Investigation of Treatment of Influential Values

Mary H. Mulry and Roxanne M. Feldpausch[1]
U.S. Census Bureau, Washington, DC 20233

Abstract

This paper investigates ways to treat an influential observation in the estimation of total sales from the Monthly Retail Trade Survey. An observation is considered influential if the estimate of total monthly sales is dominated by its weighted contribution. Influential observations occur infrequently but are problematic when they do appear. To be clear, the assumption is that the influential observation is true although unusual, and not the result of a reporting or recording error. The paper examines several methodologies for treating influential values with the goal of finding those that use the observation but in a manner that assures its contribution does not have an excessive effect on the total. Two of the methods appear effective with monthly data but will need further investigation.

Keywords: Winsorization, M-estimation, reverse calibration

1. Introduction

This paper investigates methods of identifying and treating influential observations in the estimation of total sales from the U.S. Monthly Retail Trade Survey (MRTS). An observation is considered influential if its weighted contribution has an excessive effect on the estimate of total monthly sales (Chambers et al. 2000). Influential observations occur infrequently but are problematic when they do appear. To be clear, this study assumes that the influential observation is correct although unusual, and not the result of a reporting or recording error. The goal is to find methodology that improves the estimate of total sales and uses the observation in a manner that assures its contribution does not have an excessive effect.

Black (2001) describes the current corrective procedures employed in the MRTS when sample units can have a large and possibly erroneous effect on estimates. The methods include weight adjustments and moving a unit to a different industry when the nature of the business changes. The MRTS processing already includes running the Hidiroglou-Berthelot algorithm (1986) each month to identify outliers and create the imputation base (Hunt, Johnson, and King 1999). The Hidiroglou-Berthelot algorithm designates observations that should be reviewed and sometimes suppressed from the imputation base. The intent is for the treatment of influential values that is developed to complement, not replace, the Hidiroglou-Berthelot algorithm. With the Hidiroglou-Berthelot algorithm detecting and compensating for reporting errors, the expectation is that the appearance of influential values will be fairly rare.

Basically two approaches are available for the treatment of influential observations in estimation: (1) trimming the weight, sometimes called constraining the weight, and (2) modifying the value of the influential observation so that it has less impact on the estimate of the total. If the business is expected to continue to report influential values, then possibly it could be considered as belonging to another sampling stratum and a change in the weight may be the better option. If the influential value appears to be a rare occurrence for the business, then adjusting the value is more desirable.

The basic strategy is to identify candidate methodologies and use actual data for a month that contained an influential value to identify the methodologies that demonstrate promise for subsequent research. The evaluation criteria include the number of influential observations that are detected, including the number of true and false detections made. In addition, the evaluation will include an estimate of bias and an assessment of the impact on measures of change, in particular the month-to-month ratio of sales.

The study examined weight trimming and methods that modified the influential observation. The weight trimming approach made the arbitrary choice of cutting the weight to one-third of what it was originally. The methods examined for modifying the outlying observation were as follows:

(1) Winsorization (Chambers et al 2000)

   (1a) specifying a cut-off value by stratum (Kokic and Bell 1994)

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

(1b) specifying an individual cut-off value for each observation (Clarke 1995)
(2) Reverse calibration (Chambers and Ren 2004)
(3) Generalized M-estimation (Beaumont and Alavi 2004, Beaumont 2004).

This paper describes the candidate methods and the results of a study of one month's observations for a particular industry where an influential value is present.

2. Weight trimming

For the i-th business in a survey sample of size n, Y_i is its sales for the month of observation; w_i is its sample weight, and X_i is a variable highly correlated with Y_i such as previous month's sales or its monthly sales from a pre-entry questionnaire. Total monthly sales is estimated by

$$\hat{Y} = \sum_{i=1}^{n} w_i Y_i$$

An observation Y_i is considered influential if the contribution of $w_i Y_i$ dominates the estimated total $\hat{Y}$ (Chambers et al. 2000).

A method of weight trimming that seems appropriate for influential observations is truncating the weights of the influential observations and adjusting the weights of the other observations, the inliers, to account for the remainder of the truncated weights so that the new weights have the same sum as the original set of weights (Potter 1988). Other methods of trimming weights such as those found in Stokes (1990) and Cohen and Spencer (1991) appear more suited to addressing the problems of extreme weights than of influential observations.

The procedure we describe below assumes that the sum of the weights needs to remain the same. If the sum of the weights does not have to be the same every month, as would be the case if only rates were estimated, then the weights of the inliers would not have to be adjusted.

Assume that s is the sample of size n with weights w_i, and s_1 is the subset of inliers and s_2 is the subset of influential observations. For each influential observation i in s_2, define a truncated weight c_i. Then define a trimmed set of weights as follows:

$$w_i^T = \begin{Bmatrix} c_i, i \in s_2 \\ w_i A, i \in s_1 \end{Bmatrix} \text{ where}$$

$$A = 1 + \frac{\sum_{j \in s_2} (w_j - c_j)}{\sum_{j \in s_1} w_j}$$

The estimator of the total is then

$$\hat{Y}^T = \sum_{i=1}^{n} w_i^T Y_i$$

An equivalent approach only changes the weight for the influential observations but multiplies the final total by A as follows:

$$w_i^{T_2} = \begin{Bmatrix} c_i / A, i \in s_2 \\ w_i, i \in s_1 \end{Bmatrix}$$

The estimator of the total is then calculated as follows:

$$\hat{Y}^{T_2} = A \sum_{i=1}^{n} w_i^{T_2} Y_i$$

The calculation of A probably should be done within sampling stratum.

The choice of c_i also could be investigated. We considered only the option setting $c_i = w_i/3$, which was arbitrary.

3. Modifying the observation

We consider two methods that modify the influential observations, winsorization and reverse calibration. We also consider M-estimation, which can modify either the weight or the value of the influential observation.

3.1 Winsorization

Winsorization may be one-sided or two-sided. The initial focus is on one-sided winsorization which adjusts influential values deemed to be too large. One-sided winsorization methodology sets a pre-defined rule for adjusting an outlying (positive) value Y_i downwards, leaving remaining values unchanged (Searls 1966). There are two types of winsorization to consider (Chambers, Kokic, Smith, and Cruddas 2000). Both require setting a cut-off value K and defining an alternative value of Y_i. The basic Type 1 and Type 2 estimators are

Type 1

$$Y_i^* = \begin{Bmatrix} K, Y_i > K \\ Y_i, otherwise \end{Bmatrix}$$

Type 2

$$Y_i^* = \begin{Bmatrix} K + \frac{1}{w_i}(Y_i - K), Y_i > K \\ Y_i, otherwise \end{Bmatrix}$$

For both Type 1 and Type 2, the estimator of the total is

$$\hat{Y}_W = \sum_{i=1}^{n} w_i Y_i^*$$

Type 2 has the appealing feature of using all the data, but not weighting Y_i - K, the portion deemed extreme.

The investigation considered two methods for defining K. One method defines one value of K across all the strata. Other methods define a separate K_h by stratum or by observation. The investigation includes two ways of defining K described by Chambers et al. (2000) in a discussion of the Type 2 estimator from the prediction theory perspective:
(1) defining a separate K_h for each stratum (Kokic and Bell 1994) that minimizes mean squared error
(2) defining a separate K_i for each observation (Clarke 1995) that minimizes mean squared error.

For a stratified estimator, Kokic and Bell (1994) developed a separate K_h for each stratum h under a model assuming that for each stratum h, the Y_{hi} within the stratum h were independently and identically distributed such that $E(Y_{hi}) = \mu_h$ and $var(Y_{hi}) = \sigma_h^2$.

Kokic and Bell (1994) showed that the overall mean squared error was minimized under the model by choosing a separate K_h for each stratum defined by

$$K_h = (N_h / n_h - 1)^{-1} L + \bar{y}_h$$

where $L > 0$ is a constant chosen so that the bias of the stratified winsorized estimator is $-L$. An estimator of L is shown below.

An interesting generalization by Clarke (1995) defines a separate K_i for each observation and also uses L. The approach assumes a more general model where the Y_i can be characterized as independent realizations of random variables such that $E(Y_i) = \mu_i$ and $var(Y_i) = \sigma_i^2$.

The winsorized estimator of the total is written as

$$\hat{Y}^* = \sum_{i=1}^{n} w_i Z_i$$

where $Z_i = \min\{Y_i, K_i + (Y_i - K_i)/w_i\}$.

Clarke suggests approximating the K_i that minimizes the mean squared error under the more general model by $K_i = \mu_i + L(w_i - 1)^{-1}$, which requires estimating μ_i and L.

For an estimate of μ_i, Chambers et al (2000) suggest using the results of a robust regression. Then the estimate of μ_i is bX_i where b is the regression coefficient.

To estimate L, the Clarke method has several steps. First use the estimate of μ_i to estimate weighted residuals $D_i = (Y_i - \mu_i)(w_i - 1)$ by $\hat{D}_i = (Y_i - bX_i)(w_i - 1)$.

Next arrange the estimates of the residuals in decreasing order $\hat{D}_{(1)}, \hat{D}_{(2)}, \ldots.. \hat{D}_{(n)}$.

Then find the last value of k, called k^*, such that

$$(k+1)\hat{D}_{(k)} - \sum_{j=1}^{k} \hat{D}_{(j)}$$

is positive. Finally, estimate L by

$$\hat{L} = (k^* + 1)^{-1} \sum_{j=1}^{k^*} \hat{D}_{(j)}$$

For reference and context, the investigation also examines defining K as the sample mean plus two standard deviations.

3.2 Reverse Calibration

The reverse calibration approach (Ren and Chambers 2003, Chambers and Ren 2004) has two steps:
1) use a robust estimation method to estimate the total
2) then modify the influential observations to achieve that total.

Chambers and Ren consider the population U to be divided into two subsets U_1 containing inliers, the 'clean' part of the population that does not have any influential values, and U_2 which contains the influential observations. Likewise, sample s is divided into two subsets s_1 containing inliers, the 'clean' part of the data set that does not have any influential observations, and s_2 which contains the influential observations.

The outlier resistant estimation procedure rescales all the population in a manner that permits unbiased estimation and minimizes variance. The rescaled observations are defined by

$$Y_i^{**} = \begin{Bmatrix} f(\lambda)Y_i, Y_i \in U_1 \\ \lambda Y_i, Y_i \in U_2 \end{Bmatrix}$$

The method finds an optimal value of λ and $f(\lambda)$ so that the population total (not the sample total) remains unchanged:

$$\sum_{i \in U} Y_i = \sum_{i \in U} Y_i^{**} = f(\lambda) \sum_{i \in U_1} Y_i + \lambda \sum_{i \in U_2} Y_i$$

For the equality to hold, $f(\lambda) = 1 + \delta(1 - \lambda)$ where

$$\delta = \sum_{i \in U_2} Y_i \Big/ \sum_{i \in U_1} Y_i$$

Minimization of the variance subject to $f(\lambda) = 1+\delta(1-\lambda)$ leads to the optimum value λ_{opt} . See Chambers and Ren (2004) for details.

The outlier resistant estimator of the total is then

$$\hat{Y}^{**} = f(\hat{\lambda}_{opt})\sum_{i \in s_1} w_i Y_i + \hat{\lambda}_{opt}\sum_{i \in s_2} w_i Y_i$$

To define the adjustment for the influential observations, Chambers and Ren first define the contribution that the outliers make to the estimate of the total by

$$\hat{t}_2 = \hat{Y}^{**} - \sum_{i \in s_1} w_i Y_i$$

Then Chambers and Ren have two options to consider for corrections to the influential observations in s_2 so that estimated total remains the same. Option 1 is really a ratio adjustment.

Option 1

$$Y_i^{**(1)} = \left\{ \begin{matrix} Y_i \dfrac{\hat{t}_2}{\sum_{j \in s_2} w_j Y_j}, i \in s_2 \\ Y_i, i \in s_1 \end{matrix} \right\}$$

Option 2

$$Y_i^{**(2)} = \left\{ \begin{matrix} Y_i \left[1 + w_i \dfrac{\hat{t}_2 - \sum_{j \in s_2} w_j Y_j}{\sum_{j \in s_2} w_j^2 Y_j} \right], i \in s_2 \\ Y_i, i \in s_1 \end{matrix} \right\}$$

The estimator of the total from the sample for option j = 1 and j = 2 is

$$\hat{Y}^{**(j)} = \sum_{i=1}^{n} w_i Y_i^{**(j)}$$

3.4 Weighted M-Estimation

Robust statistics are useful for studying influential observations because they relax the assumption in parametric statistics of a strict parametric model. They are designed to be insensitive to some departures from the model assumptions (Hampel et al. 1986). In contrast, nonparametric statistics relax the underlying assumption in a different way by assuming continuity and symmetry instead of a parametric model. M-estimators (Huber 1964) are robust estimators that come from a generalization of maximum likelihood estimation.

Next is a brief overview of the underlying principles of M-estimation. The maximum likelihood estimator is defined as the value $T_n = T_n(X_1, ..., X_n)$ of the parameter θ that maximizes the joint distribution of the observations X_1, ..., X_n of a random variable X with a density function f_θ as follows:

$$\prod_{i=1}^{n} f_{T_n}(X_i) .$$

This is equivalent to minimizing

$\sum_{i=1}^{n} [-\ln(f_{T_n}(X_i)]$ where ln denotes the natural logarithm.

M-estimation generalizes this by replacing -ln with a function ρ and defining T_n to be the value that minimizes the sum :

$$\sum_{i=1}^{n} [\rho(X_i, T_n)] .$$

The function ρ has a derivation ψ such that the value T_n that minimizes the sum also satisfies

$$\sum_{i=1}^{n} \psi(X_i, T_n) = 0 .$$

The application of M-estimation examined in this investigation is regression estimation. The weighted M-estimation technique proposed by Beaumont and Alavi (2004) is able to adjust the weights or the values of the influential observations. The approach for adjusting the values uses a compromise between the generalized regression estimator and the best linear unbiased estimator of the population total (Beaumont and Alavi 2004, Beaumont 2004). SAS macros have been made available by Jean-Francois Beaumont.

Briefly, the method estimates $\hat{B}^M$ which is implicitly defined by

$$\sum_{i \in s} w_i^*(\hat{B}^M)(y_i - x_i \hat{B}^M)\frac{x_i}{v_i} = 0 \quad \text{where}$$

$$v_i = \lambda x_i$$

$$w_i^* = w_i \psi\{r_i(\hat{B}^M)\} / r_i(\hat{B}^M)$$

$$r_i(\hat{B}^M) = h_i e_i(\hat{B}^M) / Q\sqrt{v_i}$$

$$e_i(\hat{B}^M) = y_i - x_i \hat{B}^M$$

Q is a constant that is specified. The variable w_i is the survey weight, which may or may not be the inverse of the probability of selection. The variable h_i is a weight that may or may not be a function of x_i.

The function ψ may have a two-sided or one-sided form. An example of a one-sided form, called the Type II Huber function, is

$$\psi\{r_i(\hat{B}^M)\} = \begin{Bmatrix} r_i(\hat{B}^M), r_i(\hat{B}^M) \le \varphi \\ \dfrac{1}{w_i} r_i(\hat{B}^M) + \dfrac{(w_i - 1)}{w_i}\varphi, otherwise \end{Bmatrix}$$

where φ is a positive tuning constant. This form is equivalent to a winzorization of $r_i(\hat{B}^M)$.

Solving for $\hat{B}^M$ requires the Iteratively Reweighted Least-Squares algorithm in many circumstances. For certain choices of the weights and variables, the solution is the standard least-squares regression estimator.

The specification of the function ψ leads to three choices for adjusting the survey weights. The Type II Huber function weight adjustment, which is the default in Beaumont's program, for the ψ above is

$$w_i^*(\hat{B}^M) = \begin{Bmatrix} w_{i,}, r_i(\hat{B}^M) \le \varphi \\ 1 + (w_i - 1)\dfrac{\varphi}{r_i(\hat{B}^M)}, otherwise \end{Bmatrix}$$

For an adjustment to the influential value, Beaumont and Alavi (2004) use a weighted average of the robust prediction $x_i \hat{B}^M$ and the observed value y_i of the form

$$y_i^* = a_i y_i + (1 - a_i) x_i \hat{B}^M$$
$$where$$
$$a_i = w_i^*(\hat{B}^M) / w_i$$

Beaumont (2004) finds an optimal value of the tuning constant φ by deriving and then minimizing a design-based estimator of the mean-square error that does not require a model to hold for all the data as in the methods of Kokic and Bell (1994) or Clarke (1995). It does not require a model to hold for the influential value, in particular. Beaumont uses a numerical analysis algorithm to solve for the optimal value of the tuning constant φ.

Under particular choices of the variables and weights, Beaumont's method reduces to the Clarke method (1995) and the Kokic and Bell method (1994) discussed earlier in this paper.

The adjustment that corresponds to the Type II Huber function is

$$y^* = \frac{1}{w_i} y_i + \frac{(w_i - 1)}{w_i}\{x_i \hat{B}^M + \frac{\sqrt{v_i}}{h_i} Q\varphi\}$$

Whether adjusting the influential observation or its weight, we obtained approximately equal weighted estimates of total sales when we calibrated the weights to maintain their sum.

4. Results

Tables 1 and 2 show the results of applying the methods to data from the MRTS for a particular industry that had an influential value in a particular month. Also included is a definition of a Winsorization K as the mean plus twice the standard error using unweighted and weighted data for reference.

For the M-estimation, the method finds the optimal φ but calls for an initial value that we set equal to 2. We used the one-sided Type II Huber function ψ shown in Section 2.4. Also, we used the program default $Q = 1$, and set $v_i = x_i$ for all units in sample. This implies that

$$h_i = (w_i - 1)\sqrt{x_i}$$
$$r_i = (w_i - 1)(y_i - x_i \hat{B}^M)$$

Notice that r_i now has the same form as $\hat{D}_i$ in the Clarke method, called "Winsor by obs" in the tables. However, the *b* in the Clarke method and $\hat{B}^M$ in the M-estimation method usually will not be equal because they use different estimation methods.

Weight trimming appeared to produce a reasonable estimate of total sales, but the choice of the trimming factor *c* = 0.3333 was arbitrary. Weight trimming also requires an adjustment of the weights of the other observations in the stratum. In a large ongoing survey, changing other values often has disadvantages. Another disadvantage is that another method is needed to identify the influential values since weight trimming only makes an adjustment after an influential value is identified.

The method of defining the *K* by stratum (Kokic and Bell 1994) identified 51 influential observations, which are too many to be considered effective with our data.

The Clarke winsorization of defining a separate *K* for each observation identified exactly one influential value and seemed to produce a reasonable estimate of total sales. The resulting month-to-month change for the total was to the observed month-to-month change than M-estimation that adjusts the observation or M-estimation that adjusts the weight.

The winsorization that defined *K* as the mean plus twice the standard error with unweighted data identified no influential values so it was not effective in this case. However, when the same definition for *K* used weighted data, the method identified four influential values. The three additional values are not extreme enough to be considered influential. The resulting month-to-month change for the total was further from the observed month-to-month change than M-estimation by observation, M-estimation by weight, or winsorization for each observation.

Reverse calibration did not provide an adjustment for the influential value. Evidently the situation is too extreme for it to be effective.

M-estimation identified one influential value. The method that adjusts the influential value produced the month-to-month change closest to the observed month-to-month change for the total. The method that adjusts the weight for the influential value produced the same month-to-month change as the method that changed the value of the influential observation.

5. Summary

These results illustrate our investigations of these methods with data from other month and other industries in the MRTS. Winsorization by each observation, M-estimation by observation, and M-estimation by weight appear to have the most promise. M-estimation has the feature of being able to be deployed as a two-sided method that raises weighted low influential values as well as lowering weighted high influential values. The winsorization by each observation only applies to weighted high influential observations.

Subsequent research focuses on these three methods and expands to other industries and views the effect of adjustments on a continuing basis. In addition, consideration of whether the methods are able to handle situations where more than one influential value is present work. Some of the results are contained in a paper by Mulry and Feldpausch (2007).

References

Beaumont, J.-F. (2004) "Robust Estimation of a Finite Population Total in the Presence of Influential Units". Report for the Office for National Statistics, dated July 23, 2004. Office for National Statistics, Newport, U.K.

Beaumont, J.-F. And Alavi, A. (2004) "Robust Generalized Regression Estimation". *Survey Methodology*, 30, 2. 195-208.

Black, J. (2001) "Changes in Sampling Units in Surveys of Businesses". *Proceedings of the Federal Committee on Statistical Methods Research Conference*. Office of Management and Budget. Washington, DC. http://www.fcsm.gov/01papers/Black.pdf

Chambers, R. L. and Ren, R. (2004) "Outlier Robust Imputation of Survey Data". *2004 Proceedings of the American Statistical Association, Section on Survey Research Methods* [CD-ROM]. American Statistical Association. Alexandria, VA. 3336-3344.

Chambers, R. L., Hentges, Adao, and Zhao, Xinquiang (2003) "Robust Automatic Methods for Outlier and Error Detection". S^3RI Methodology Working Paper M03/17. Southampton Statistical Sciences Research Institute, Univ of Southampton, U.K. http://www.s3ri.soton.ac.uk/publications/methodology.php?s=20

Chambers, R. L., Kokic, P., Smith, P. And Cruddas, M. (2000) "Winsorization for Identifying and Treating Outliers in Business Surveys". *Proceedings of the 2nd International Conference on Establishment Surveys.* Statistics Canada. Ottawa, Canada. 717-726.

Clarke, M. (1995) "Winsorization Methods in Sample Surveys". Masters Thesis. Department of Statistics. Australia National University.

Cohen, T. and Spencer, B. D. (1991) "Shrinkage Weights for Unequal Probability Samples". *Proceedings of the Section on Survey Research Methods.* American Statistical Association. Alexandria, VA. 625 - 630.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Werner, S. A. (1986) *Robust Statistics. An Approach Based on Influence Functions.* John Wiley & Sons. New York, NY.

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

Hunt, J. W., Johnson, J. S., and King, K. S. (1999) "Detecting Outliers in the Monthly Retail Trade Survey Using the Hidiroglou-Berthelot Method". *Proceedings of the Section on Survey Research Methods.* American Statistical Association. Alexandria, VA. 539-543.

Huber, P. J. (1964) "Robust Estimation of a location parameter". *Annals of Mathematical Statistics*. Institute of Mathematical Statistics. 35. 73-101.

Kokic, P. N. and Bell, P. A. (1994) "Optimal Winsorising Cut-Offs for a Stratified Finite Population Estimator". *Journal of Official Statistics.* Stockholm, Sweden. 10. 419-435.

Latouche, M. and Berthelot, J.-M. (1992) "Use of a Score Function to Prioritize and Limit Recontacts in Editing Business Surveys". *Journal of Official Statistics*. Stockholm, Sweden. 8. 389-400.

Mulry, M. H. and Feldpausch, R. M. (2007) "Treating Influential Values in a Monthly Retail Trade Survey". Paper prepared for presentation at the Annual Meeting of the Statistical Society of Canada, St. John's, Newfoundland, Canada, June 10-13. U.S. Census Bureau, Washington, DC.

Potter, F. (1988) "Survey of Procedures to Control Extreme Sampling Weights". *Proceedings of the Section on Survey Research Methods.* American Statistical Association. Alexandria, VA. 453-458.

Ren, R. and Chambers, R. L. (2003) "Outlier Robust Imputation of Survey Data via Reverse Calibration". S[3]RI Methodology Working Paper M03/19. Southampton Statistical Sciences Research Institute, University of Southampton, U.K. http://www.s3ri.soton.ac.uk/publications/methodology.php?s=20

Searls, D. T. (1966) "An Estimator Which Reduces Large True Observations". *Journal of the American Statistical Association*. 61. 1200-1204.

Stokes, S. L. (1990) "A Comparison of Truncation and Shrinking of Sampling Weights". *Proceedings of the 1990 Annual Research Conference.* U.S. Census Bureau. Washington, DC. 463 - 471.

Table 1. Results for total and month-to-month change for different treatments of influential values

Method	Total (billions)	Month-to-month percent change	Number of Influential Units Identified
previous month	42.4		
current month	38.6	-0.090	
weight trim	38.3	-0.097	1*
Winsor by stratum	37.6	-0.113	51
Winsor by obs	38.5	-0.092	1
Winsor $\mu + 2s$	38.6	-0.090	0
Winsor wgt $\mu + 2s$	38.2	-0.099	4
Reverse Calibration	38.7	-0.087	1*
M-est obs	38.4	-0.094	1
M-est wgt	38.4	-0.094	1

*Method does not detect influential units, one influential unit was specified

Table 2. Results for different treatments for the influential value of sales

	Value (millions)	Weight	Weighted Value (millions)
previous month	0.57	55	31
current month	7.50	55	413
weight trim[1]	7.50	18	135
Winsor by obs	4.00	55	220
Winsor wgt $\mu + 2s$ [2]	1.60	55	87
M-est obs	4.30	55	234
M-est wgt	7.50	30	225

[1] Weight trimming adjusts the other 18 weights in the stratum by a factor of 1.020.

[2]Winsor wgt $\mu + 2s$ identified three other values.