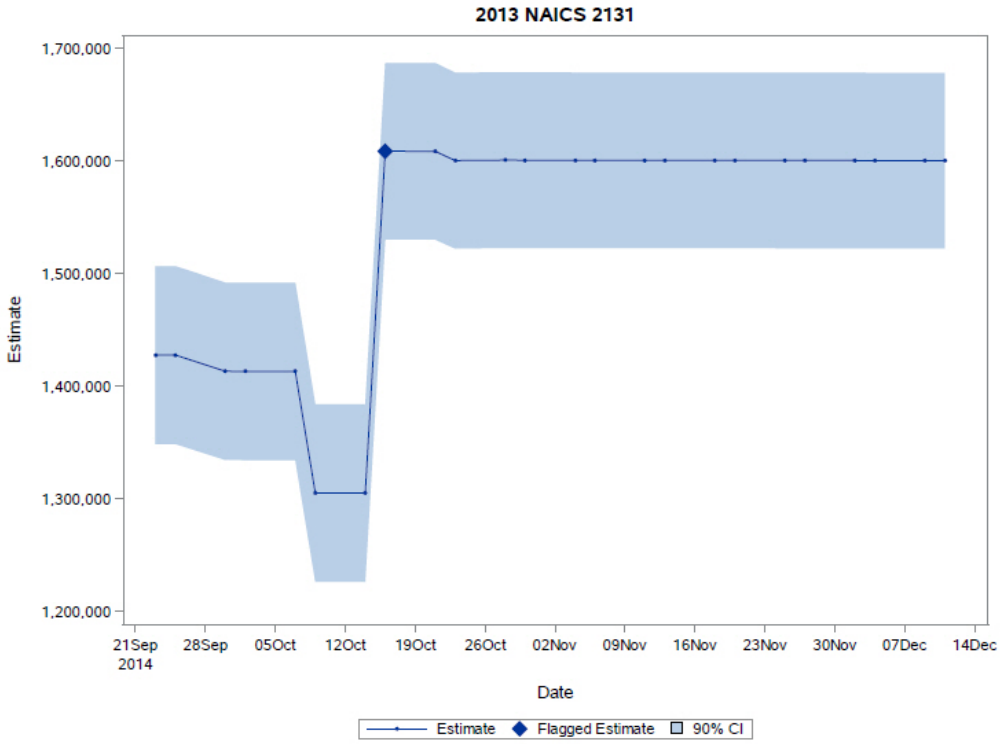


Figure C7: 2013 plot of estimates over time and 90 percent confidence band for NAICS 2131 with estimates flagged according to percent difference criteria.