STATISTICAL METHODS FOR DEVELOPING RATIO EDIT TOLERANCES FOR ECONOMIC CENSUSES

Katherine Jenny Thompson and Richard S. Sigman, U.S. Census Bureau
Katherine Jenny Thompson, Bureau of the Census, ESMPD, Room 3108, Washington, DC 20233

Key Words: EDA, Nonparametric, Resistant, Robust

1. Introduction

Data collected by the Economic Census Programs are subjected to ratio edits as a part of the overall data-review process. In a ratio edit, the ratio of two highly correlated items is compared to upper and lower bounds, known as tolerances. Ratio edits are useful because it is difficult to evaluate the "reasonableness" of a data item's value by itself. By comparing an item to other related values, the analyst can determine if the response appears valid (e.g., total annual hours to total employees should be approximately 2000). Ratios outside the tolerances are considered edit failures, and one or both of the items in an edit-failing ratio are either imputed or flagged for analyst review. The efficiency of the ratio edit is therefore dependent on the selected tolerances.

When historical data are available, the development of ratio-edit tolerances usually begins with data analysis. Ratio-edit tolerances separate a distribution of ratios into two regions: an acceptance region and an outlier region. Determining ratio-edit tolerances via data analysis thus falls into the category of univariate outlier detection applied to ratios. Because the distributions of ratios are generally unknown, nonparametric techniques are often preferred.

This report compares three different tolerance development methodologies. In Section 2, we describe our considered outlier detection methods. Most of the techniques presented assume the ratios are symmetrically distributed. Distributions of ratios created from skewed data may or may not be symmetric. In Section 3, we present our approach for symmetrizing skewed distributions of ratios. Section 4 describes our evaluation of the outlier detection methods. Section 5 contains final remarks and recommendations.

2. Outlier Detection Methods

The historic data we used to develop tolerances contained outliers. We can minimize the effects of deviate observations by using methods that are said to be robust or resistant. A robust method is insensitive to departures from assumptions surrounding an underlying probabilistic model. Resistant methods are insensitive to localized misbehavior in data and produce results that change only slightly when a part of the data is replaced by new (entirely different) numbers. (Hoaglin et al., 1983).

We examined three approaches to setting tolerance limits: a robust approach (fifteen-percent trimmed mean and standard deviation); an outlier-resistant approach (resistant fences); and a gap analysis approach.

2.1 Robust Mean and Standard Deviation

In an earlier economic census editing cycle, the first step of tolerance development was to eliminate all of the ratios within a class that were more than two standard deviations from the mean and to base the tolerances on the resultant data set. We attempted an analogous technique using robust estimates of the population mean and standard deviation. We used the fifteen-percent trimmed mean as our robust estimate of the population mean and a robust estimate of population standard deviation based on the Winsorized sum of squared deviations (a consistent estimator of the variance of the trimmed mean. See Gross (1976)).

If the distribution of ratios is approximately symmetric, the robust estimates of the population mean and standard deviation define a robust confidence interval for the mean. See Gross (1976). Though we are not interested in the mean, the underlying principle is that the distribution is "normal in the middle," so that ($\bar{x}_k \pm 3\sigma_k$) is a plausible tolerance interval for the data. A liberal rule for setting tolerances defines an outlier as any ratio greater than ($\bar{x}_k + 2\sigma_k$) or less than ($\bar{x}_k - 2\sigma_k$). A more conservative rule uses $\pm 3\sigma_k$.

2.2 Resistant Fences

West (1995) recommended using an exploratory data analysis (EDA) outlier detection method called resistant fences to develop tolerances. Resistant-fences rules are based on sample quartiles. Given an ordered distribution of ratios, let

q_{25} = the first quartile (the lower fourth)
q_{75} = the third quartile (the upper fourth)
$H = q_{75} - q_{25}$, the interquartile range

The resistant-fences rules define outliers as ratios less than $q_{25} - k*H$ or greater than $q_{75} + k*H$, where k is a constant. The inner fences rule sets k equal to 1.5. The Tukey boxplot uses the inner fences rule. The outer fences rule sets k equal to 3. A compromise rule -- the middle fences rule -- sets k equal to 2.

We calculated the quartiles by setting the cumulative probability level for the i<u>th</u> order statistic, $x_{(i)}$, equal to $i/(n+1)$ as recommended by Hoaglin and Iglewicz (1987).

The resistant-fences rules implicitly assume that the ratios are symmetrically distributed. Symmetry is not required, however. We applied the resistant-fences rules to both transformed and untransformed data.

2.3 Distance Measurement Algorithm for Selection of Outliers (D_MASO)

The D_MASO gap analysis approach was developed at the Census Bureau to develop ratio edit parameters for the 1992 Enterprise report (Oh et al. (1994)). D_MASO examines successive ratios of ordered explicit ratios, considering distances (actually proportions) between adjacent ordered observations to find potential edit bounds. An "unusually" large gap between adjacent observations at either end of the distribution may indicate that the observations between the gap and the end of the distribution are outliers. The observation on the "center side" of the gap is a potential tolerance.

The user specifies the maximum percentage of observations that can be labeled as outliers and specifies a cut-off value for distance comparisons.

3. Symmetrizing Skewed Distributions

3.1 Background

In general, economic data are highly skewed. However, most of the statistical methods considered for tolerance-development assume that the ratios are symmetrically distributed. We use transformations called power transformations to symmetrize skewed distributions of ratios. Once transformed, many ratios that originally appeared to be outlying become consistent with the rest of the distribution. A power transformation with parameter equal to p is a function of the following form:

$$T_p(x) = \begin{cases} x^p & (p>0) \\ \log(x) & (p=0) \\ -x^p & (p<0) \end{cases} \qquad (1)$$

Our approach to determining the appropriate p parameter for the power transformation of a skewed distribution of ratios consisted of the following steps:

- Employ a modification of an EDA method for determining p, described by Hoaglin et al. (1983). See section 3.2;
- Apply the natural logarithm transformation to the same distribution of ratios;
- Select the transformation that results in the smallest absolute value of the skewness coefficient.

We include the comparison to the natural logarithm transformation because the EDA method's resistance breaks down when the data set contains more than the expected number of outliers. We further discuss this issue in Section 3.2.

We did not symmetrize every distribution of ratios. Several of the distributions contain legitimate outliers and remain skewed even after being transformed. Moreover, a symmetrizing power transform may not exist for some distributions.

3.2 EDA Method For Symmetrizing Skewed Distributions

Most of the skewed distributions we examined could be symmetrized by applying the natural logarithm transformation to the ratios. This is equivalent to applying a power transformation with p=0. When the log transformation was not appropriate, we obtained p from the following modification of Hoaglin et al.'s (1983) transformation plot for symmetry.

Given an ordered sample of size n', let M be the median of the sample, and x_L and x_U represent the lower and upper values of a set i of k-percent approximate quantiles. The EDA transformation plot for symmetry places

$$x_{vi} = [(x_U + x_L)/2] - M \qquad (2)$$

on the vertical axis (v), and

$$x_{hi} = [(x_U - M)^2 + (M - x_L)^2]/[4M] \qquad (3)$$

on the horizontal axis (h), so that

$$\text{slope}_i = x_{vi}/x_{hi} \qquad (4)$$

If the resultant graph is **nearly linear** -- that is slope $\approx c$ -- then p = 1-c. We determined a value for p without graphing the points, by setting p = 1-median(x_{vi}/x_{hi}). Hoaglin recommends rounding p to the nearest power of ½. We did not employ this rounding. The procedure's resistance breaks down when the actual number of outliers exceeds the number of observations in one (or both) of the quantiles with the smallest tail probability. This is known as exceeding the breakdown point. The breakdown point can be raised by decreasing the number of quantiles employed. However, the lower the breakdown point, the more data points available to calculate the median slope.

To minimize the effect of multiple outliers, we deleted extreme observations from the untransformed sample using the outlier resistant fences rule described in section 2.2. We refer to the reduced sample size as n'. To maximize the resistance, we linked sample size with expected number of outliers before breakdown by using the sets of quantiles specified in Table 1.

Table 1: Quantiles Used for EDA Transformation Plot for Symmetry

Range of n',	Breakdown Point	Quantiles Used
33-64	1/8	1/4, 1/8
65-128	1/16	1/4, 1/8, 1/16
129-256	1/32	1/4, 1/8, 1/16, 1/32
257-512	1/64	1/4, 1/8, 1/16, 1/32, 1/64
513-∞	1/128	1/4, 1/8, 1/16, 1/32, 1/64, 1/128

Our variation of the EDA method allows for four expected outliers in the subsetted dataset before breakdown.

4. Determining Ratio Edit Tolerances

4.1 The Ratio Edit As A Hypothesis Test

A ratio edit is a hypothesis test, in which the null hypothesis is that both data fields in the ratio are correct (Pierce and Gillis, 1995). One rejects the null hypothesis when the ratio falls outside of the tolerances.

Given this definition, we can define Type I and Type II errors for each ratio edit. A Type I error flags a ratio value as an error when it is in fact correct: these are **good** ratios that fall **outside** of the tolerances. A Type II error flags a ratio as correct when it is in fact an error: these are **bad** ratios that lie **inside** of the tolerances. Type I error increases unnecessary analyst work and is generally controlled by widening the tolerances. However, the wider the tolerances, the greater the probability of Type II error. Some caution must be used in defining Type II error. Only a portion of the Type II error for an individual ratio test is controlled by the tolerance limits. With ratio edits, there is usually an "inlier" set of bad ratios, where an inlier is defined as a **bad** ratio whose value is consistent with the rest of the distribution. For example, ratio edits rarely identify rounding errors: if **both** items are reported in the wrong units (e.g. thousands instead of units), the ratio value will be acceptable even though both data items are scaled incorrectly.

Because item values are often tested in more than one ratio edit, the individual ratio-edit Type II error is a poor measure of the **overall** proportion of uncorrected (unidentified) **bad** items left remaining in the edited data. Consequently, we define the Type II error of an **entire set** of ratio edits as bad data that passes all containing ratio edits. At the ratio-edit level, the **all-ratio-test** Type II error (for the **complete** set of ratio edits) is the number of bad ratios that are not outside of **any** tolerances. At the individual data-item level, the **all-item** Type II error (for the **complete** set of ratio edits) is the bad data items that are not contained in any ratios outside the tolerances, i.e., the bad items that are **not** identified as outliers by any of the ratio edit tests.

We can also examine the power of a set of ratio edits. For outlier detection, the power is the probability of correctly concluding that a bad ratio is an outlier. The power of a set of ratio edits is the proportion of bad ratios that fall outside of one or more tolerances.

An alternative measure for evaluating the tolerances for a given ratio test is the **hit rate** -- the ratio of the number of bad ratios outside of the tolerances to the total number of ratios outside of the tolerances (Granquist, 1995).

To examine the different methods, we used historic data (see section 4.2) and classified each ratio based on its historic edit outcomes. First, we considered each data item separately using the reported data item, the edited data item, and the data item's edit action flag. We classified each nonblank/nonzero item as:

Good- The data item should be located within the edit bounds;

Bad- The data item should be located outside of the edit bounds;

Questionable- The data item cannot be classified as "good" or "bad." The edit action flags for these items may be difficult to interpret. On the average, approximately five percent of the non-zero data items were flagged as questionable.

A **ratio** is **good** if both the numerator and the denominator are flagged as **good.** A ratio is **bad** if either the numerator or the denominator is flagged as bad. Ratios that contain blank or questionable values are excluded from all evaluations. The flowchart in Figure 1 shows how we classified and used the historical data.

4.2 Historic Data Sources

We had two sources of historic data: the 1994 Annual Survey of Manufactures (ASM) and the 1992 Business Census. Our data included only full-year reporter establishments and excluded all full-impute cases.

The ASM is a mail-out/mail-back survey representing all establishments that received a form in the previous census of manufactures. The survey provides detailed annual statistics on the location, activities, and products of approximately 58,000 U.S. manufactures.

Prior to ratio-editing, the ASM reported data undergoes some clerical edits. We used this partially edited data to develop the tolerances. Twelve of the ASM ratio edits use statistically developed tolerances. The ASM uses the same set of ratio edits in each standard industrial classification (SIC).

The Business Census comprises five trade areas: Retail Trade; Wholesale Trade; Service Industries; Transportation, Communication, and Utility Industries (Utilities); and Finance, Insurance, and Real Estate Industries (FIRE). It is a mail-out/mail-back census and is conducted once every five years.

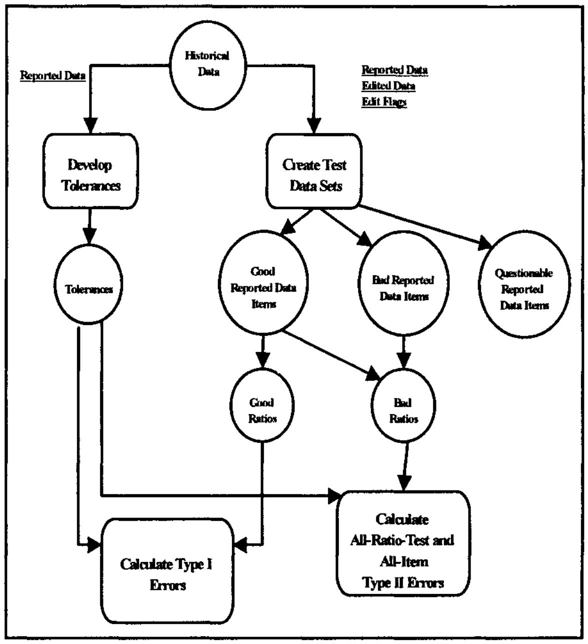

Figure 1

The Business Census reported data are directly input into ratio editing without a prior clerical edit. Administrative data is substituted for blank data whenever possible to develop tolerances, so we used reported and administrative data to develop the tolerances. Some trade areas classify the establishments within SIC by legal form of organization, type of operation, and tax status. We used the trade area classifications for our evaluation, but refer to each classification as an SIC. Each trade area in the Business Census employs a common set of core ratio edits. Four of these ratio edits require statistically developed tolerances. We performed our evaluation by trade area within census for the four statistically determined ratio edits.

4.3 Evaluation Methodology

We generated nine sets of edit tolerances per ratio test in each SIC: two sets of robust edit limits for symmetrized distributions; three sets of resistant edit limits for symmetrized and unsymmetrized distributions (six sets total); and one set of gap analysis (D_MASO) edit limits.

We used sixteen SICs for the ASM evaluation. Because some industries contained less than sixteen establishments with non-zero ratios, we could not produce tolerances for all twelve of the ratio tests in each SIC. We used thirty SICs for the Business Census evaluation. Our objective was to find a technique that balanced the goals of maximizing the number of rejected bad items and minimizing the number of rejected good items.

Thompson and Sigman (1996) presents the average Type I error rates for each ratio test, the all-ratio-test Type II error rate, and the all-item Type II error rate obtained using each tolerance development method on 1994 ASM data and 1992 Business Census data by trade area.

4.3.1 Comparison of Robust and Resistant Methods: Symmetrized Distributions

We examined two outlier-detection methods for symmetrized distributions: robust methods and resistant methods. Our goal was to eliminate the method that performed the worst overall in terms of Type I error (false reject rate). A high proportion of ratio tests with a Type I error greater than 0.10 was considered unacceptable.

Across the board, the tolerances generated with two robust standard deviations were too narrow. The robust methods perform poorly because the tails of the symmetrized distributions are heavier than those of a normal distribution with σ^2 equal to that estimated by the Winsorized variance estimator. Thus, $\pm 2\sigma$ did not cover the expected 95 percent. We concluded that there was no apparent advantage in further pursuing the robust estimation techniques. The remainder of our evaluation concentrated on the resistant methods.

4.3.2 Comparison of Resistant Fences Methods: Symmetrized and Unsymmetrized Data

We first compared the resistant methods separately on symmetrized and unsymmetrized data. For each SIC/ratio, we selected a "best" resistant fence rule for the unsymmetrized data and for the symmetrized data. We then examined whether the symmetrizing was necessary for the historic data used. For more details, see Thompson and Sigman (1996).

For most of the ASM ratios, the same resistant fence rule worked best on both the symmetrized and unsymmetrized data. In fact, the tolerances generated on both data sets are similar. Although the ASM distributions of ratios are generally positively skewed, the degree of skewness is not often severe. The symmetrizing compressed the ASM distributions of ratios but did not dramatically change their shapes. In many cases, the transformed distributions remained skewed because the longer tail consisted entirely of outliers. Consequently, the tolerances developed from the symmetrized data were slightly narrower than those from the unsymmetrized data.

In general, the tolerances calculated from the unsymmetrized ASM data yielded tests with slightly higher hit rates (proportion of rejected ratios that were bad) than those calculated from symmetrized data. Consequently, the Type I error rate (proportion of rejected good ratios) is also lower for the unsymmetrized distributions. Moreover, the proportion of ratios inside of the tolerances that was bad was essentially the same for both the symmetrized and unsymmetrized distributions.

In contrast with the ASM results, the resistant methods performed quite differently on the Business Census symmetrized and unsymmetrized distributions of ratios. The Business Census data is highly positively skewed, more so than the ASM data. Because of the degree of skewness, the interquartile range (*H*) is generally larger than (q_{25} -$x_{(1)}$) for the unsymmetrized data: the lower bound is almost always negative, and the upper bound is near the center of the distribution. For most of our data sets, applying the natural log transformation to the distributions of ratios corrected most of the skewness (occasionally another power transformation was required). When the resistant fences rules were applied to the symmetrized data, the tolerances were generally near the ends of the distributions. The Type I error rate decreased when we applied the resistant fences rules to the symmetrized data, while the proportion of bad ratios inside of the tolerances remained nearly constant.

Applying the resistant outer fences (k=3) to the orginal **unsymmetrized** distribution of ratios usually worked well for the ASM data. This was not always the case with the Business Census data. Although **symmetrizing** the distributions improved the hit rate of the tests, a distance of three interquartile ranges from the upper and lower quartiles was not often sufficient. In three of the five trade areas, the Type I error rate was too high for k=3 (on the average, greater than 0.05). We found that specifying four interquartile ranges (k=4) improved the Type I error rate with very little loss in individual hit rates or total power for the census of Retail Trade, the census of Service Industries, and the census of Transportation, Communication, and Utitilities Industries.

4.3.3 Comparison of Resistant Fences and the D_MASO Algorithm

The D_MASO algorithm was developed at the Census Bureau to generate tolerances for the 1992 Enterprise Report (Oh et al., 1994). There are some key differences between the D_MASO approach and the resistant approaches:

- the D_MASO algorithm does not use a probability model;
- the D_MASO algorithm looks for separate groups of observations to determine outlier zones, rather than looking for extreme observations;
- with D_MASO, the user specifies *a priori* the maximum proportion of the data that can be labeled as outliers.

For our application, we specified a maximum outlier proportion of five percent per tail and used the default cut-off factors of 1.2 for the lower and upper tail as specified in Oh et al. (1994). We found that the algorithm was fairly insensitive to the cut-off percent when the default cut-off factors were between 1.2 and 3. The selection of the cut-off factor had much more effect on the bounds than the starting location of the algorithm.

The resistant methods do not generate tolerances for distributions containing less than sixteen observations. The D_MASO procedure does. If the SIC did not have the two estimates of bounds, we excluded it from our comparison. We compared the D_MASO procedure to the most successful resistant procedure for each historic data set: outer fences with **unsymmetrized** ASM data; and outer fences (k=3) or "big" fences (k=4) with **symmmetrized** Business Census data, depending on trade area. See Thompson and Sigman (1996) for more details.

For most of the ASM ratios, the resistant fences methods usually performed better than D_MASO. In the few cases where the D_MASO bounds were clearly superior, the original distributions were very positively skewed; the resistant fences bounds were too narrow and were negative at the lower end. In general, however, the ASM resistant fences tolerances were usually slightly wider than the D_MASO tolerances and identified the same bad ratios. Consequently, the Type I error rate is usually higher for D_MASO. The power is about the same for the two methods, although the hit rate is generally higher for the resistant fences tolerances. However, the difference in error rates and hit rates between the two methods is usually caused by a small number (two or three) of rejected good ratios.

For the Business Census data, the resistant methods outperformed D_MASO in three trade areas: Wholesale, Utilities, and FIRE. In the other two trade areas, the two methods tied in terms of overall performance. The D_MASO procedure limits the number of observations that can be flagged as outliers, so the procedure begins at the tail ends of the distribution. In cases where the interquartile range was small and the range of the distribution of ratios was large, the resistant bounds were much narrower than the D_MASO bounds. In these cases, the D_MASO bounds outperformed the resistant fences bounds by a large margin in terms of rejected good ratios (Type I error).

4.4 Discussion

We examined three different approaches to setting tolerance limits. None of these approaches incorporated specialized subject-matter knowledge of the distribution of ratios. Analysts who work with economic data develop an understanding of the distributions of ratios in a given industry. A statistical methodology cannot replace this knowledge. However, it can serve as a good starting point, especially when there is no known mathematical relationship to rely upon.

Outlier detection methods can fail to work properly when more than one outlier is present. Problems that arise in the presence of multiple outliers are of two types: masking and swamping. Masking occurs when the presence of several outliers makes each individual outlier difficult to detect. Swamping occurs when multiple outliers cause the procedure to erroneously flag too many observations as outliers. These two problems can adversely impact tolerance development.

The resistant fences rules were designed to reduce masking. Because they are based on quartiles, they have a breakdown point of approximately 25%, i.e., "up to n/4 observations can be replaced by arbitrary values without causing the lower or upper cutoff value to become unbounded" (Hoaglin et al., 1986). Swamping must be controlled by the choice of k, the number of interquartile ranges between the quartiles and the fences.

Hidiroglou and Berthelot (1986) note that resistant fences methods are not free from masking. They cite two specific masking effects, both of which were present in our analysis. First, if the distribution is very positively skewed, then outliers on the left tail of the distribution are undetectable (as they are with generated negative lower tolerances for data that is always non-negative). Second, the resistant fences method does not make a specific provision for the size of the establishment, and the variability of ratios for small establishments is larger than the variability of ratios for large establishments. If the establishment size varies widely within a SIC, then too many small units will be flagged as outliers, and not enough large units will be considered. Hidiroglou and Berthelot refer to this as the "size masking effect."

D_MASO was designed to reduce swamping. The user specifies the maximum percentage of the data set that can be identified as outliers. In this case, the masking is controlled by the choice of lower and upper cut-off factors. The larger gaps in proportional distances are usually due to the smaller establishments, so the D_MASO algorithm is also prone to the size masking effect.

In terms of outlier detection, the resistant methods were the most consistently successful, in terms of balancing minimum Type I error and maximum power. They worked best when the distribution was approximately symmetric. After fine-tuning some of the values of k, we were able to develop tolerances with low Type I error rates (proportion of rejected good ratios) and reasonable power (proportion of rejected bad ratios) for most of the ASM and the Business Census ratio edits. As always, there is a trade-off between Type I error and Type II error: by minimizing the Type I error rate, we increase the Type II error rate for the set of ratio edits and correspondingly reduce the power of the set of ratio tests.

D_MASO worked quite differently for the two sets of historic data. Usually, the D_MASO bounds were too tight with the ASM data: the algorithm appeared to be quite prone to swamping. This result surprised us because it was counter to the design of the algorithm. We expected the D_MASO bounds to be wider than the resistant bounds in most cases. There was no clear pattern for the Business Census data.

The appeal of the D_MASO approach is the user's control over the maximum number of outliers From a statistical perspective, this is not necessarily a strength. Deciding *a priori* on the number of outliers that can be detected has quality implications for the final edited data; tabulations may use data that contains several unexamined erroneous observations. If the number of establishments is large, then the Type II error rate (proportion of accepted bad ratios) can have a significant effect on the final tabulations.

For us, D_MASO was a "black box." We did not have an intuitive understanding of the ordered distribution of gaps for a ratio and had a difficult time relating the D_MASO breaks to histograms of ratios in a SIC.

In contrast, the resistant fences rules were fairly intuitive. This approach takes the skewness of the distribution into consideration, without any parametric assumptions. The resistant fences methods do not allow explicit control over the number of flagged outliers per ratio test. In practice, however, we found that the percent of observations in the rejection region could be controlled through the choice of k.

Based on our evaluation results, we recommend the following steps for developing tolerances:

Step One- Use a power-transformation to symmetrize skewed distributions of ratios, implementing the procedure described in Section 3;

Step Two - Apply resistant outer fences rule (k=3) to the (symmetrized) data to obtain outlier bounds;

Step Three - Use the inverse power transformation on the initial bounds to obtain final bounds (if necessary);

Step Four - Examine the bounds. If they are too wide, decrease the value of k. If they are too narrow, increase the value of k. Repeat steps two through four as necessary.

For a new survey or census, the user might prefer starting with a less conservative rule, such as the middle fences rule (k=2).

5. Conclusions

In this paper, we examined a variety of methods for developing ratio edit tolerances. Our proposed approach for tolerance development uses power transformations and EDA resistant fences rules. As we have shown, this approach must be modified for different data sets. Each economic census and survey collects unique data. The distributions of ratios will correspondingly be very different.

No matter how much we refine a statistical methodology, however, statistical methods cannot always provide the "best" tolerances. Statistical methods do not always replace subject matter expertise and common sense. As a general rule, a mathematical relationship that governs the upper and lower bounds of a ratio edit should preempt any statistical techniques. For example, the Business Census tests the ratio of Annual Payroll to First Quarter Payroll. Logically, the lower bound of this ratio is one. When we used the resistant fences methods to generate tolerances, our lower tolerances were as low as 0.85. For this distribution, a ratio value of 0.86 would not be flagged as an outlier. However, one of the two items being edited is obviously wrong.

At the Census Bureau, the traditional approach to tolerance evaluation examines summary statistics. Analysts review the tolerances, the industry averages, and the percentage of items that fall in the rejection region. This approach is limited. A better approach combines this review with some form of graphical analysis.

Developing good ratio edit tolerances is an iterative process. Our proposed approach provides an initial set of parameters. Data users should examine these parameters using a combination of graphical techniques and subject-matter expertise and modify them accordingly.

6. Acknowledgements

We thank Jock Black, William C. Davie, Jr., Lisa Draper, Vicki Farri, and Nash Monsour for their useful comments.

7. References

Barnett, V. and Lewis, T. (1978). *Outliers in Statistical Data.* John Wiley and Sons, New York.

Granquist, Leopold (1995). Improving the Traditional Editing Process. *Business Survey Methods.* New York: John Wiley and Sons, 385-481.

Gross, Alan M (1976). Confidence Interval Robustness with Long-Tailed Symmetric Distributions. *Journal of the American Statistical Association,* **71**, 409-419.

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

Hoaglin, D.C; Mosteller, F; and Tukey J.W. (Eds.) (1983). *Understanding Robust and Exploratory Data Analysis.* NY: Wiley.

Hoaglin, D.C, and Iglewicz, B.(1987). Fine-Tuning Some Resistant Rules for Outlier Labeling. *Journal of the Americal Statistical Organization,* **83**, 1147-1149.

Oh, Sungsoo; Paletz, David; Kim, Jay Jong-Ik; Salyers, Eddie (1994). Development of Edit Parameters for 1992 Economic Census Enterprise Reports. *Proceedings of the Section on Survey Research Methods,* American Statistical Association, 1144-1149.

Thompson, Katherine J. and Sigman, Richard (1996). Evaluation of Statistical Methods for Developing Ratio Edit Module Parameters. Bureau of the Census: Economic Statistical Methods and Programming Division. Report ESM-9601.

West, Sandra A. (1995). Discussion of paper presented at the Seminar on New Directions in Statistical Methodology. Washington, DC: Office of Management and Budget (Statistical Policy Working Paper 23, Part 1 of 3).

This paper reports the general results of research undertaken by the Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.