

## **Developing Macro Edits for the Census and Annual Survey of Governments Finance**

Loretta A. McKenzie, Terri L. Craig, Jennifer N. Whitaker  
Governments Division  
U.S. Census Bureau<sup>1</sup>, Washington D.C. 20233

**Abstract:** The Census Bureau conducts a Census of Governments Finance every five years and an annual sample survey in the intercensal period. For the 2006 annual survey, a macro edit procedure was developed to isolate suspicious data at the aggregate level. Simple parameters were established based on subject-matter expertise. For the 2007 Census, the Hidioglou-Berthelot method of determining edit bounds (H-B method) was used as a macro editing procedure. Both methods result in the identification of data that require manual review and analysis. This paper compares the methods used in the two years and provides recommendations for future applications.

**Key words:** Macro edits, Hidioglou-Berthelot Edit

### **1. Introduction**

For most surveys, data are edited before information is presented so they are accurate and consistent. The first stage of editing consists of a micro (unit-by-unit) review of individual units. These procedures are done on a flow basis as the forms are received, and often result in a lot of false edit failures to review. This paper presents the results of research undertaken to develop a systematic, statistical review of macro data (tabulation). Macro editing procedures follow micro editing. Macro editing detects errors by analyzing the weighted data aggregates. Various methods exist for establishing macro edits. Data totals are often compared with earlier versions of the same data (historical edit), but different totals within the current collection period (e.g., revenue to expenditures) can also be compared. The macro edit parameters can be obtained by analysts expertise or through a statistical analysis of the distribution of the aggregates.

The Census Bureau conducts a Census of Governments Finance (COGF) every five years and a sample survey in the intercensal period. For the 2006 Annual Survey of Governments Finance (ASGF), a team of analysts, computer specialists, and mathematical statisticians was organized to identify typical state and local government survey estimates or census totals. This group developed a macro edit procedure to isolate suspicious data at the aggregate level that relied on simple ratio comparisons of current to prior tabulations using edit parameters based on subject-matter expertise. This methodology is described in detail in Section 3.1 and is hereafter referred to as the “simple parameter” method. A limitation of this macro editing approach is its subjectivity and reliance on subject-matter expertise. Consequently, for the 2007 COGF, we researched and implemented the Hidioglou-Berthelot statistical data edit, hereafter referred to as the H-B method (Hidioglou and Berthelot, 1986). This paper presents our research and analysis.

---

<sup>1</sup>

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

The following section gives background information on the COGF and the ASGF. Section 3 discusses the development of simple parameters and the H-B method. Macro editing methodology for the 2006 ASGF and the 2007 COGF are discussed in Section 4. Conclusions, recommendations, and topics for further research are presented in Section 5.

## **2. Survey Background**

The ASGF releases detailed financial data on revenue, expenditures, debt, and assets from state and local governments (counties, cities, townships, special districts, and independent school districts) each year. There are statistics for the 50 state areas and the District of Columbia, as well as a national summary. Statistics are also available by level of government (all states, all locals, and all states plus all locals aggregates). The survey mails out in October and requests data for the government's fiscal year ending between July 1 of the previous year and June 30 of the survey year.

In years ending in '2' and '7', a census of all state and local governments is conducted. For the 2007 COGF, 89,527 governments were counted, of which, 89,476 were local governments. In the intervening years, a sample of counties, cities, townships, and special districts is surveyed. All independent school districts and state governments are surveyed each year. For 2006, the total sample size including all types of government (state, counties, cities, townships, special districts, and independent school districts) was 25,978.

Counties, cities, and townships are called general purpose governments because they perform multiple functions (education, transportation, public safety, etc.). Special district governments are fiscally and administratively independent entities that generally have one purpose and are referred to as single function special districts. These governments include water districts, fire protection districts, and airport authorities among others. Each special district is categorized by a function code that indicates its function. There are some special districts that have multiple functions such as sewer and water supply or natural resources and water supply. These are referred to as multi-function special districts.

The items collected on the questionnaire are the same for census and sample years. Data on several hundred variables are provided by state and local governments. The variables include various taxes, charges, assessments, intergovernmental revenue, utility revenue, liquor store revenue, and insurance trust revenue. Capital outlay and current operation expenditures are collected for various functions (education, transportation, public safety, etc.). Long-term and short-term debts, as well as data on various cash and security holdings, are also included in the data collection. For more information on the survey go to <http://www.census.gov/govs/www/financegen.html>. For information on the sample design, see "Sampling Procedure for the Annual Survey of Government Finance" by McLaughlin and McKenzie (2006) which is available upon request. Results from the 2007 Census of Governments Finance were released in September 2009.

## **3. Methodology**

Two methods of doing macro edits are compared in this paper: simple parameters and the H-B method. These methods are described in Sections 3.1 and 3.2 respectively.

### **3.1 Simple Parameters**

Economic surveys collect quantitative data with highly skewed distributions. A few units can contribute substantial amounts to the table totals. Consequently, it can be a large benefit to survey methodologists to isolate divergent totals and the units that contribute the most to these unexpectedly high or low totals, in a cost and time effective manner using little effort. Macro editing is an additional validation at the aggregate level to editing at the individual level. The macro editing procedure looks for suspicious data according to its potential impact on survey estimates or census totals.

Macro editing procedures often compare tabulations in the current period (time =  $t$ ) to tabulations obtained in the prior period from the same cells. These comparison are called “historic cell comparisons” and are generally of two forms: an absolute difference between the two corresponding estimates in cell  $i$  ( $\hat{D}_i = |\hat{Y}_{t,i} - \hat{Y}_{t-1,i}|$ ) or; a ratio comparison within the same cell ( $\hat{R}_i = \hat{Y}_{t,i} / \hat{Y}_{t-1,i}$ ). For either comparison, the  $\hat{D}_i$  or the  $\hat{R}_i$  are evaluated against specified limits, known as tolerances. As mentioned in Section 1, the simple parameter approach develops one set of tolerances per test: for example, the analysts would be asked to further examine the data comprising any tabulation whose absolute difference was greater than \$50 million or whose percentage change (increase or decrease) was greater than 20 percent in concurrent collection periods. In each of these identified tabulations, the individual units that impact the failed totals the most are manually reviewed. Reviewing from the top to bottom of the list of units ends when changing the micro data has no further effect on the survey estimates or census totals in the originally identified cell (Granquist, 1994). Note that this micro review procedure is performed in any cell that is identified as an outlier tabulation, regardless of macro editing procedure (simple parameters or H-B method).

### 3.2 The Hidioglou-Berthelot Method

Traditionally, ratio edits compare the ratio of two items to predetermined limits or tolerances (cf. the simple parameter procedure). This approach can be problematic when used on highly skewed economic data. First, since the large cell ratios will tend to dominate the distribution, smaller outliers on the left tail of the distribution can be undetectable (a “swamping” effect). Second, in many economic programs, the period-to-period variability of smaller cell estimates is often legitimately larger than the variability of larger ones. Hidioglou and Berthelot (1986) call this the “size masking effect.” Thus, using traditional ratio editing methods, can lead to too many small cell estimates being identified as outliers and not enough larger ones.

The Hidioglou-Berthelot edit method addresses these two problems for distributions of positive ratios. This method applies two transformations to the original set of tabulations to generate upper and lower tolerances. The first transformation uses the median to symmetrize the  $\hat{R}_i$  about 0, which ensures outliers are detected at both tails of the distribution. In the macro edit application described here, the  $i$  indexes the state and the ratios are calculated from the lines in the aggregated tables of financial data ( $Y_j$ ).

$$\begin{aligned} s_i &= (\hat{R}_i / R_m) - 1, \quad \hat{R}_i \geq R_m \\ &= 1 - (R_m / \hat{R}_i), \quad 0 < \hat{R}_i < R_m, \text{ where } R_m \text{ is the median of the } \hat{R}_i. \end{aligned}$$

The second transformation is performed to take into account the magnitude of the data.

$$E_i = \varepsilon_i \times \left( \text{MAX}(\hat{Y}_i^t, \hat{Y}_i^{t-1}) \right)^U, \text{ where } 0 \leq U \leq 1.$$

The exponent U in the transformation is used to “provide a control on the importance associated with the magnitude of the data”. The  $E_i$ ’s are also distributed around zero. If the parameter U equals zero, there is no emphasis placed on the size of the data and  $E_i = \varepsilon_i$ . As the parameter U gets closer to one, the larger cases affect  $E_i$  more, and  $\varepsilon_i$  becomes less significant. Often, with economic data, smaller units have the larger  $\varepsilon_i$  values. As U gets closer to one, the outliers will consist of large units only. Let  $E_{q1}$ ,  $E_m$ , and  $E_{q3}$  be the first quartile, median, and third quartile of the  $E_i$ . Define the equations  $d_{q1} = \text{MAX}(E_m - E_{q1}, |AE_m|)$ , and  $d_{q3} = \text{MAX}(E_{q3} - E_m, |AE_m|)$ . The  $AE_m$  term avoids detecting false outliers when the value of either  $(E_m - E_{q1})$  or  $(E_{q3} - E_m)$  is very small. The  $E_i$ ’s are clustered around a single value with a few true outliers. Hidioglou and Berthelot suggest setting A=0.5. Outliers are units outside the interval  $(E_m - Cd_{q1}, E_m + Cd_{q3})$ . The parameter C, set by trial and error, controls the width of the acceptance interval. Sigman (2005) set C to 4. Thompson and Ozcoskun (2007), Thompson (2007), and Banim (2000) set C to 10 and 20. Davila (2004) varied C between 15 and 47. Values of C >47 were also tried, without diverging results.

To summarize, the H-B method performs a series of transformations on historic cell ratios, then dynamically determines the tolerance limits from the data set at hand. This has the advantage of using the shape of the existing distribution to locate suspicious values, instead of using arbitrary limits or limits based on historic data. The disadvantage in the application of the H-B method is appropriate values need to be determined for three key input parameters (U, the influence parameter; C, the tolerance limit parameter; and A, the narrow range parameter) for the data sets and ratio tests under consideration.

Although the H-B method was designed to isolate outlier historical ratios in periodically collected micro data, it has also been used in macro editing: see Banim (2000) and Sigman (2005). Thompson (2007) and Thompson and Ozcoskun (2007) also report promising research results by applying the H-B method to data from a variety of ongoing economic programs, when input ratios are highly correlated and H-B method parameters are allowed to differ by ratio pair.

#### 4.0 Macro Editing Methodology for 2006 and 2007

##### 4.1 Application of Macro Editing to the 2006 ASGF

The ASGF summary tables contain 139 estimates, for each state and the District of Columbia, at the national, state, and total local levels of government. In 2006, we implemented the simple parameter approach described in Section 3.1 for 79 estimates (or table lines) at the total local government level for each state.

For the 2006 macro edits, the subject matter specialists designated any cell where the absolute difference was greater than a \$50 million difference **and** the cell ratio change was

greater than 20% as a macro edit failure targeted for review. We also failed estimates when either the current year tabulation value or the prior year tabulation value was zero but not both. For each flagged tabulation, analysts were provided with the five units' data for manual review that contributed the most to the weighted absolute difference of a failure as a drill down, in a top down manner.

We did four rounds of macro editing using these simple parameters. Our first round of macro editing isolated 1,183 failed units. Note that a unit could be listed more than once if it was contained in more than one simple parameter failing edit. After the analysts finished making corrections, our last round had 1,040 failed units. There were 312 units on the first listing only, and 169 on the last listing only. There were 871 units and items in common between the first and last round listings, 46 which had data changed. With each identified macro edit failing tabulation, we always provided five units from within the tabulation cell for analyst review, so different units could be on a later listing when units on prior listings were resolved. Thirty-seven states and lines failed during the first round of macro editing that did not fail on the last. Three states and lines failed the last round that did not fail the first. Some estimates that failed the first round of macro edits had large changes by the last round. However, many changes did not appear to change the level of the state estimate, and a lot of estimates were not changed. The analysts also changed data not isolated by macro edits and the counts of these data changes could not be separated out. The macro edit listing was only one of the tools that was used to verify the data after estimation. Subject matter specialists considered the macro edit workload reasonable, and agreed many of the units the simple parameters isolated should have been reviewed, and either changed or verified.

A required peer review was conducted after the last round of macro editing. This is not a statistical review and is conducted by analysts outside the survey. The purpose of the peer review is to help identify problems in data, documentation, table structure, and presentation before publication. Historical consistencies of the estimates were reviewed. The 2006 peer review questioned approximately 70 estimates for current to prior changes of which 49 had macro edits attached. Thirty of the 49 failed the final macro edit. We were interested in the estimates the peer review isolated that our macro edits did not fail. Many of these estimates had a large percent change but less than a \$50 million difference. The peer review did not question the estimates the analysts resolved as a result of the macro edit process. With hindsight, it is easy to set simple parameters to fail all the totals questioned by peer review, or to fail a certain number of states per aggregate.

#### **4.2 Development of the H-B Method Parameter Settings for the 2007 Census**

The 2006 AGSF macro edit application demonstrated the feasibility of using a macro editing approach on a large recurrent program. However, since the tolerance levels were subjectively determined, it was thought preferable to investigate statistical methods. We decided to research the H-B method using 2006 and 2005 data and compare it to the simple parameter method. We wanted a smaller workload for aggregates such as interest earnings, where more than 20 states failed with the simple parameters. The subject matter specialists decided to macro edit all 139 table lines beginning in 2007. Since the 2007 data were not available at this point, the initial evaluations were done using historical data from the 2006 and 2005 data estimates. The current year estimates were those at the start of 2006 macro edit production, and the prior year estimates were the latest 2005 estimates. With the expansion in 2007 to 139 estimates from 79, the 2006 simple parameters provided too much editing workload. To alleviate this problem, we relaxed the simple parameters for certain

volatile estimates such as Capital Outlay.

Initially for the H-B method, we set A to 0.5 as originally recommended by Hidioglou and Berthelot (1986). To assess the impact of the width of the tolerance region on outlier identification, we considered five different tolerance limit parameters, ranging from very narrow to very wide limits [C = 5, 10, 20, 30, and 40]. To assess the impact of large tabulations on the H-B method's effectiveness, we considered four different settings of the influence parameter [U = 0.0, 0.3, 0.5, and 1.0]. After looking at the states that failed in the first 32 table lines, it became apparent we should set U to 0.3 for all lines. We set C to 5 if the table line had been macro edited (with simple parameters) during 2006 production; for the new table lines C was set to 10. This appeared to balance all factors: failure of actual problems, and analysts' workload. After the H-B method was run, there were 80 state and line combinations that failed both the simple parameters and the H-B method, 250 states and line combinations failed the H-B method only, and 211 failed with the adjusted simple parameters only. We adjusted the H-B method parameters to minimize analyst workload and took into consideration those cases that the analysts indicated were real problems.

When reviewing the 250 states and lines that failed the only the H-B method, we noticed a large number were concentrated in 12 table lines. Each of these lines had 5 or more states that failed the H-B method that the subject matter specialists didn't want to review. They tended to be large percentage changes in typically small estimates, or more volatile estimates such as capital outlay. For these table lines, we raised C to widen the acceptance interval and raised U to place more emphasis on the size of the data. This reduced the number of state and line combination failures from 250 to 183. Interestingly for public utility, motor vehicle license, school lunch sales, and parking facilities, the H-B method application with adjusted parameters only failed the states isolated by simple parameters or questioned by the 2006 peer review.

**Table 1:**  
**Examples Where the H-B Method Originally Didn't Fail States the Analysts Wanted to Review**

Variable	C	U	States that Failed	Analyst's list	Parameter decision
Special Assessments	5	0.3	AZ, ME, SC	CA	x
	5	0.8	AR, CO, SC		
	2	0.8	AZ, AR, CA, CO, FL, KS, LA, MD, MO, NM, SC, UT, WV		
Intergovernmental Expenditures	10	0.3	MT, NV, OR, TN	CA, OH	
	20	0.8	CA, NV, NY, OH, OR, TN		x
	30	0.8	NY, OR, TN		

**Table 2:**  
**Comparison of H-B Method and Simple Parameters for Macro Editing the 2006 ASGF**

Description	Simple parameters to final H-B parameters
Number of edits (table lines) that did not fail any states for both methods.	23
Number of edits that failed the same states for both methods.	7
Number of edits where more states failed using the H-B method parameters, and the simple parameters were a subset.	13
Number of edits where more states failed using the simple parameters, and the H-B method parameters were a subset.	9
Number of edits that did not fail any states with the simple parameters, and failed some states with the H-B method.	20
Number of edits that did not fail any states with the H-B method, and failed some states with the simple parameters.	19
Number of edits where H-B method failed more states, and the states were different.	11
Number of estimates where simple parameters failed more states, and the states were different.	34
Number of edits that failed the same number of states for both methods but the states were different.	3
Total number of estimates.	139

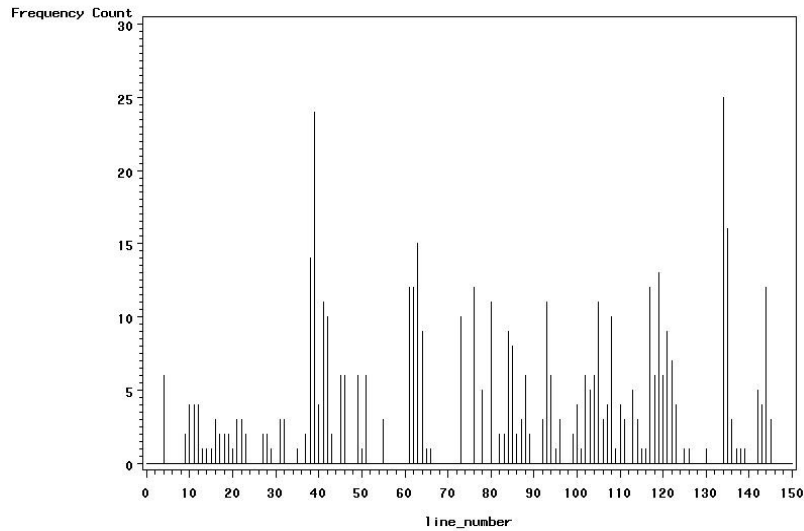
The subject matter specialists reviewed the 211 estimates (covering 53 table lines) isolated by simple parameters that did not fail with the H-B method. They determined that 68 failures were cases they wanted to see. For several of the involved edits, it was impossible to adjust the H-B method parameters to have a reasonable macro editing workload. Table 1 gives examples for the Special Assessments and Intergovernmental Expenditures edit failures. Column 1 indicates the edit variable; columns 2 and 3 give the attempted H-B method parameters. Column 4 indicates which states failed the H-B method with those parameters. Column 5 indicates which states the analysts wanted to see. Column 6 indicates the final decision.

Table 2 summarizes how the H-B method with final parameters performed against the 2006 simple parameters for all 139 table lines. The 2006 simple parameters failed 495 states and lines, yet 313 states and lines failed with H-B. The adjusted H-B parameters reduced the editing workload in comparison to the original simple parameters. Table 2 shows a total of 44 edits had more failures with H-B, and 62 had more failures with the simple parameters.

There was less variability of failure amongst table lines using the H-B method over the original simple parameters. Figures 1 and 2 present SAS needle scatter plots of table line number by the number of failed states per line. The H-B method did not fail more than 12 states per line, whereas the original simple parameters failed up to 24.

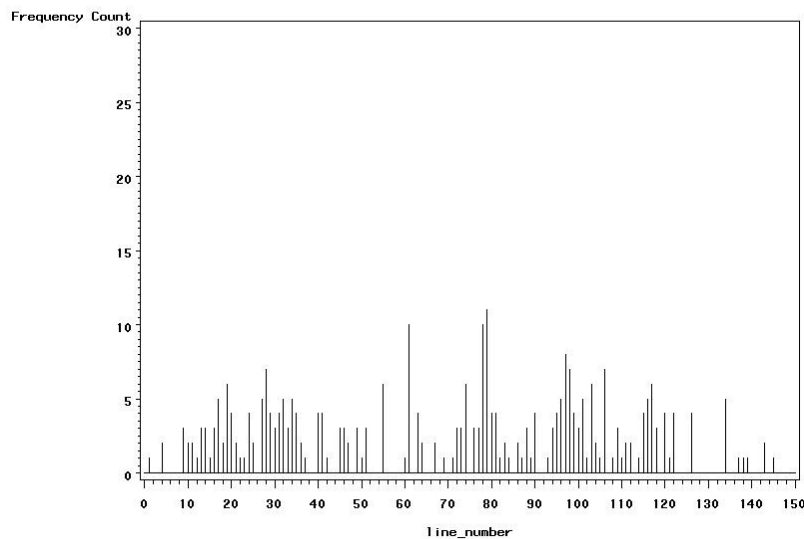
**Figure 1:**

Plot of Table line number by edit failures per line using original simple parameters



**Figure 2:**

Plot of Table line number by edit failures per line using H-B parameters





## 5. Discussion

### 5.1 Conclusions

We decided to research the H-B method as a more robust statistical method for developing macro edit parameters for the 2007 COGF and beyond. From the data we could construct the ratio edit bounds as opposed to the less reliable method of having analysts supply the bounds. During research, we found that the H-B method provided a smaller editing workload than simple parameters and did a better job of isolating outliers. Often, estimates that the 2006 peer review questioned and were not isolated by the 2006 simple parameters method were isolated by the H-B method. The peer review, although not statistical, was a barometer of how well the H-B method is working as well as the insight of the subject matter specialists. Both methods require subject matter expertise to develop the parameters.

### 5.2 Recommendations

Based on the 2006 research, we recommend using the H-B method to macro edit the ASGF and the COGF. The actual data are used to determine the outliers. We should continue to review estimates questioned by the peer review but not isolated by macro editing (and vice-versa) as one way to monitor the parameters used in the H-B method. For the future we will continue to look for other methods of analyzing the data at the estimate level to determine if there are any improvements needed.

### 5.3 Further Research

There will be upcoming research for the 2007/2002 COGF. We used the H-B method parameters to macro edit the 2007 Census.

All 50 state and D.C. current and prior estimates are used to develop the parameters for the H-B method for a given table line. We would like to research developing parameters for large and small states for each table line. This could produce more salient outliers, although failure would be based on fewer states. Macro edits of current year relationships should be researched. Also subject matter specialists want to develop different methods for isolating individual government units that cause the edit failures.

### Acknowledgments

The authors would sincerely like to thank Rita Petroni, Jenny Thompson, Carma Hogue, and Lisa Blumerman for their assistance in reviewing this abstract and paper. The authors would also like to thank Kenneth Lederman and Kheang Hang for their participation on this project.

### References

- Banim, J. (2000). "An Assessment of Macro Editing Methods," UN/ECE Work Session on Statistical Data Editing, Cardiff, United Kingdom, 18-20 October, 2000.  
[www.unece.org/stats/documents/2000/10/sde/7.e.pdf](http://www.unece.org/stats/documents/2000/10/sde/7.e.pdf).
- Cornett, E., McLaughlin, J.F., Hogue, C. R., and Owens, S.D. (2006). "A Comparison of Two Ratio Edit Methods for the Annual Survey of Government Finances,"

- Proceedings of the Section on Survey Research Methods, American Statistical Association.
- Hidioglou, M.A. and Bertheolot, J.M. (1986). “Statistical Editing and Imputation for Periodic Business Surveys,” *Survey Methodology*, Vol. 12. 73-83.
- Davila, E. H. (1994). “Macro-Editing - The Hidioglou-Bertheolot Method (Statistical Edits),” United Nations Statistical Commission and Economic Commission for Europe
- Granquist, L. (1994). “Macro-Editing - A Review of Some Methods for Rationalizing the Editing of Survey Data,” United Nations Statistical Commission and Economic Commission for Europe.
- McKenzie, L., Craig, T., and Hogue, C .R. (2007). “Investigation of Selective Editing Procedures for the Annual Survey of Government Finances,” Federal Committee on Statistical Methodology.
- Sigman, R.S. (2005). “Statistical Methods Used to Detect Cell-Level and Respondent-Level Outliers in the 2002 Economic Census of the Services Sector,” Proceedings of the Section on Survey Research Methods, American Statistical Association.
- Thompson, K.J. (2007). (2007). “Investigation of Macro Editing Techniques for Outlier Detection in Survey Data,” Proceedings of the Third International Conference on Establishment Surveys, American Statistical Association.
- Thompson, K.J., and Ozcoskun L. (2007). “An Empirical Investigation into Macro Editing,” Federal Committee on Statistical Methodology.