WP. 31
ENGLISH ONLY

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Bonn, Germany, 25-27 September 2006)

Topic (v): New and emerging methods

HANDLING OF OUTLIERS AT SFSO

Supporting paper

Submitted by Swiss Federal Statistical Office[1]

I. INTRODUCTION

1. The present paper shows the experience with robust estimators and robust methods for detection of outliers at the Swiss Federal Statistical Office. The evolution of the robust methods since the beginning of the 1990 is pointed out by different applications in several statistics at SFSO. The main methods used are shortly discussed. In a further paragraph the lessons learned are described, followed by a short description of the software developed for robust estimation, detection of outliers and nearest neighbor imputation based on a robust multivariate distance function.

II. EXPERIENCE AT SFSO

A. Survey on Housing Rents 1993

2. The quarterly survey on housing rents is an important source of data for the Swiss Consumer Price Index (CPI). It accounts for 21% of the total CPI weight. The redesign in 1993 had as its main objective a much more up to date representation of the Swiss housing, in particular recently built houses and apartments. A sample design with Bernoulli-sampling was applied. A second objective was to dominate better the problem with outliers in the survey data. Housing rents have a very skew distribution and mistakes when reporting yearly housing rent instead of monthly housing rent as well as simple transcription errors occur.

3. In addition to the traditional weighted means for housing rents trimmed means and a trimming one-step W-estimator was developed (Hulliger 1995a). The definition involved the housing rents $y_i, i = 1, \ldots, n$, the sampling weight w_i and two parameters $0 \leq \alpha_l \leq 0.5 \leq \alpha_u \leq 1$. Suppose the housing rents y_i are sorted in ascending order. Then the two indices $i_l = \min\{i : \sum_{j=1}^{i} w_j / \sum_{j=1}^{n} w_j \geq \alpha_l\}$ and $i_u = \max\{i : \sum_{j=1}^{i} w_j / \sum_{j=1}^{n} w_j < \alpha_u\}$ are determined. The trimmed mean is $T(y_S, \alpha_l, \alpha_u) =$

[1] Prepared by Beat Hulliger (Beat.Hulliger@bfs.admin.ch) and Daniel Kilchmann (Daniel.Kilchmann@bfs.admin.ch).

$\sum_{j=i_l}^{i_u} w_j y_j / \sum_{j=i_l}^{i_u} w_j$. For the trimming one-step estimator a robustness weight u_i was calculated with

$$u_i = \frac{\max(y_{i_l}, \min(y_{i_u}, y_i)) - T(y_S, \alpha_l, \alpha_u)}{y_i - T(y_S, \alpha_l, \alpha_u)} \tag{1}$$

Then the trimming one-step estimator is $T_{1W}(y_S, \alpha_l, \alpha_u) = \frac{\sum_{j=1}^{n} w_j u_j y_j}{\sum_{j=1}^{n} w_j u_j}$. A crude variance approximation for the two estimators which did not differentiate between sampling and robustness weights was used.

4. The implementation of these estimators was very cumbersome due to the properties of the data base language ADABAS. The estimators were tested intensively with trial data and on the first round of the survey. It turned out that the robust estimator gave very similar results as the non-robust weighted one. Furthermore the gain in variance of the robust estimators (at least as far as the crude variance estimators showed it) was negligible. Closer inspection showed that the clerical editors of the survey actually did such a good job that no outliers wer left in the final data which were used for extrapolation. Therefore, in spite of having the robust estimators at disposition, it was actually the weighted mean which was used for calculating the input to the CPI.

B. **Earning Structure Survey**

5. When the revised Swiss Earning Structure Survey was introduced in 1994 an intensive discussion took place about the correct estimand to use. The up-to-then traditionally used mean was abandoned and quantiles were used as the estimands, in particular the median. Therefore (weighted) quantiles were used as estimators and outliers were not considered problematic anymore (very small and very large quantiles were avoided) (Peters and Hulliger 1996). Though the method of Woodruff, back-transforming a confidence interval with the inverse of the empirical distribution function on the probability, was proposed for implementation in (Peters and Hulliger 1996) the estimation of the variance of these quantiles was for long handled with a rough normal approximation. Only in 2000 the Woodruff method was actually refined and implemented (Graf 2002)

C. **Enterprise census 1995**

6. The Swiss Enterprise Census of 1995 had as a novelty, that the businesses should indicate their turnover in the questionnaire. About 21% of the enterprises did not answer this question. Multiple imputation based on propensity scores and a ratio imputation within cells was used (Hüsler and Müller 2001). The main problem were outliers in the remaining data. Besides traditional ratio imputations ad-hoc robust ratio imputations were used. It turned out that the variance of aggregates due to the imputations was negligible compared with sensitivity to the tuning constant used in the robust estimates. In other words, the outliers were so influential that without treating them manually the uncertainty about the aggregates was considered too large for publication.

D. **Environment protection expenditures 1993**

7. In the pilot survey on environment protection expenditures of Swiss enterprises 1993 the problem of outliers was recognized. However, the implementation was difficult because the survey was very small and modeling had to be done on the level of branch of economic activity (INFRAS 1996). Combined ratio estimates were used and the residuals from this non-robust model werde scrutinized. Observations with residuals larger then 3 or 5 times the residual standard error were eliminated from the extrapolation. This is not a robust procedure since both the original ratio as well as the residual

standard error are non-robust statistics. Nevertheless, the approach shows that the outlier problem was taken seriously.

8. The data of the Environment Protection Expenditure Survey was reanalysed in (Hulliger and Kassab 1998). A robust one-step procedure was used and, in particular, the semi-continuous distribution of the expenditures was treated explicitly. The problem was that many expenditures were 0, while the positive expenditures had a very skew distribution, resulting in a semi-continuous distribution. For outlier detection the probability of a 0 expenditure was estimated and robust estimation was applied only for the positive expenditures. For the aggregates the two parts of the distribution had to be joined again. The robust estimation was a one-step ratio estimator as described in (Hulliger 1999). However, the variance estimators were crude approximations at the most.

E. Survey on Production and Value Added 1998

9. The Swiss Survey on Production and Value Added (SPVA) is an annual survey based on a stratified random sample of enterprises. Outliers were detected and treated by a one-step ratio estimator adapted to sampling weights shortly discussed below, see (Hulliger 1995b) and (Hulliger 1999) for details of the estimators and (Peters, Renfer, and Hulliger) for the application in the SPVA98.

10. For the construction of the robust one-step ratio estimator we assume an auxiliary variable x_i, $x_i > 0, i \in$ the population U, positively correlated with the target variable y_i. And the population mean $\bar{x}_U$ is known. The construction of the estimator itself starts with a robust initial estimate, β_0, of the slope, e.g.

$$\beta_0 = \hat{\beta}_{\mathrm{med}} = \mathrm{med}(y_i, w_i)/\mathrm{med}(x_i, w_i) \tag{2}$$

with $\mathrm{med}(y_i, w_i) = T(y_s, 0.5, 0.5)$ the weighted median of y. The weighted median of the slopes $\mathrm{med}(y_i/x_i, w_i)$ could be used as another starting value. However, it seems to downweight the observations with large x_i too heavily and thus looses too much efficiency. In the second phase of the construction of the one-step ratio estimator the scale, σ, of the residuals is estimated by the weighted median of the absolute standardized residuals

$$\hat{\sigma} = \mathrm{med}\left(|r_i(\beta_0)|, w_i\right) \tag{3}$$

with the standardized residual $r_i(\beta_0) = (y_i - \beta_0 x_i)/\sqrt{x_i}$.

11. Robustness weights are defined as

$$u_i = \begin{cases} 1 & \text{if } |r_i(\beta_0)| \le c\hat{\sigma} \\ c\hat{\sigma}/|r_i(\beta_0)| & \text{if } |r_i(\beta_0)| > c\hat{\sigma} \end{cases} \tag{4}$$

These weights u_i robustify only against extreme residuals. Extreme values of x_i may still have undue influence on the estimate. In practice, x_i is often negatively correlated with w_i because the underlying model $\mathbf{E}[y_i] \propto x_i$ is already taken into account in the sample design. In that case leverage values in x_i are at least partially compensated by low w_i.

12. Finally a robustness and sampling weighted estimate of the slope and the robust ratio estimate for the mean of y, $\bar{y}_U$, adapted to sampling weights are computed as follows:

$$\hat{\beta}_{RS} = \frac{\sum w_i u_i y_i}{\sum w_i u_i x_i}, \qquad T_{RS} = \bar{x}_U \hat{\beta}_{RS}. \tag{5}$$

TABLE 1. Different estimates of the SPVA for a particular economic activity. 'sme' denotes the stratified mean Hajek-estimator, 'ds' the univariate one-step estimator (6), 'cr' is the combined ratio estimator, 'rs' is the one-step ratio estimator (5) and 'rm' is the fully iterated M-estimator. T is the mean production, SD(T) its estimated standard deviation and CV its coefficient of variation. '$\bar{u}_S$' denotes the average robustness weight.

Statistic	**sme**	**ds**	**cr**	**rs**	**rm**
large c	0	40	0	20	20
T	2009.54	1811.35	2009.54	1984.19	1982.67
SD(T)	97.45	75.67	62.33	57.32	57.18
CV(T)	4.85	4.18	3.10	2.89	2.88
$\bar{u}_S$	0	96.40	0	99.56	99.52
moderate c	0	20	0	10	10
T	2009.54	1663.85	2009.54	1953.87	1946.00
SD(T)	97.45	73.02	62.33	48.46	47.48
CV(T)	4.85	4.39	3.10	2.48	2.44
$\bar{u}_S$	0	89.27	0	98.20	97.73

13. The estimated slope $\hat{\beta}_{RS}$ is a one-step estimator. It can be used as initial estimation of the slope for the next step of an iteration. The weighted M-estimate is found by convergence of the iteration algorithm.

14. (Hulliger 1999) developed also a univariate one-step estimator for the weighted mean of y:

$$T_{0S}(c) = \frac{\sum w_i u_i y_i}{\sum w_i u_i} \tag{6}$$

with the robustness weights

$$u_i = \begin{cases} 1 & \text{if } |y_i - T_0| \leq c\hat{\sigma} \\ c\hat{\sigma}/|y_i - T_0| & \text{if } |y_i - T_0| > c\hat{\sigma} \end{cases} \tag{7}$$

and the weighted robust scale estimate $\hat{\sigma} = \text{med}\left(|y_i - T_0|, w_i\right)/0.67$. Further iteration leads to a univariate M-estimate adapted to sampling weights.

15. The robust variance estimators adapted to sampling weights are derived from the implicit estimating equation (Hulliger 1999). They differentiate between robustness and sampling weights. The resulting estimator moderately underestimates the variance, mainly because the variance of the scale estimator is not taken into account.

16. In our example the one-step ratio estimator adapted for sampling weights reduced bias considerably, cf. 1. In comparison the univariate one-step estimator downweights the extreme values of y too much. The one-step ratio estimator is even more efficient in variance reduction. The coefficient of variation decreases accordingly from 4.85% to 2.89% with $c = 20$ and to 2.48% with $c = 10$. The gain resulting from fully iterating the M-estimator can be neglected for the estimation of production 1998. Mean robustness weight indicates how much downweighting a robust estimator uses. The univariate one-step estimator leads to a considerable downweighting with a moderate tuning constant $c = 20$. Several tuning constants were tried out and the function of the difference of the corresponding estimators to the stratified mean and the standard deviation of the estimators versus the tuning constant was plotted. A tuning constants that yields roughly a difference of half the standard deviation seemed to work well. The average robustness weights shown in table 1 indicate that for these data rather large tuning constants might be suitable, e.g. $c = 15$.

17. The robustness of the weighted median is undermined by large differences in the sampling weights. Some very large weights may dominate the weighted median completely. Sampling weights must then be robustified. Correct non-representative outliers, cf.(Chambers 1986), often have a sampling weight 1 and should be discarded from the data for the robustification step. Their final weight should usually equal their sampling weight.

F. Survey on energy consumption

18. The Swiss Survey on Energy Consumption (SEC) is an annual business survey based on a stratified random sample, cf. (Salamin 2005) and (Bendel, Scherer, Salamin, and Gülden 2006). The size and an adapted economic activity nomenclature were used as stratification variables. In exceptional cases the adaptation consisted in building classes which can be defined by a very detailed economic activity, comparable to NACE3, on the one hand and by groups of several classes comparable to NACE2 classes on the other hand. This classification leads to quite homogeneous strata regarding the energy consumption of the enterprises but rather inhomogeneous regarding sampling weights.

19. A slightly adapted version of the one-step ratio estimator (5) was applied in the SEC to estimate the total energy consumption in domains. The only difference to the construction described in E was that not a weighted median was used in (2) and (3) but the usual median. The inhomogeneous weights leading to dominant observations were the reason for doing so.

20. The choice of the tuning constant c depended on the homogeneity of the data inside the strata and was discussed with the subject matter specialists.

G. Survey on Income and Expenditure 1998

21. The Swiss Survey of Income and Expenditure (SIE) collects information about household consumption to establish the consumption basket of a Swiss standard household which is used for weighting the Consumer Price Index. The SIE is an annual stratified survey of about 10'000 responding households. Unit nonresponse is treated by weighting the respondents accordingly. The global item nonresponse rate is about 5.5%.

22. The problem of detection and handling of outliers and missing values in the SIE98 was analyzed in (Oetliker 2002). Different types of robust regression methods were used to model the income with the expenditure. Based on the standardized residuals resulting from these regression models rejection rules for outlier detection were defined.

23. A logarithmic transformation was applied to reduce the asymmetry of the data before testing the different regression models. The least absolute deviation regression, LAD-regression or L_1-regression, minimises the sum of the absolute residuals at $\hat{\beta}_{L_1}$: $\min_\beta \sum_i |r_i|$, where the residual of observation i is $r_i = y_i - x_i^T \hat{\beta}$. Further, least median of squares regression, least trimmed squares regression, robust MM regression estimator implemented in S-Plus, cf. (MathSoft 1999) and (Yohai, A., and Zamar 1991), and the BACON[2] algorithm for regression data, a forward search algorithm cf. (Billor, Hadi, and Vellemann 2000), were applied. The robust regression methods improved the quality of the model 'income = expenditure + $Varepsilon$' considerably in comparison to non-robust regression.

[2] Blocked Adaptative Computationally-efficient Outlier Nominators

24. The lower and the upper limits for outlier detection were defined as $l = r_{0.25} - 1.5 \times (r_{0.75} - r_{0.25})$ and $u = r_{0.75} + 1.5 \times (r_{0.75} - r_{0.25})$, where the p-th quantile of the distribution of the residuals is noted r_p. These limits correspond to the boxplot limits. 0.7% of normal distributed data would be flagged as outliers with these limits, 0.35% on each side.

25. All robust methods performed similarly in the detection of outliers in income data. Finally, L_1-regression was chosen because of its slightly more satisfying performance compared to the other robust methods.

26. In a second step outliers in expenditures for several consumption goods were detected. On the one hand the 99.7%-quantile was used as upper limit for expenditure and on the other hand the 0.3%-quantile and the 99.7%-quantile were used as lower und upper limits for the expentiture/quantity ratio. The rejection area was then defined by the combination of these rejection rules. Outliers and both univariate missing values were imputed using the L_1-regression. If expenditure and quantity were missing then the means of the non-outlying values was imputed.

27. In homogenous expenditure categories these methods were very accurate. In categories with less homogeneous content, the regression model is less appropriate for imputation.

H. Retail trade statistics 2001

28. The Retail Trade Statistics is a monthly survey based on a random sample, stratified by the size and the economic activity nomenclature (NOGA). The main variable of this survey is turnover. The sample design used a Neyman allocation based on register data of the number of employees and turnover collected in the enterprise census 1995, cf. paragraph C and (Hüsler and Müller 2001). Turnover values with kilo errors and 'randomly' added one digit errors had to be treated before the elaboration of the sampling design, (Renfer 2006). The initial correlation between the number of employees and turnover before this treatment was only 0.07.

29. L_1-regression was used to define a robust slope $\hat{\beta}_{L_1}$. Observations with $y > c\hat{\beta}_{L_1}$ were flagged as outliers. After testing several values for c on the 70 strata, it was agreed that the most appropriate value was $c = 4$, leading to more or less 4% outliers. The correlation between the number of employees and turnover climbed to 0.97 by leaving out the outlying observations.

I. Swiss Hospital Statistics 1999-2004

30. The Swiss Hospital Statistics (SHS) covers the statistics of hospitals (SH) and social medical institutions (SMI). These two annual exhaustif surveys are mainly used to analyse the healthcare costs and rentability of hospitals and institutions. Several institutions entered the survey, vanished and re-entered between 1999 and 2004. In 2005 and 2006 missing values for the surveys of 1999 to 2004 were imputed separately for the SH and the SMI with the objective to have complete data for about 40 variables in each survey and for each year (Kilchmann 2006). Furthermore, consistency for the annual and longitudinal totals was asked on canton and institution type level. These totals had already been published and discussed. The unit nonresponse rate was given in a separate table. The main goal for the subject matter specialists was to be able to calculate about 25 indicators based on totals for each statistic and year after the imputation. This condition results in a multivariate imputation problem.

31. Investigation of the data showed that the data contained representative outliers, detected by a univariate one-step estimator for each year (Huber 1981), the Hidiroglou-Berthelot algorithm

(Hidiroglou and Berthelot 1986) and multivariate outlier detection for each year based on transformed rank correlations, TRC, see below and (EUREDIT Project 2004b), (EUREDIT Project 2004a) and (Béguin and Hulliger 2004). Most of the outliers were detected by all three methods but each method detected particular outliers in addition. Outliers were excluded from the imputation as donors.

32. Furthermore, most of the data with missing values did not have any data at all or only a few. Some of the variables have mainly zeros and only a few have positive values. Hence, zeros must be imputed to respect the covariance matrix as much as possible. The 'last carried forward' method was not appropriate because it generated inconsistencies concerning filter variables, of the questionnaire, which could change easily from year to year. Therefore, most of the missing values could not be imputed.

33. It was agreed to apply a nearest neighbor method (NNI) with robust multivariate distance function. The NNI allows to impute zeros and enhances respect of the covariance matrix. Representative outliers with missing values could only be imputed without creating a huge amount of outliers in additional directions by using a robust multivariate distance function.

34. Using the covariance matrix of TRC, the marginal Mahalanobis distance was defined on the observed values (Béguin and Hulliger 2004):

$$\mathrm{MD}_{marg} = \frac{p}{q}(x_o - m_o)^T S_{oo}^{-1}(x_o - m_o) \tag{8}$$

where p is the total number of variables, q the number of variables the distance was calculated on, x_o the observed part of the variables, m_o the median vector of the observed variables and S_{oo} the part of the covariance matrix corresponding to x_o. S_{oo} is based on the transformation of the Spearman rank correlations adapted to sampling weights. MD_{marg} was used inside the nearest neighbor algorithm where the donor was chosen randomly among the c nearest neighbors.

35. The eigenvalue of S_{oo} for variables with few positive observations was 0 resulting in a non-positive definite covariance matrix. These variables had to be excluded from the calculations of the distance function.

36. Compared to a nearest neighbor imputation with an L_1 distance function the NNI coupled with the TRC gave more reliable results. On the one hand, the number of outliers did not change substantially, and on the other hand the 25 indicators did not behave in an unexpected way as they did before imputation and with the L_1-NNI.

III. Lessons learned

37. Impact of sampling weights

The impact of sampling weights on the robustness properties of estimates can be severe. The weighted median, for example, may depend on very few observations. In particular in business surveys, where the range of sampling weights may be rather big, the median may be completely dominated by the smallest businesses and an outlier among them can completely destroy the robustness of the weighted median.

The range of sampling weights and the distribution of them should always be analysed. In particular, at SFSO, the median-dominance overall and in domains is used as an indicator for possible threats to robustness from sampling weights. The median-dominance is defined as the smallest number of observations from the upper tail of the weight distribution which have more than 50% of the total weight. More formally let $w_{(i)}, i = 1, \ldots, n$ denote the order statistics of the sampling weights. Then the median-dominance is defined as

$$\mathrm{dom}_{0.5} = \min\{k : \sum_{i=n-k+1}^{n} w_{(i)} / \sum_{i=1}^{n} w_i \geq 0.5\} \tag{9}$$

The lower $\mathrm{dom}_{0.5}$ is the more unbalanced are the weights. If the median-dominance is, e.g., 30% then the empirical breakdown point of the weighted median is 30% instead of the 50% of an unweighted median. When weighted medians are calculated for domains then the median-dominance may be much lower even.

38. Total weight
Another problem is that sometimes the resulting robustness weights are so small that the product of robustness weight u_i and sampling weight w_i, say the total weight $w_i u_i$, becomes much smaller than 1. This may be undesirable for outliers which are correct observations. The problem becomes visible when the weighted total of a variable is in fact smaller than the unweighted sum of the variable over the sample. Therefore, it may be necessary to establish a lower limit (1 or slightly lower, like 0.2) to avoid total weights close to 0.

39. Models
Since the definition of outliers depends on models these must be checked carefully. For example a ratio estimator is not very vulnerable to deviations from a regression-through-the-origin model if there are no outliers. However, the capability of outlier detection rapidly deteriorates if the model is not adequate for the bulk of the data. And, of course, outliers may have very large effects on ratio-estimators.

40. Choice of tuning constants
For asymmetric distributions robust procedures usually lead to a bias because correct observations may be downweighted. This is compensated by a gain in variance. The degree of robustification can be chosen by tuning constants. The choice of the tuning constant is not straightforward. External information may be helpful. It may be necessary to submit the results of an outlier detection procedure with certain tuning constants to the subject matter specialists and editors of the survey. By checking observations flagged as outliers they can give a feedback on whether the degree of robustification is reasonable. Generally the mean of the robustness weights should not be much below 1. In other words the robustification should be conservative, declaring only a very small fraction of the data as outlying. A plot of a robust estimator and its variance vs. the tuning constant may also help in choosing the tuning constant.

IV. **Software**

41. SFSO uses SAS as its standard statistical package. Algorithms which are used in production should be implemented in SAS. For developing purposes S-Plus and R are also used. We report here on the SAS programmes only.

42. One-step ratio estimator
The SAS programme `oneratx` implements a onestep ratio estimator. Additionally to limiting the influence of the residuals the influence of leverage points in the explanatory variable may be limited. Furthermore a lower limit for the total weight may be imposed. An approximate variance is calculated with the help of SAS procedure `surveymeans`. The new SAS procedure `robustreg` will be used to obtain a fully iterated robust ratio estimator.

43. TRC

The macro TRC (Transformed Rank Correlations) is a program for multivariate outlier detection with incomplete survey data and developed using SAS Version 8.2 . To run the macro, the SAS Modules Base, IML and Graph must be installed. The input is a SAS dataset with n observations and p variables and a sampling weight variable, which is not required. There are a number of parameters to control the execution of the macro: the robust regression method used inside TRC, a tuning constant being the minimal proportion of observations which are observed for both variables involved in the bivariate regression and the lower limit value for the robustness weight u indicating outlyingness. For the theoretical background of the implemented algorithms refer to (Béguin and Hulliger 2004), (EUREDIT Project 2004b) and (EUREDIT Project 2004a).

44. NNI

The SAS macro 'NNI' is an implementation of a nearest neighbor imputation algorithm. It is possible to use the Mahalanobis distance of the TRC macro mentioned above to calculate the distance between the recipient and the donors. Otherwise L_1 and matching distance is available. The size of the donor pool can be influenced by several parameters. Draft documentation is available in German only.

References

Béguin, C. and B. Hulliger (2004). Multivariate Oulier Detection in Incomplete Survey Data: The Epidemic Algorithm and Transformed Rank Correlations. *J.R.Statist.Soc.A 167*(Part 2.), 275–294.

Bendel, R., R. Scherer, P.-A. Salamin, and J. Gülden (2006). Energieverbrauch in der Industrie und im Dienstleistungssektor. Resultate 2002 bis 2004. Report, Swiss Federal Institute of Energie, Bern.

Billor, N., A. S. Hadi, and P. F. Vellemann (2000). BACON: Blocked Adaptative Computationally-efficient Outlier Nominators. *Computational Statistics and Data Analysis 34*(3), 279–298.

Chambers, R. (1986). Outlier Robust Finite Population Estimation. *Journal of the American Statistical Association 81*(396), 1063–1069.

EUREDIT Project (2004a). *Methods and Experimental Results from the Euredit Project*, Volume 2. http://www.cs.york.ac.uk/euredit/results/results.html.

EUREDIT Project (2004b). *Towards Effective Statistical Editing and Imputation Strategies - Findings of the Euredit project*, Volume 1. http://www.cs.york.ac.uk/euredit/results/results.html.

Graf, M. (2002). Enquête suiss sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé. Methodological Report 338-0010, Swiss Federal Statistical Office.

Hidiroglou, M. A. and J.-M. Berthelot (1986). Statistical Editing and Imputation for Periodic Business Surveys. *Survey Methodology 12*(1), 73–83.

Hüsler, J. and S. Müller (2001). Schlussbericht Betriebszählung 1995 (BZ95), Mehrfach imputierte Umsatzzahlen. Methodological report, Swiss Federal Statistical Office. Author affiliation: University of Bern.

Huber, P. J. (1981). *Robust Statistics*. New York: Wiley.

Hulliger, B. (1995a). Konjunkturelle Mietpreiserhebung: Stichprobenplan und Schätzverfahren. Methodological report, Swiss Federal Statistical Office.

Hulliger, B. (1995b). Outlier Robust Horvitz-Thompson Estimators. *Survey Methodology 21*(1), 79–87. Statistics Canada.

Hulliger, B. (1999). Simple and Robust Estimators for Sampling. In *Proceedings of the Section on Survey Research Methods*, pp. 54–63. American Statistical Association.

Hulliger, B. and M. Kassab (1998). Evaluation of Estimation Methods for the Survey on Environment Protection Expenditures of Swiss Communes. Methodological report, Swiss Federal Statistical Office.

INFRAS (1996). Umweltausgaben in der Schweiz, Technischer Bericht zur Piloterhebung. Technical report, Swiss Federal Statistical Office.

Kilchmann, D. (appears 2006). Krankenhausstatistik und Statistik der sozialmedizinischen Institutionen 1999-2004. Einsetzungsverfahren. Methodological Report 338-00XX, Swiss Federal Statistical Office, Neuchâtel.

MathSoft (1999). *S-Plus 2000 Guide to Statistics*. Seattle, Washington: MathSoft.

Oetliker, U. (2002, Août). Traitement des données manquantes et aberrantes dans le domaine des revenus de l'enquête sur les revenus et la consommation (ERC98). Mémoire du diplôme postgrade en statistique, Université de Neuchâtel.

Peters, R. and B. Hulliger (1996, June). Schätzverfahren für die Lohnstruktur-Erhebung. Methodological report, Swiss Federal Statistical Office.

Peters, R., J.-P. Renfer, and B. Hulliger. Technical report.

Renfer, J.-P. (appears 2006). Enquête sur les chiffres d'affaire du commerce de détail. Elaboration du plan d'échantillonnage et méthodes d'estimation. Methodological Report 338-00XX, Swiss Federal Statistical Office, Neuchâtel.

Salamin, P.-A. (2005). Extrapolation pour la statistique de la consommation d'énergie. In *Swiss Statistics Meeting*.

Yohai, V., S. W. A., and R. H. Zamar (1991). A Procedure for Robust Estimation and Inference in Linear Regression. In S. W. Stahel W. A., Weisberg (Ed.), *Linear Regression*, Volume 2 of *Directions in Robust Statistics and Diagnostics*, New York. Springer-Verlag.