

Census Scientific Advisory Council

# Edit Reduction Efforts in the Economic Directorate

March 31, 2017

# Introduction to Editing in the Economic Directorate

- From monthly, quarterly, annual, and quinquennial surveys and censuses, the Economic Directorate collects, processes, and releases data on Retail Sales, Wholesale Sales, Services, Manufacturing, Construction, International Trade, and Governments.
- After data collection, the data are typically analytically reviewed, corrected (edited), and imputed prior to estimation, variance estimation, benchmarking, seasonal adjustment, etc.
- Data review
  - May be manual (by an analyst) or automated (system)
  - May be at an item level (micro) or aggregated level (macro)
  - Should improve the validity of the estimates
- Edits
  - Check internal form consistency
  - Trend checks
  - Additivity

# State of Data Review Processes in the Economic Directorate

It's a mix!

- From analyst reviews of individual items to Standardized Economic Processing System (StEPS)
- From ratio edits with non-data-driven bounds to bounds determined by proven statistical methods (Hidiroglou-Berthelot, resistant fences, asymmetric fences, etc.)
- From regular evaluations of edits to no evaluations of how well the edits have done
- From standardized detailed flags for monitoring the process to minimal flags that tell little
- From audit trails and timely snapshots of the estimation system throughout the process to no audit trails

What we know:

- Editing processes and procedures are not standardized.
- Editing using the current procedures takes a lot of time.
- Editing generally stops when the allotted time runs out.

# Motivation for the Edit Reduction Effort

- Over-editing affects timeliness
- Over-editing may introduce bias
- Editing is costly
- Editing processes must be accurate and repeatable
- Identifying and dealing with outliers is necessary, but detailed microdata cleanup may not be

# Purpose of the research

- To identify areas of improvement in our editing processes in order to improve the timeliness and quality of our estimates while reducing our cost
  - Find stopping rules for when “good enough is good enough to stop editing”
  - Identify areas that need to be automated
  - Find question changes that can be made that will alleviate errors
  - Identify edits that can be dropped

# Research Questions

- Are Econ Surveys able to stop the editing process earlier?
- If so, are there clear “stopping points” where data quality remains high while saving resources on edits?
- If there are no clear “stopping points,” what steps can be made to determine stopping points in the future?
- What steps can be taken to allow the development of a model to signal when to discontinue edits?

# Annual Capital Expenditures Survey Edit Reduction Project – Background

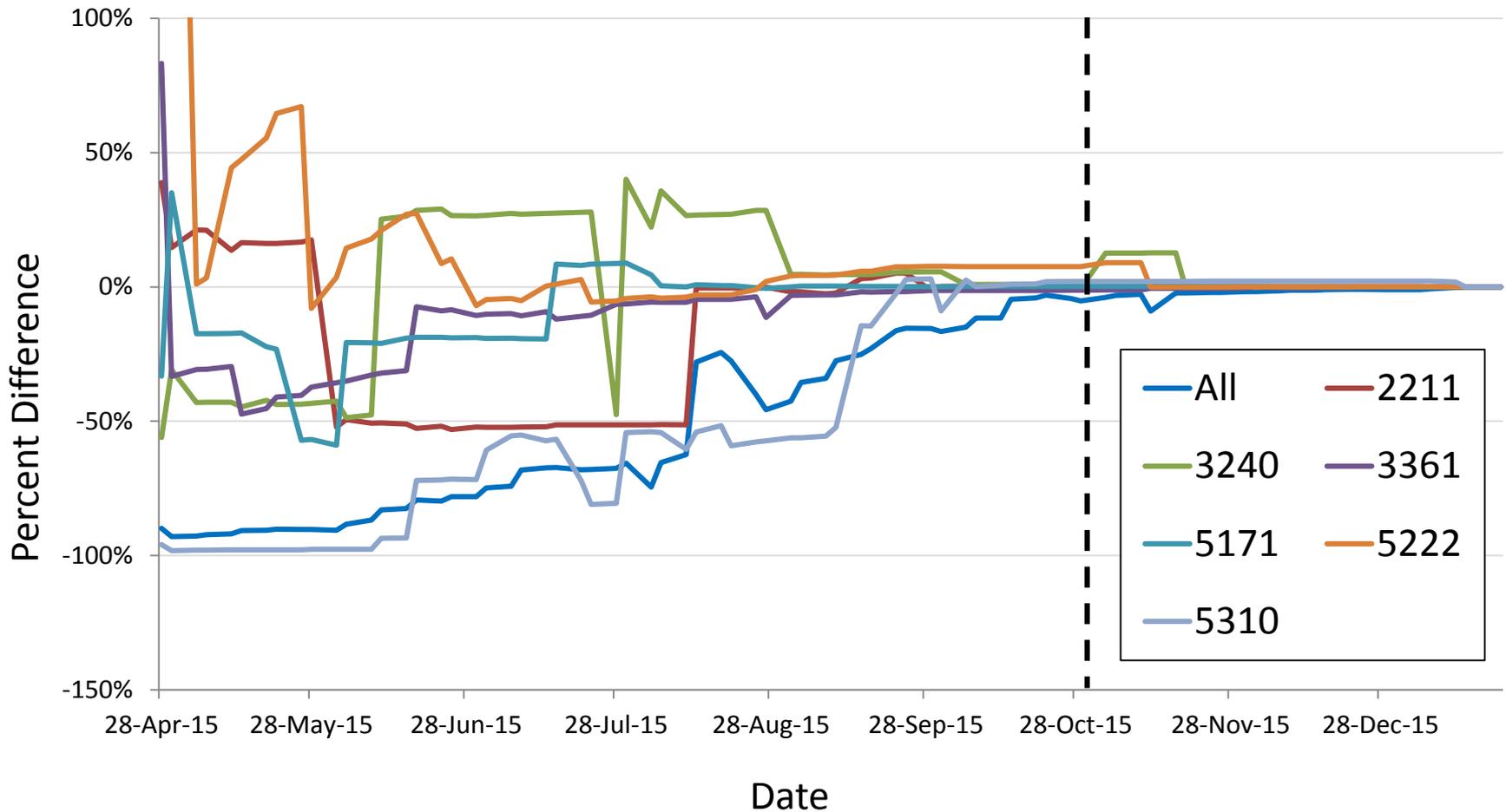
- This survey provides data on capital spending for new and used structures and equipment by U.S. nonfarm businesses.
- Capital expenditure is not highly correlated with any other variable, so current to prior edits are usually not helpful.
- Data are supposed to be reported in thousands, but respondents often report in dollars (data slides).
- Securities and Exchange Commission (SEC) filings are available for public companies and can be used for missing data.

# Annual Capital Expenditures Survey Edit Reduction Process – Approach

We conducted a series of experiments on 2014 ACES data

- To examine raw sums, estimates, standard errors, and the number of edit failures over time
- To understand the nature of edits that are made to adjust data but not to correct for edit failures.
- To quantify the impact of editing on estimates by NAICS and by edit type

# Percent Difference Between Current and Final Estimates for Selected NAICS Codes for 2014 ACES Processing



# Annual Capital Expenditures Survey Edit Reduction Process – Results/Findings

- The most impactful edits are edits that can be nearly eliminated by simple changes to the questionnaire
- Many item estimates for ACES (at the four-digit NAICS level) stabilize and are within final confidence interval limits 15 weeks earlier than they are normally released
- Moving editing resources from stabilized items to those that still require review could bring all estimates into the final confidence intervals earlier – adaptive editing
- In several instances, analysts changed data for two items that did not fail an edit which may indicate that perhaps an edit was needed for these items
- Readily available administrative data could be used earlier for some variables.

# Challenges

- IT storage
- Sometimes a lack of skills for dealing with large audit trails
- Lack of standardized procedures: flags, evaluations, and audit trails

# Economic Directorate Edit Reduction Project

## Next Steps

- We have examined ACES data from 2010, 2011, and 2013 and found similar results. We omitted 2012 because of the government shutdown which occurred during the ACES data review. The final report will be issued in March
- We are currently examining a quarterly survey, the Quarterly Services Survey
- We will eventually look at monthly surveys and the census

# Economic Directorate Edit Reduction Project

## Next Steps, cont.

- We will make recommendations for standardization of editing procedures, audit trails, and edit flags
- We have proposed and will work towards implementing stopping rules for when to stop editing certain ACES NAICS codes and switch resources to other codes in a more adaptive design
- We will research Big Data editing methods to see if those methods can be used