Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

Robust calibration estimators

by Pierre Duchesne

June 1999

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 1999
Vol. 25, No. 1, pp. 43-56
Statistics Canada, Catalogue No. 12-001

Robust calibration estimators

Pierre Duchesne [1]

Abstract

We consider the use of calibration estimators when outliers occur. An extension is obtained for the class of Deville and Särndal (1992) calibration estimators based on Wright (1983) QR estimators. It is also obtained by minimizing a general metric subject to constraints on the calibration variables and weights. As an application, this class of estimators helps us consider robust calibration estimators by choosing parameters carefully. This makes it possible, *e.g.*, for cosmetic reasons, to limit robust weights to a predetermined interval. The use of robust estimators with a high breakdown point is also considered. In the specific case of the mean square metric, the estimator proposed by the author is a generalization of a Lee (1991) proposition. The new methodology is illustrated by means of a short simulation study.

Key Words: Calibration estimator; Regression estimator; Range restrictions; Robustness.

1. Introduction

The problem of outliers is an important one in all branches of statistics. In sampling theory, the background is different from that of parametric statistics since the objective is often to estimate the total of a variable of interest y. An outlier may have its full weight within the population total. Moreover, methodologists may assume, at the estimation stage, that the values of units are recorded without error, since the gathered units are often processed within an editing system (Särndal, Swensson and Wretman 1992, section 1.7). This step is part of the sampling procedure in large statistical agencies such as Statistics Canada. Lee (1995) has provided an overview of robustness developments within sampling theory.

Nevertheless, since populations for economic surveys are often asymmetric, some units might be extreme as compared to others, as was discussed by Kish (1965). The complete elimination of such units would lead to biased estimates, while maintaining them with their full weight might make an estimator such as the generalized regression (GREG) estimator highly variable. This would suggest a compromise between bias and variance. When outliers occur, the challenge is to propose robust estimators of the total that are little affected by certain units that deviate sharply from others. Such estimators should have little bias and a small mean square error. Traditionally, sampling theory has been deeply involved in the development of unbiased or asymptotically design unbiased (ADU) estimators. See for example Särndal *et al.*, (1992, Section 7.12). However, this ADU property is perhaps undesirable within the context of outliers. This was discussed by Chambers and Kokic (1993), who showed the conflict between the ADU property and the robustness of an estimator.

We consider the Horvitz-Thompson (HT) estimator defined by $\hat{T}_{y\mathrm{HT}} = \sum_s d_k\, y_k$, where $d_k = \pi_k^{-1}$, π_k being the inclusion probability (If A is a set of units, $A \subseteq U$, then $\sum_A$ is a notation signifying $\sum_{k\in A}$). Let us assume a positive variable of interest y and an asymmetric population. As the HT estimator is a mean weighted by the d_k, it is vulnerable to large values of y. A unit with a high weight d_k may also have a considerable impact on the estimation step by including variable estimates. Lee (1995) defined these units as influential. An extreme unit is not necessarily influential if its weight d_k is sufficiently small. Traditionally, methodologists have sought to limit the impact of influential units when they are known prior to sampling, by assigning for example sampling weights close to 1 to extreme units. Gambino (1987) and Lee (1995) have nevertheless discussed situations in which this cannot be done. In a major article, Hidiroglou and Srinath (1981) considered changing the sampling weights when outliers occur. Their approach gave much legitimacy to weight modification within sampling procedures.

Many of the first robust alternatives to the total were based on M estimators and GM-estimators. Nevertheless, much interest has been shown recently for estimators that also provide good overall robustness, as measured by the breakdown point of an estimator. These concepts are discussed for example in Donoho and Huber (1983), Hampel, Ronchetti, Rousseeuw and Stahel (1986) and Rousseeuw and Leroy (1987). The breakdown point measures the percentage of outliers within the sample that the estimator can tolerate while providing nonetheless a good estimate of a given characteristic of the population. Lee, Ghangurde, Mach and Yung (1992) required estimators of the total that were based on robust estimators with a high breakdown point.

We will be considering calibration estimators of the total T_y written as $\sum_s w_k y_k$. These estimators were developed for example in Deville and Särndal (1992). We are looking for weights w_k that are as close as possible to sampling weights $d_k = \pi_k^{-1}$, while meeting benchmark constraints, denoted CE (also known as calibration constraints),

1. Pierre Duchesne, Département de Mathématiques et de Statistique, Université de Montréal, C.P. 6128, succursale centre-ville, Montréal, Québec, H3C 3J7.

$$\sum_s w_k x_k = T_x, \tag{1.1}$$

where x_k is a vector of dimension m that corresponds to the available auxiliary information of known total $T_x = \sum_U x_k$. These estimators are popular as they are easily interpreted, since methodologists are used to assigning weights w_k to units y_k. Several metrics are studied to measure the proximity between d_k and w_k. The GREG estimator is an important example with $w_k = d_k(1 + (T_x - \hat{T}_{x\mathrm{HT}})' M_s^{-1} x_k/c_k)$, where $M_s = \sum_s d_k x_k x_k'/c_k$. It is obtained by minimizing the mean square metric $\sum_s c_k(w_k - d_k)^2/d_k$. Constants c_k are weighting factors which can take into account problems of heteroscedasticity (for example). Särndal (1996) discussed the selection of these constants. However, since the *g*-weights $g_k = w_k/d_k$ of the GREG estimator are not generally restricted, other metrics are proposed as a means of limiting them so that they might meet certain constraints applicable to the range of values (CARV). Specifically, this makes it possible to avoid undesirable negative weights w_k. See Deville and Särndal (1992), Singh and Mohl (1996) and Stukel, Hidiroglou and Särndal (1996).

As was noted by Fuller, Loughin and Baker (1994, page 81), there is a link between calibration estimators and robust methods. However, it is wrong to assume that calibration estimators necessarily have good properties of robustness, given that all the calibration estimators considered by Deville and Särndal (1992) were asymptotically equivalent to the GREG estimator, which, being ADU, is not robust. Moreover, a traditional calibration estimator is not robust as it depends linearly on w_k, and w_k and does not take into account y_k.

The purpose of this paper is to build estimators in the form of $\sum_s w_k y_k$ where the weights w_k provide robustness while meeting constraints on the calibration variables and the weights w_k. The starting point of our approach is the class of Wright (1983) estimators QR. Let us assume we have available constants $\{(q_k, r_k), q_k > 0, r_k \geq 0, \forall k \in U\}$, such that $\sum_U \pi_k q_k x_k x_k' > 0$ and $\sum_s q_k x_k x_k' > 0, \forall s$. (If A is a symmetric matrix, $A > 0$ means that A is definite as positive.) The QR estimators are defined on the basis of q_k and r_k by the relation

$$\hat{T}_{y\mathrm{QR}} = T_x' \hat{B}_q + \sum_s r_k e_k, \tag{1.2}$$

where $\hat{B}_q$ assumes a form weighted by the q_k

$$\hat{B}_q = \left(\sum_s q_k x_k x_k'\right)^{-1} \sum_s q_k x_k y_k, \tag{1.3}$$

and

$$e_k = y_k - x_k' \hat{B}_q. \tag{1.4}$$

It will be shown in section 2 that the QR estimators are calibration estimators, and a new class of estimators, denoted RQR, will be introduced, also based on the choice of constants q_k and r_k. It generalizes to a certain extent the QR estimators as well as the class of Deville and Särndal (1992) estimators. The RQR class is interesting in that it makes it possible to obtain weights w_k that are limited to a given interval, say $[L, U]$. Some of the properties of classes QR and RQR are provided in section 2.

Section 3 describes applications of the RQR class in the building of robust calibration estimators. The main goal is to modify robust default weights so that they meet calibration constraints. Section 3.1 discusses the choice of constants q_k and r_k using arguments suited to calibration estimators. This is a new and unifying approach, and in section 3.2 it guides our choice of q_k and r_k when there is auxiliary information. One important element is the use of a robust estimator allowing for the weighted form (1.3), providing the q_k. Note that this is the case for GM-estimators. Usually, estimators with a high breakdown point do not have a weighted form. Consideration is given to reweighting these estimators, allowing the breakdown point to be kept under control and making it possible to have estimators written in the form (1.3). See Rousseeuw and Leroy (1987) and Simpson and Chang (1997). We then discuss the choice of q_k and r_k, so as to calculate an RQR estimator and obtain a robust calibration estimator with restricted weights. Various robust estimators, including the Lee (1991) estimator and the Chambers (1986) estimator, are compared in section 4 with RQR estimators as well as with the GREG estimator and one calibration estimator considered in Deville and Särndal (1992) whose weights are limited. The Lee (1991) estimator can be considered a specific case of our approach. It allows us to also consider a new estimator with restricted weights. Four populations that have already been studied in the literature are considered. It will be noted that estimators free of weight constraints are subject to negative weighting problems. With the RQR class of estimators, robust estimators having positive weights can be obtained, and they compare well with estimators free of weight constraints. Finally, conclusions are drawn in section 5. Appendix B contains a list of abbreviations, and Appendix C contains a list of the various constants found in this paper, with definitions.

2. RQR class estimators

Consider a finite population $U = \{1, 2, ..., N\}$ of size N whose total $T_y = \sum_U y_k$ we wish to estimate for a variable of interest y that is positive. A sample s of size n_s is drawn following a sampling design $p(s)$. The inclusion probability of a unit k is denoted π_k, and the second-order inclusion probabilities are denoted π_{kl}. We assume that the auxiliary information x_k is of unit value, *i.e.*, x_k is known from a reliable source $\forall k \in U$.

Wright (1983) introduced a class of QR estimators written in the form (1.2) with the primary objective of unifying a large number of common estimators. We find the best linear unbiased prediction (BLUP) estimator of Royall (1970) derived from the model-based theory,

obtained by assuming $(q_k, r_k) = (1/c_k, 1)$, and the GREG estimator of Cassel, Särndal and Wretman (1976) by considering the choice $(q_k, r_k) = (d_k/c_k, d_k)$. Alternately, (1.2) can be written as

$$\hat{T}_{y\mathrm{QR}} = \sum_s d_k\, g_k\, y_k,$$

where $d_k\, g_k$ satisfies

$$d_k\, g_k = r_k + (T_x - \hat{T}_{xr})' \Big(\sum_s q_k\, x_k\, x_k'\Big)^{-1} q_k\, x_k, \quad (2.1)$$

with $\hat{T}_{xr} = \sum_s r_k\, x_k$. Assuming $r_k = d_k$, g_k corresponds to the *g*-weight of the GREG estimator.

The QR estimators are calibration estimators, obtained by minimizing the mean square metric subject to the CEs

$$\min \frac{1}{2} \sum_s (w_k - r_k)^2 / q_k, \text{ as of } \sum_s w_k x_k = T_x. \quad (2.2)$$

The weights w_k are chosen as close as possible to the r_k and the q_k are weighting factors. In other words, the starting weights r_k are transformed into calibration weights w_k. The solution to problem (2.2) is $w_k = d_k\, g_k$, where $d_k\, g_k$ is given by the formula (2.1).

Nothing, however, guarantees that the weights w_k of the QR estimator are positive, which might be undesirable in practice. See Brewer (1994), who formalized the interpretation of weights. To limit the weights w_k in $[L, U]$, we wish to resolve

$$\min \sum_s G(w_k; q_k, r_k), \quad \text{as of } \sum_s w_k\, x_k = T_x \text{ and } w_k \in [L, U]. \quad (2.3)$$

The calibration estimator of the total is

$$\hat{T}_{y\mathrm{RQR}} = \sum_s w_k\, y_k, \quad (2.4)$$

where the w_k are obtained by resolving problem (2.3). It is assumed that function $G(w; q, r)$ is strictly convex and can be derived in w for fixed r and q. We denote $g(u; q, r) = G'(u; q, r)$ and $h(u; q, r) = g^{-1}(u; q, r)$. Moreover, it is assumed that $h(0; q, r) = r$ and $h'(0; q, r) = q$. The resulting estimators are called QR (RQR) restricted calibration estimators.

Fuller *et al.*, (1994) favoured regression estimators having reasonable invariance properties. It can be shown that RQR estimators are regression equivariant and to scale when constants q_k and r_k are transformation invariant. Useful definitions may be found in Bolfarine and Zacks (1992).

There is no guarantee that there is a solution to problem (2.3). We refer to the simulation study in Stukel *et al.*, (1996). There may, for example, be realizations of the sample for which even the CEs cannot be satisfied (1.1). Thus, the sample is so imbalanced that it is impossible for the weighted sum of the components for each dimension to provide the corresponding population total. The only recourse for the practitioner, then, is to relax the constraints by reducing the dimension of the number of auxiliary variables. See also the discussion in Fuller *et al.*, (1994). As for the calibration estimators considered in Deville and Särndal (1992), it was shown, in result 1, that there is a solution with a probability approaching one. Under certain conditions, this result can be adapted to class RQR estimators.

The metric on which we will focus our attention so that the weights may satisfy the CARVs is a slight modification of case No. 7 in Deville and Särndal (1992). We call it the restricted mean square metric. The *G*-function that corresponds to the choice of this metric is

$$G(w_k; q_k, r_k) = \begin{cases} \dfrac{1}{2}(w_k - r_k)^2/q_k & \text{if } w_k \in [L, U], \\ \infty & \text{otherwise,} \end{cases}$$

whereas the *h*-function is

$$h(x_k'\lambda; q_k, r_k) = \begin{cases} L & r_k + q_k\, x_k'\lambda < L, \\ r_k + q_k\, x_k'\lambda & r_k + q_k\, x_k'\lambda \in [L, U], \\ U & r_k + q_k\, x_k'\lambda > U. \end{cases}$$

Given this modification, it is the weight w_k that is constrained and not only w_k / d_k as for case No. 7 in Deville and Särndal (1992). In our situation, w_k can "correct" an initial weight that is an outlier. It will be noted that, as it is formulated, the Deville and Särndal metric (1992) subtly inserts the constraints on the w_k in the *G*-function. In order to calculate the estimator (2.4) according to this metric, it is sufficient to follow the same approach as Deville and Särndal (1992), which leads us to a solution, using Newton's method, for the following equation in λ

$$\sum_s h(x_k'\lambda; q_k, r_k)\, x_k = T_x. \quad (2.5)$$

The final estimator is $\hat{T}_{y\mathrm{RQR}} = \sum_s h(x_k'\lambda_\infty; q_k, r_k)\, y_k$, where λ_∞ is the solution to equation (2.5).

It is interesting to know whether the weight constraint changes the properties of the estimator as compared to a QR estimator that is free of weight constraints. The following result (as proven in the Appendix) shows that, under certain conditions, the two estimators are asymptotically equivalent. In practice, using the restricted mean square metric, we have not observed any significant deviations.

Proposition 1. According to hypotheses C_1 and C_2 given in the Appendix,

$$N^{-1} |\, \hat{T}_{y\mathrm{QR}} - \hat{T}_{y\mathrm{RQR}} \,| = o_p(n^{-1/2}). \quad (2.6)$$

This result can possibly be obtained using the approach leading to result No. 5 in Deville and Särndal (1992) dealing with the asymptotic equivalence between the GREG and calibration estimators considered by the authors.

However, proposition 1 is of some use to understand the type of conditions needed to reach the result described in our situation.

Since (1.2) shows the same asymptotic behaviour as quantity $T_x' B_q + \sum_s w_k E_k$, where $E_k = y_k - x_k' B_q$ and $B_q = (\sum_U \pi_k q_k x_k x_k')^{-1} \sum_U \pi_k q_k x_k y_k$, this would suggest as variance estimator

$$v_L = \sum\sum_s \breve{\Delta}_{kl} (w_k e_k)(w_l e_l), \qquad (2.7)$$

where $\Delta_{kl} = \pi_{kl} - \pi_k \pi_l$, $\breve{\Delta}_{kl} = \Delta_{kl}/\pi_{kl}$ and e_k are given by (1.4). See Särndal *et al.*, (1989) and Särndal *et al.*, (1992, page 234). It can be shown that the asymptotic bias of a QR estimator is given under general conditions by

$$E_p(\hat{T}_{y\mathrm{QR}} - T_y) = \sum_U (\pi_k r_k - 1) E_k.$$

Then, a possible bias estimator is $b = \sum_s d_k (\pi_k r_k - 1) e_k$, which can be used in conjunction with formula (2.7) to build an estimator of the mean square error for a QR and RQR estimator, using proposition 1.

RQR estimators make it possible to obtain calibration estimators with constrained weights. Given set q_k and r_k, it is sufficient to resolve problem (2.3). In the sections which follow, the RQR class is applied within a context of robustness. We will show how to direct the selection of constants q_k and r_k, chosen in practice using sample s.

3. Building robust and calibrated estimators

3.1 Methods based on weight reduction and value modification

Lee (1995) discussed various propositions based on the weight reduction method for simple random sampling. Once outlier observations have been detected, these methods consist in reducing the weight of extreme observations. These methods are to be preferred to those which eliminate doubtful observations entirely, since all the observations in the sample are legitimate, as was discussed by Lee *et al.*, (1992).

With respect to calibration estimators, we begin by considering the situation in which there is no auxiliary information available and the only constraint is $\sum_s w_k = N$. This case will guide our path. Consider the QR estimator with $q_k = r_k$. For the sake of our discussion, we consider constants r_k, known and fixed. The weights minimizing (2.2) subject to $\sum_s w_k = N$ are $w_k = C_s(r) r_k$, where $C_s(r) = N/\sum_s r_k$, so that $\hat{T}_{y\mathrm{QR}}$ becomes

$$\hat{T}_{y\mathrm{QR}} = C_s(r) \sum_s r_k y_k. \qquad (3.1)$$

Whenever an observation is extreme, it might represent few units like itself within the population, and its weight should perhaps be reduced. In order to satisfy the CEs, this means finding weights w_k that come closest to the sampling weights d_k for units which are not outliers but come as close as possible to a reduction factor r for outlier units, where r is chosen by the statistician. Specifically, we denote $s = s_1 \cup s_2$, where s_1 of cardinality n_1 represents those units that are not reported as outliers, whereas $s_2 = s - s_1$ of cardinality $n_2 = n - n_1$ represents the outlier units of s. The reduction factor r will typically satisfy $r \le d_k$, $\forall k \in s_2$. For example, consider the estimator (3.1) with $q_k = r_k = B_k = d_k I_{k1} + r(1 - I_{k1})$, where I_{k1} is the variable indicating affiliation to s_1. In this way, constants q_k and r_k are reduced for units of s_2 so as to reflect the fact that units of s_2 are extreme. The estimator (3.1) becomes

$$\hat{T}_{y\mathrm{QR}} = C_s(B)\left(\sum_{s1} d_k y_k + r \sum_{s2} y_k\right).$$

In the case of simple random sampling, $d_k = N/n$ and we obtain

$$\hat{T}_{y\mathrm{QR}} = C_s(B) \hat{T}_{yB},$$

where $\hat{T}_{yB} = N/n \sum_{s1} y_k + r \sum_{s2} y_k$ is the Bershad (1960) estimator discussed in Lee (1995). Other methods based on weight reduction have been discussed in Lee (1995), who also discussed the choice of r.

One disadvantage of methods based on weight reduction is that the analyst must identify the outlier units. Methods based on value modification avoid this difficulty by providing gradual weight reduction for units that are more extreme. We consider a case of simple random sampling. We assume

$$m(y_k; t, a, b) = b + (a - b) \min(1, t/y_k). \qquad (3.2)$$

Thus, this function assigns a starting weight of value a for the $y_k < t$, and gradually reduces this to a final weight b, as y_k becomes extreme. Value t is called the threshold. The constants a, b and t are chosen by the statistician. Several values for a and b have been considered in the literature. Thus, instead of assigning a fixed reduction factor to the units of s_2, we select $q_k = r_k = W_k = m(y_k; t, N/n, fN/n)$, where f is a constant between 0 and 1. The estimator (3.1) becomes

$$\begin{aligned} \hat{T}_{y\mathrm{QR}} &= C_s(W) \sum_S W_k y_k \\ &= C_s(W) \hat{T}_{yW}. \end{aligned}$$

The estimator $\hat{T}_{yW}$ has been discussed in Gross, Taylor and Lloyd-Smith (1986) as well as in Chambers and Kokic (1993), who called it the winsorized estimator. This is a special case of the approach used by Chambers (1982, 1986). When $f = 0$, the estimator (3.1) becomes $\hat{T}_{y\mathrm{QR}} = C_s(W_I) \hat{T}_{yW1}$ with $q_k = r_k = W_{Ik} = m(y_k; t, N/n, 0)$,

where $\hat{T}_{yW1} = N/n(\sum_{s1} y_k + n_2 t)$, denoting the part of s containing units that satisfy $y_k < t$. The estimator $\hat{T}_{yW1}$ has been discussed in Lee (1995), as well as in Gross *et al*., (1986), who called it the type I winsorized estimator. When $f = n/N$, Gross *et al*., (1986) called it the type II winsorized estimator. It has also been discussed in Bruce (1991).

In a design πps, Dalén (1987) inserted the design by assuming $D_k = m(y_k; \pi_k t, d_k, 1)$. Thus, if k and l are two extreme observations such that $y_k = y_l$, then the observation whose sampling weight is largest will have a higher weight D_k. Selecting $r_k = q_k = D_k$ makes it possible to obtain essentially the Dalén estimator, $\hat{T}_{y\mathrm{QR}} = C_s(D)\sum_s D_k y_k$. The estimator $\hat{T}_{yD} = \sum_s D_k y_k$ has been studied for example in Tambay (1988).

Table 3.1
Estimator (3.1) based on weight reduction and value modification

Estimator	Value of $q_k = r_k$
Bershad	$B_k = d_k I_{k1} + r(1 - I_{k1})$
Winsorised	$W_k = m(y_k; t, N/n, fN/n)$
Winsorised, type I	$W_{Ik} = m(y_k; t, N/n, 0)$
Winsorised, type II	$W_{IIk} = m(y_k; t, N/n, 1)$
Dalén	$D_k = m(y_k; \pi_k t, d_k, 1)$

Note: $m(y_k; t, a, b) = b + (a - b)\min(1, t/y_k)$.

The approach used in this section suggests that we may occasionally seek estimators whose weights are close to r_k rather than the sampling weights d_k. The constants r_k will themselves be chosen close to d_k for the proper units, but will be reduced once a unit is declared extreme. The QR estimators allow the weight reduction and value modification methods to be unified. Methods based on value modification help us choose weights that are adapted to the specific sample s chosen. As was noted in Chambers and Kokic (1993), this is not surprising since the problem of outliers occurs after the selection of sample s. We must use the sample at our disposal to overcome the problem. These methods are generalized in the following section using auxiliary information.

3.2 Estimators of the total based on robust statistics

One of the first attempts to obtain robust alternatives to population totals using auxiliary information can be found in Chambers (1982, 1986), who proposed a robust ratio estimator based on BLUP estimator decomposition. One recent extension of the work carried out by Chambers can be found in Welsh and Ronchetti (1998). Gwet and Rivest (1992) also proposed a robust version of the ratio estimator using an approach based on the design in simple random sampling. Rivest and Rouillard (1991) carried out a comparative study of several robust estimators, and examined several estimators of the mean square error. For designs with unequal probabilities, Hulliger (1995) considered robustifying the HT estimator when inclusion probabilities are obtained using auxiliary information. Gwet and Rivest (1992) and Hulliger (1995) considered a version of the influence function for finite populations, emphasizing the need for procedures having good properties of local robustness and the use of estimators having limited influence functions. Influence functions were discussed generally in Hampel *et al*., (1986).

The following sections will deal with building robust estimators having constrained weights. The building of such estimators is based on the following steps:

- Identifying the constants q_k and r_k; this provides a QR estimator.
- Resolving the problem (2.3) so as to provide an RQR estimator.

In terms of robustness, the coefficients q_k are selected such that $\hat{B}_q$ is a robust estimator. Thus, the first part of the QR estimator, $T_x'\hat{B}_q$, provides a good predicted value for the entire population. The second part of the QR estimator, $\sum_s r_k e_k$, corrects the first part for the y_k observed in the sample. The constants r_k ensure that with this correction, the outliers in the sample will not return with full weight.

3.2.1 Choice of q_k based on a GM-estimator

Consider the estimator (1.2) in which $\hat{B}_q$ is replaced by a robust estimator of a regression coefficient. Such estimators have been discussed for example in Huber (1981) and Hampel *et al*., (1986). We thus obtain

$$\hat{T}_g = T_x'\hat{B}_g + \sum_s r_k(y_k - x_k'\hat{B}_g). \qquad (3.3)$$

The estimator (3.3) does not have the form of QR estimators unless $\hat{B}_g$ assumes a weighted form. This is the case if $\hat{B}_g$ is a GM-estimator defined by the equation

$$\sum_s d_k h_k x_k \psi((y_k - x_k'B)/(\sigma h_k^{\alpha}\sqrt{c_k}))\sqrt{c_k} = 0, \qquad (3.4)$$

since the solution to (3.4) can be expressed as

$$\hat{B}_g = \left(\sum_s d_k h_k^{1-\alpha} u_k x_k x_k'/c_k\right)^{-1} \sum_s d_k h_k^{1-\alpha} u_k x_k y_k/c_k,$$

where

$$u_k = \frac{\psi((y_k - x_k'\hat{B}_g)/(\sigma h_k^{\alpha}\sqrt{c_k}))}{(y_k - x_k'\hat{B}_g)/(\sigma h_k^{\alpha}\sqrt{c_k})}.$$

The properties of GM-estimators have been discussed in Simpson and Chang (1997). To simplify our discussion, σ is assumed to be known, and the role of c_k is the same as in the case of the GREG estimator. The function ψ is determined by the analyst. A current example would be the Huber function

$$\psi_{Hub}(x;\, c) = \begin{cases} c & \text{if } x > c, \\ x & \text{if } |x| \le c, \\ -c & \text{if } x < -c. \end{cases} \tag{3.5}$$

A value of c around 2 is often used in calculating GM-estimators. See for example Hampel *et al*., (1986), Gwet and Rivest (1992) and Hulliger (1995).

The choice of h_k makes it possible to limit the influence of auxiliary information that is too extreme. The constant $\alpha = 0$ leads to the choice of Mallows whereas $\alpha = 1$ makes it possible to obtain the Schweppe version. The Schweppe version is sometimes preferred. See Coakley and Hettmansperger (1993) and Hampel *et al*., (1986, page 322). When there is minimal auxiliary information, *i.e.*, when we only have available a real variable x_k, $\forall k \in U$, a possible choice for function h_k is

$$h_k = \min\left(1, \frac{t}{x_k/\text{med}(x_k)}\right). \tag{3.6}$$

For a design πps, a modification of h_k following Dalén (1987) so as to take various sampling weights into consideration would perhaps be desirable. The constant t must be specified by the statistician. A value of t around 1.5 is found in the applications. See for example Rivest and Rouillard (1991), who also provide other choices for functions h_k.

Writing $\hat{B}_g$ as a weighted estimator makes it possible to write estimator (3.3) as a QR estimator with

$$(q_k,\, r_k) = (d_k h_k^{1-\alpha} u_k / c_k,\, r_k).$$

The choice of constants r_k is discussed in section 3.2.3.

3.2.2 Choice of q_k based on a high-breaking-point estimator

The choice of a GM-estimator is only a first step towards obtaining a very robust estimator of the total. In fact, although the influence function of GM-estimators is restricted, the fact remains that such estimators do not have a high breakdown point, which usually diminishes according to the dimension of the auxiliary information (Rousseeuw and Leroy 1987, page 13). This section will explain how to build robust calibration estimators based on high breakdown point estimators. As such estimators do not usually assume a weighted form, we will consider reweighting them. This will allow us to obtain, as in the previous section, the constants q_k needed to compute the RQR estimator metric. Specifically, the following weights $\hat{u}_k$ are considered:

$$\hat{u}_k = \frac{\psi\left((y_k - x_k'\hat{B}_0)/(\sigma h_k^{\alpha}\sqrt{c_k})\right)}{(y_k - x_k'\hat{B}_0)/(\sigma h_k^{\alpha}\sqrt{c_k})}, \tag{3.7}$$

where $\hat{B}_0$ is an equivariant estimator with a high breakdown point meeting certain regularity conditions. The reweighted estimator is

$$\hat{B}_r = \left(\sum\nolimits_s d_k h_k^{1-\alpha}\hat{u}_k x_k x_k' / c_k\right)^{-1} \times \sum\nolimits_s d_k h_k^{1-\alpha}\hat{u}_k x_k y_k / c_k. \tag{3.8}$$

The asymptotic properties of this type of estimator have been studied in Simpson and Chang (1997).

The estimator $\hat{B}_0$ that is considered is the one-step GM-estimator of Coakley and Hettmansperger (1993). This estimator has a high breakdown point. It is obtained as the first iteration of the Newton formula in equation (3.4), where the Schweppe version is used, assuming $\alpha = 1$. Other robust estimators could have been chosen. However, the efficiency and robust properties of the Coakley and Hettmansperger (1993) estimator make it a good choice. Thus, the proposed constant q_k is

$$q_k = d_k h_k^{1-\alpha}\hat{u}_k / c_k,$$

with $\hat{B}_0 = \hat{B}_{\text{CH}}$, $\hat{B}_{\text{CH}}$ denoting the Coakley and Hettmansperger (1993) estimator.

3.2.3 Choice of r_k

Once the constants q_k, have been determined, the constants r_k must be selected. If $d_k = r_k$ then under general conditions, the QR estimator is an ADU estimator. However, such a choice of r_k yields an estimator that is sensitive to outliers. Alternately, choosing $r_k = 0$ provides a robust estimator that might be very biased as was emphasized in Gwet and Rivest (1992, page 1180). Lee (1991) suggested choosing $r_k = \theta d_k$, where $\theta \in [0,\, 1]$. The asymptotic bias becomes under general conditions $(\theta - 1)\sum_U E_k$, where E_k represents the residuals obtained by adjusting a robust estimator for the entire population. Choosing θ makes it possible to control estimator bias. The discussion in section 3 leads us to suggest constants r_k that are close to the d_k for good units, and reduced gradually for doubtful observations. We suggest choosing

$$r_k = d_k\, u_k^*, \tag{3.9}$$

where

$$u_k^* = \frac{\psi^*\left((y_k - x_k'\hat{B}_r)/(\sigma h_k^{\alpha}\sqrt{c_k})\right)}{(y_k - x_k'\hat{B}_r)/(\sigma h_k^{\alpha}\sqrt{c_k})}.$$

The function ψ^* which we will be considering is a modification of the Huber function

$$\psi^*(x) = \begin{cases} x & \text{if } |x| \le a, \\ a\ \text{sign}\,(x) & \text{if } |x| > a \text{ and } |x| < a/b, \\ bx & \text{if } |x| > a/b. \end{cases} \tag{3.10}$$

We choose $a = 9$, $b = 1/4$. The reason for this modification is that we do not want the outliers comprising large residuals or extreme auxiliary information to have weights that are too reduced. In this way, the sampling weight is fully maintained when the argument for u_k^* is between -9 and 9, and reduced gradually to one quarter. If the weight of large residuals is reduced too much, then the bias becomes too great, leading to the choice of ψ^*. The choice of constants r_k has been done empirically, and seems to work well in practice.

Thus, we will consider the choice of constants q_k and r_k following

$$(q_k, r_k) = (d_k h_k^{1-\alpha} \hat{u}_k / c_k, d_k u_k^*). \qquad (3.11)$$

We suggest a generalization of the Lee proposition (1991), since instead of considering $r_k = d_k \theta$ where θ is fixed, $r_k = d_k u_k^*$ will adapt automatically (or adaptively) to the sample. Having this choice of constants q_k and r_k at our disposal, and with the usual mean square metric, we obtain a QR estimator, but it is subject to negative weighting problems. However, with the constants (3.11), we can consider the restricted mean square metric, solve the problem (2.3) and obtain a robust estimator meeting the CEs and the CARVs.

There are in proposition 1 possible solutions for the asymptotic behaviour of the resulting RQR estimator as compared to the QR estimator free of weight constraints. However, since the constants (3.11) depend on s in a complex way, there can be no automatic conclusion about asymptotic equivalence. Nevertheless, the simulation study in section 4 seems to suggest a very comparable behaviour for the estimator with and without constraint on the weights, with respect to the Monte Carlo mean square error. Thus, empirical evidence shows that if the q_k and r_k are chosen in such a way that the estimator without constraint on the weights is robust, then the version with constraints on the weights will also be robust.

Finally, the following is a summary of the steps in the proposed method used to obtain a robust RQR estimator.

1. Choice of constants q_k and r_k. We suggest the constants found in equation (3.11). For this step, it is necessary to compute $\hat{B}_{\text{CH}}$.
2. Choice of metric. If need be, choice of constants L and U. These constants are chosen such that $L \le r_k \le U, \ \forall k \in s$.
3. Solution using Newton's method for equation (2.5).
4. Assume $w_k = h(x_k' \lambda_\infty; q_k, r_k)$ for λ_∞ solution to step 3.
5. Assume $\hat{T}_{yr} = \sum_s w_k y_k$, which is the proposed RQR estimator.

The procedure requires a certain number of constants. The constants α, t and c are found in the calculation of q_k and r_k. The choice of these values is nevertheless justified using robustness theory, which helps guide the practitioner. Thus, the value for c in the Huber function can be obtained by taking into account efficiency concerns under normal errors. See Hampel *et al.*, (1986, page 333) and Gwet and Rivest (1992). Constants a and b are also found; they are more directly linked to the proposed estimators. Constant b represents the maximum weight reduction that can be allowed when specifying the default weights r_k, and for this reason there is a link with the suggestion made by Lee (1991). The constant which it is most important to specify is possibly the value of a. We suggest here $a = 9$. However, in our simulations, a value of a between 6 and 12 yielded relatively comparable results. The choice of limits L and U rests on cosmetic considerations, so that the weights may be limited to one interval. This last consideration is perhaps secondary for the practitioner. As a result, it would seem that the most important aspect is to choose a value of r_k that is close to d_k for the proper values, then reduced as an observation is deemed extreme, and that is the goal which has guided our choice of r_k in this section. Nevertheless, it would be useful to make a choice of r_k that satisfies a certain optimality criterion.

3.3 Chambers model-assisted estimator

Another approach is based on a decomposition proposed by Chambers (1982, 1986) which we now apply to QR estimators. Note that a QR estimator can always be written in the form

$$\hat{T}_{y\text{QR}} = \sum_s r_k y_k + (T_x - \hat{T}_{xr})' B + \sum_s z_k \sqrt{q_k}\,(y_k - x_k' B),$$

where $z_k = (T_x - \hat{T}_{xr})' (\sum_s q_k x_k x_k')^{-1} x_k \sqrt{q_k}$, $\hat{T}_{xr} = \sum_s r_k x_k$ and B are arbitrary. Chambers (1986) had considered the specific case $(q_k, r_k) = (1/\sigma_k^2, 1)$ for the ratio estimator. In order to limit the influence of outlier units, Chambers proposed

$$\hat{T}_{y\text{CHAM}} = \sum_s r_k x_k + (T_x - \hat{T}_{xr})' B + \sum_s z_k \psi_p\,(\sqrt{q_k}\,(y_k - x_k' B)). \qquad (3.12)$$

The function ψ_p helps limit the influence of large residuals. The choice for B is a robust estimator, *e.g.*, $\hat{B}_g$. One function ψ_p considered in Chambers (1986) was

$$\psi_p(t) = t \exp(-0.25(|t| - 6)^2). \qquad (3.13)$$

It is interesting to note that (3.12) can be written as

$$\hat{T}_{y\text{CHAM}} = T_x' B + \sum_s (r_k + (d_k g_k - r_k)\,\lambda_k)\, e_k(B),$$

where $e_k(B) = y_k - x_k' B$, g_k is defined in formula (2.1) calculated using q_k and r_k, and

$$\lambda_k = \frac{\psi_p\,(\sqrt{q_k}\,(y_k - x_k' B))}{\sqrt{q_k}\,(y_k - x_k' B)}. \qquad (3.14)$$

Thus, the residuals $e_k(B)$ are weighted using a relation referring to formula (3.2). If $\lambda_k \equiv 1$, then $d_k g_k$ is applied to residuals $e_k(B)$ and it is easy to verify that we have the estimator $\hat{T}_{y\mathrm{QR}}$. Alternately, if $\lambda_k \equiv 0$, we obtain (3.3) if we assume $B = \hat{B}_g$. If in (3.12) we assume $(q_k, r_k) = (1/c_k, 1)$ and $B = \hat{B}_g$, then the Chambers estimator represents a compromise between the BLUP and a robust estimator based on a GM-estimator. Note that formally (3.12) is a QR estimator with

$$(q_k^*, r_k^*) = (d_k h_k^{1-\alpha} u_k / c_k,\ r_k + (d_k g_k - r_k)\lambda_k).$$

However, since r_k^* is not necessarily positive, it is not always possible to undertake a change of metric in this case.

4. Empirical study

To study the performance of robust calibration estimators, we carried out a Monte Carlo simulation study. We considered four populations comprising data from readily available works on sampling theory. For each population, $K = 2{,}000$ samples were drawn using simple random sampling for various sample sizes. Our main objective was to determine whether it is possible to obtain estimators having good empirical properties (bias, mean square error) while satisfying the CEs and the CARVs. Note that all the programs were written in S-PLUS (Statistical Sciences 1991) and are available from the author.

4.1 Populations under study

The population graphs can be found in Figure (4.1). The first population, comprising 51 units, can be found in Mosteller and Tukey (1977, page 560). It consists of the U.S. population in 1960 and in 1970 for each of the 50 states and the federal district of Columbia. It is called POPUSA. Looking at the scattergram of the 1970 population in terms of the 1960 population, we notice that all units seem to be on the same straight line, with some good leverage points. An example of a good leverage point is the point surrounded in this population. The second population, with 34 units, can be found in Singh and Chaudhary (1986, page 177). It deals with the area of fields sown in 1971 and in 1974. This population is called AREA. There is a bad leverage point (see the surrounded point) in this population since the point (4,170.99) does not respect the linear trend

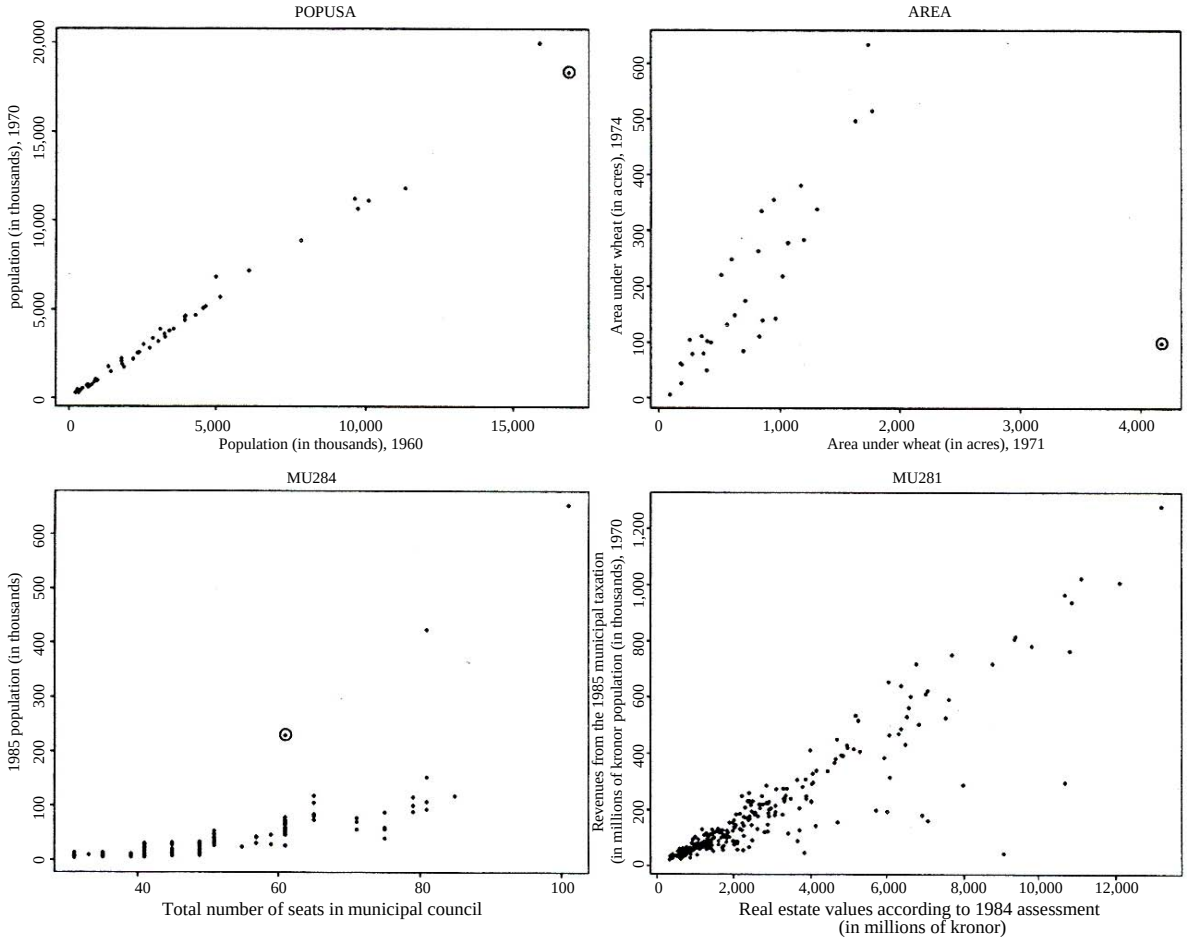

Figure 4.1 The four populations under study

of the majority of units. Samples of size 10 and 15 are drawn from POPUSA and AREA. The third population, *i.e.* the MU284 in Särndal *et al.*, (1992), comprises the 284 municipalities of Sweden. We considered variables $x = \text{S82}$ concerning the total number of seats in the municipal council, and $y = P85$ representing the population of Sweden in 1985. There are vertical outliers (*e.g.*, the surrounded point) and one bad leverage point. Finally, we considered population made up of MU281 made up of MU284 from which the three largest municipalities were excluded. The variables considered were $x = \text{REV84}$ representing the values of landed property based on the 1984 assessment, and $y = \text{RMT85}$ representing municipal tax revenues in 1985. The unit of measurement was one million kronor for both variables. It seems this population has several bad leverage points. Samples of size $n = 30$ and $n = 60$ were drawn from MU284 and MU281. Table (4.1) contains totals for various populations.

Table 4.1
Totals for various populations and totals known from auxiliary information

Population	T_x	T_y	N
POPUSA	179,972	203,923	51
AREA	29,118	6,781	34
MU284	13,500	8,339	284
MU281	757,246	53,124	281

4.2 Description of the estimators

The two basic estimators were the GREG estimator and the estimator obtained by considering case No. 7 in Deville and Särndal (1992), *i.e.*, a GREG estimator with restricted weights. These estimators were denoted GREG/U and GREG/R respectively. We selected $c_k \equiv 1$ for populations POPUSA and AREA, and chose $c_k = x_k$ for populations MU284 and MU281. Our choice for the c_k was motivated by the relationship between these constants and the heteroscedasticity of the superpopulation model. Of the robust estimators, we studied the Chambers (1986) estimator by considering

$$(q_k^*,\, r_k^*) = (d_k \hat{u}_k(\hat{B}_{\text{CH}})/c_k,\, 1 + (d_k g_k - 1)\lambda_k(\hat{B}_r)),$$

where in the formula (2.1) $(q_k,\, r_k) = (1/c_k,\, 1)$, denoted CHAM, based on $\hat{B}_r$. The constants $\hat{u}_k(\hat{B}_{\text{CH}})$ were obtained from formula (3.7). Selection $\alpha = 1$ was used throughout the simulation. Huber's function ψ was used with the constant $c = 1.345$ for $\hat{B}_r$. The functions h_k are those given by formula (3.6), where we selected $t = 1.46$. The function λ_k is defined by equation (3.14). The function ψ_p considered was that given by equation (3.13). The scale was estimated as in Coakley and Hettmansperger (1993). We also considered the model-assisted BLUP estimator in which the generalized least squares estimator was replaced by estimator $\hat{B}_r$, which we called MODEL. Moreover, we considered the Lee (1991) estimator on the basis of $\hat{B}_r$ where $r_k = 0,25\, d_k$, using the mean square metric. We also studied an extension of the Lee (1991) estimator by considering the limited mean square metric. These estimators were denoted by LEE25/U and LEE25/R respectively. Finally, we considered the new method in section 3.2.3, selecting $(q_k,\, r_k)$ as given by equation (3.11) in accordance with the mean square metric and the limited mean square metric. They were denoted by QRROB/U and QRROB/R respectively. The choice of function ψ^* was given by formula (3.10).

Table 4.2
Monte Carlo results for sampling from the POPUSA population

Estimators		VAR_{M}	MSE_{M}	CV_{M}	BR_{M}	MIN	MAX	CARV [1]	CONV
$n = 10$									
	GREG/U	34.90	34.92	2.90	-0.07	-6.24	26.75	86.7	
	GREG/R	35.29	35.30	2.91	-0.04	0.20	32.00	100.0	98.4
	CHAM	32.43	33.75	2.85	-0.56	-19.61	40.96	84.0	
	MODEL	27.66	30.69	2.72	-0.85	-19.71	40.86	82.8	
	LEE25/U	27.48	30.07	2.69	-0.79	-19.38	39.64	83.2	
	LEE25/R	28.67	30.90	2.73	-0.73	0.20	32.00	100.0	98.4
	QRROB/U	27.40	28.40	2.61	-0.49	-15.68	40.10	83.2	
	QRROB/R	28.33	29.18	2.65	-0.45	0.20	32.00	100.0	98.4
$n = 15$									
	GREG/U	21.90	21.95	2.30	-0.10	-3.13	15.32	94.7	
	GREG/R	22.12	22.15	2.31	-0.09	0.20	16.00	100.0	99.5
	CHAM	18.11	20.14	2.20	-0.70	-5.79	16.44	92.4	
	MODEL	15.43	19.03	2.14	-0.93	-6.09	16.92	91.0	
	LEE25/U	15.44	19.54	2.17	-0.99	-6.19	17.06	90.8	
	LEE25/R	15.72	19.68	2.18	-0.98	0.20	16.00	100.0	99.5
	QRROB/U	14.68	16.44	1.99	-0.65	-4.48	16.41	90.9	
	QRROB/R	14.85	16.56	2.00	-0.64	0.20	16.00	100.0	99.5

[1] The limits for the CARVs are [0.20, 32] for $n = 10$ and [0.20, 16] for $n = 15$.

4.3 Frequency measurements

The eight estimators in Section (4.2) were calculated for each sample. The results can be found in Tables 4.2, 4.3, 4.4 and 4.5. Since one asset of the new methods is the CARVs, statistics were calculated on these weights. Columns MIN and MAX in the tables of results contain the minimum and maximum values of weights calculated during the simulation for each estimator. Also shown is the percentage of samples for which the weights are within the CARVs in the CARV column in the tables of results. We also considered the percentage of samples for which there were convergent limited estimators in the CONV column. The intervals used $[L, U]$ for the limited intervals are specified in the different tables. In all cases, the various statistics were calculated using samples for which all estimators were convergent.

Another significant feature is related to the bias and efficiency of the proposed methods. Let $\hat{T}$ denote an estimator of the total T_y. Assume $\hat{T}_i$ is the estimator of the total calculated using sample i, $i = 1, ..., K$. The relative Monte Carlo bias BR_M, the mean value E_M and the variance V_M are given by the usual formulas, *i.e.*,

$$\mathrm{BR}_M = (E_M(\hat{T}) - T_y)/T_y \times 100,$$

$$E_M(\hat{T}) = \frac{1}{K}\sum_{i=1}^{K}\hat{T}_i \text{ and } V_M = \frac{1}{K}\sum_{i=1}^{K}(\hat{T}_i - E_M(\hat{T}))^2.$$

Our main criterion for efficiency will be the Monte Carlo mean square error defined by $\mathrm{MSE}_M = 1/K \sum_{i=1}^{K}(\hat{T}_i - T_y)^2$. The coefficients of variation CV_M are calculated in accordance with $\sqrt{\mathrm{MSE}_M}/T_y$. The variance and mean square error are expressed in millions. The coefficient of variation, the relative bias, the CARVs and the convergence of limited versions are expressed as percentages.

4.4 Discussion

The POPUSA population had no outliers that did not satisfy the linear model. During sampling, the coefficients of variation of the estimators were small, which could be expected given the trend of the population. Columns MSE and VAR are very similar, indicating that bias is not a problem for this population. All relative bias was less than 1%. The QRROB/U estimator provided a reduction in variance as compared to GREG/U that exceeded 21% for $n = 10$ and 30% for $n = 15$.

The size of the AREA population was small. This population had a bad leverage point leading to very high empirical relative bias for all the estimators. The GREG/U estimator had a relative bias of more than 7% in spite of a 44% sampling for this population. The robust estimators had the most significant bias, though it was relatively comparable to the bias of the GREG/U estimator. The most significant reduction in variance was achieved for the QRROB/U estimator, but at the cost of a relative bias of about 10%.

Population MU284 had a vertical outlier and bad leverage points. Robust estimators reduced the variance radically, since they were not affected by the three extreme units in y which were clearly moving away from the linear trend. The CHAM, QRROB/R and QRROB/U estimators were more than four times less variable than the GREG/U estimator. However, this led to a much higher negative bias. All the robust estimators were severely biased. The MODEL estimator showed a negative bias of more than 13%, whereas QRROB/U had a negative bias of the order of 11%. As for QRROB, a better choice of constants in function ψ^* might help reduce a larger part of the bias at the cost of a lower variance reduction. Increasing the sample size to $n = 60$ made it possible to reduce the bias below the 10% of the CHAM and QRROB estimators, but the other robust estimators remained more biased.

Population MU281 contained a fairly large number of bad leverage points. The variance dominated the MSE share of this population. The LEE25 estimator was the least variable, with a reduction of more than 35% as compared to GREG/U. However, although $\theta = 0.25$ functions well for this population, our study shows that it is not always the best choice.

Note that all the robust estimators were more efficient than the GREG or its limited version. As was confirmed by the results of Deville and Särndal (1992), the limited version of the GREG estimator showed essentially the same behaviour as the GREG in terms of both bias and Monte Carlo variance for each population. Of all the estimators that were considered, GREG/U and GREG/R were the least biased. The robust versions all exhibited greater bias. However, this is more than offset by the reduction in variance so that the efficiency of robust estimators is always greater than that of GREG/U or of GREG/R estimators.

Concerning the constraints on the weights, it will be noted that the GREG/U, CHAM, MODEL, LEE25 and QRROB/U estimators are all subject to problems of negative weighting, as can be seen in column MIN. This problem is avoided with limited estimators. The CARV column shows that the constraints were not met relatively frequently, depending on population and sample size, varying between 5% and 60%. The general behaviour of the two limited robust estimators was comparable to that of their non-limited versions. Moreover, QRROB/R, in addition to meeting the CARVs, provided interesting properties of efficiency, as compared to other robust estimators. Limited versions were not as prone to convergence problems when sample sizes were greater. Note that we had to use wider bands in the case of POPUSA in order to obtain satisfactory convergence rates.

Table 4.3
Monte Carlo results for sampling from the area population

Estimators	VAR_M	MSE_M	CV_M	BR_M	MIN	MAX	CARV [1]	CONV
$n = 10$								
GREG/U	1.334	1.700	19.23	8.92	-3.35	14.94	86.6	
GREG/R	1.295	1.629	18.82	8.53	0.20	14.00	100.0	99.0
CHAM	1.187	1.541	18.30	8.77	-4.09	14.90	87.2	
MODEL	1.291	1.580	18.54	7.93	-5.23	16.75	86.8	
LEE25/U	1.279	1.593	18.61	8.26	-5.28	16.89	86.6	
LEE25/R	1.284	1.596	18.63	8.24	0.20	14.00	100.0	99.0
QRROB/U	1.026	1.440	17.70	9.50	-4.74	15.38	87.6	
QRROB/R	1.028	1.437	17.68	9.43	0.20	14.00	100.0	99.0
$n = 15$								
GREG/U	0.940	1.178	16.00	7.18	-1.40	7.03	93.0	
GREG/R	0.928	1.154	15.85	7.01	0.20	6.00	100.0	99.8
CHAM	0.708	0.989	14.67	7.82	-1.52	7.92	93.7	
MODEL	0.757	0.997	14.73	7.22	-1.66	8.39	93.1	
LEE25/U	0.672	1.059	15.18	9.18	-1.68	9.40	92.0	
LEE25/R	0.671	1.056	15.15	9.15	0.20	6.00	100.0	99.8
QRROB/U	0.485	0.990	14.68	10.48	-1.59	8.90	93.9	
QRROB/R	0.485	0.986	14.64	10.44	0.20	6.00	100.0	99.8

[1] The limits for the CARVs are [0.20, 14] for $n = 10$ and [0.20, 6] for $n = 15$.

Table 4.4
Monte Carlo results for sampling from the MU284 population

Estimators	VAR_M	MSE_M	CV_M	BR_M	MIN	MAX	CARV [1]	CONV
$n = 30$								
GREG/U	2.833	2.925	20.51	-3.64	-6.83	23.90	89.8	
GREG/R	2.813	2.910	20.46	-3.73	0.20	16.00	100.0	99.2
CHAM	0.645	1.639	15.35	-11.95	-11.80	31.26	77.0	
MODEL	0.709	2.037	17.11	-13.82	-12.06	31.91	68.40	
LEE25/U	0.887	1.877	16.43	-11.93	-11.06	30.93	73.5	
LEE25/R	0.871	1.847	16.30	-11.85	0.20	26.00	100.0	99.2
QRROB/U	0.719	1.532	14.84	-10.81	-9.46	25.84	86.5	
QRROB/R	0.720	1.525	14.81	-10.76	0.20	16.00	100.0	99.2
$n = 60$								
GREG/U	1.473	1.489	14.63	-1.49	-1.19	10.03	90.1	
GREG/R	1.467	1.484	14.61	-1.57	0.20	7.00	100.0	99.7
CHAM	0.357	0.990	11.93	-9.54	-2.53	15.59	69.8	
MODEL	0.380	1.255	13.43	-11.22	-4.93	14.52	58.1	
LEE25/U	0.403	1.201	13.14	-10.72	-4.80	14.20	60.3	
LEE25/R	0.396	1.203	13.16	-10.78	0.20	7.00	100.0	99.7
QRROB/U	0.308	0.976	11.85	-9.80	-2.36	10.99	86.1	
QRROB/R	0.308	0.979	11.87	-9.82	0.20	7.00	100.0	99.7

[1] The limits for the CARVs are [0.20, 16] for $n = 30$ and [0.20, 7] for $n = 60$.

Table 4.5
Monte Carlo results for sampling from the MU281 population

Estimators	VAR_M	MSE_M	CV_M	BR_M	MIN	MAX	CARV [1]	CONV
$n = 30$								
GREG/U	17.33	17.35	7.84	-0.26	-38.97	34.56	86.0	
GREG/R	17.40	17.41	7.86	-0.24	0.20	25.00	100.0	99.8
CHAM	13.23	13.26	6.86	-0.33	-47.09	39.08	56.9	
MODEL	11.30	11.91	6.50	1.47	-66.22	41.43	47.9	
LEE25/U	11.21	11.60	6.41	1.17	-59.75	37.03	53.3	
LEE25/R	11.26	11.73	6.45	1.29	0.20	25.00	100.0	99.8
QRROB/U	12.92	13.29	6.86	1.15	-54.14	39.73	70.8	
QRROB/R	12.94	13.34	6.88	1.20	0.20	25.00	100.0	99.8
$n = 60$								
GREG/U	7.57	7.57	5.18	-0.10	-12.77	15.34	86.4	
GREG/R	7.58	7.58	5.18	-0.09	0.20	9.00	100.0	99.9
CHAM	5.85	5.90	4.57	-0.43	-22.97	11.49	51.4	
MODEL	4.53	5.23	4.30	1.57	-24.02	14.58	38.7	
LEE25/U	4.55	5.18	4.28	1.49	-23.74	14.41	41.2	
LEE25/R	4.50	5.21	4.30	1.58	0.20	9.00	100.0	99.9
QRROB/U	5.40	6.16	4.67	1.64	-21.08	21.07	68.6	
QRROB/R	5.39	6.17	4.67	1.66	0.20	9.00	100.0	99.9

[1] The limits for the CARVs are [0.20, 25] for $n = 30$ and [0.20, 9] for $n = 60$.

5. Conclusion

The goal of this paper has been to introduce calibration estimators having good properties of robustness. Traditional calibration estimators are easy to use, since it is sufficient to have a set of starting weights, usually the sampling weights d_k, which are transformed into calibrated weights. The steps used in this paper have been the same, *i.e.*, the robust default weights r_k have been transformed into calibrated weights, and the constants q_k have been chosen such that $\hat{B}_q$ is a robust estimator. The proposed choice of r_k is given by the formula (3.9), with $a = 9$, $b = 1/4$. There remains to develop a theory for the optimal choice of r_k. The suggestion is made for applications to vary constant a, between say 6 and 12, in order to determine the influence of the constant on the estimation. The limits L and U can be used to limit the weights, *e.g.*, to make them all positive. We suggest the general use of $L = 0.2$, $U = kN/n$, where k is about 3.

Note that robust calibration estimators are not meant to replace the GREG estimator, but to be used in conjunction with it. Thus, if the robust estimator and the GREG estimator are very different, a more in-depth analysis might help determine the reason. The proposed estimators could be useful as diagnostic tools.

It would be interesting to pursue the empirical studies of section 4, by examining for example the effect of sampling design on the proposed procedures. Another important area of development is the estimation of variance. Multipurpose surveys are yet another area of interest. In fact, for applications, there is rarely a single variable of interest, and methodologists would like to use a single set of weights for all the variables of interest. In terms of robustness, a solution has been proposed in the conclusion of a paper by Gwet and Rivest (1992), where robust weights were calculated for each variable of interest $y^{(i)}$, $i = 1, ..., I$. For one unit, the final weight corresponds to the minimum weight among the weights obtained. Alternately, to obtain robust and calibrated estimators, we could calculate robust default weights for each variable of interest, providing a set of $r_k(y^{(i)})$, and assume $r_k = \min r_k(y^{(i)})$, where the minimum is on $i = 1, ..., I$. These weights could then be transformed into calibrated weights. This procedure should be assessed in greater detail.

Acknowledgements

I wish to thank Carl-Erik Särndal for introducing me to sampling theory and for suggesting that I consider the problem of outliers in sampling theory. I also thank Roch Roy and Christian Léger for helping me during various stages of development of this paper. My sincere thanks go to the Associate Editor, the Assistant Editor and two referees for comments which led to significant improvements in both content and layout.

Appendix A

Proof of proposition 1

Let $\Delta_h(u; q, r) = r + qu - h(u; q, r)$ and z_k be a variable of interest. We assume the following conditions

$$C_1. N^{-1}\sum_s q_k z_k = O_p(1).$$

$$C_2. N^{-1}\sum_s \Delta_h(x_k'\lambda_\infty; q_k, r_k) z_k = o_p(n^{-1/2}),$$

where λ_∞ is a solution to equation (2.5).

Note that $\sum_s (r_k + q_k x_k' \lambda_1) x_k = T_x$, where $\lambda_1 = -(\sum_s q_k x_k x_k')^{-1}(\hat{T}_{xr} - T_x)$, and also that $\sum_s h(x_k'\lambda_\infty; q_k, r_k) x_k = T_x$. Thus, in using C_2, we find that

$$N^{-1}\sum_s q_k x_k x_k' (\lambda_1 = \lambda_\infty) = o_p(n^{-1/2}),$$

$\lambda_1 - \lambda_\infty = o_p(n^{-1/2})$, and therefore using C_1, with $z_k = x_k x_k'$. It is also easily shown that

$$\begin{aligned} &N^{-1}(\hat{T}_{\mathrm{QR}} - \hat{T}_{\mathrm{RQR}}) \\ &= N^{-1}\sum_s (r_k + q_k x_k'\lambda_1) y_k - N^{-1}\sum_s h(x_k'\lambda_\infty; q_k, r_k) y_k \\ &= N^{-1}\sum_s q_k x_k' y_k (\lambda_1 - \lambda_\infty) + N^{-1}\sum_s \Delta_h(x_k'\lambda_\infty; q_k, r_k) y_k \\ &= o_p(n^{-1/2}). \end{aligned}$$

Appendix B

List of abbreviations

ADU: Asymptotically Design Unbiased.

BLUP: Best Linear Unbiased Predictor (Royall 1970).

CARV: Constraints applicable to the range of values for the weights w_k, by requiring for example that all the $w_k \in [L, U]$.

CE: Calibration constraints, $\sum_s w_k x_k = T_x$, where $T_x = \sum_U x_k$.

CH: Robust estimator proposed by Coakley and Hettmansperger (1993), a single-step GM-estimator that is robust and efficient.

CHAM: Robust Chambers (1982, 1986) estimator.

GM: Generalized M estimators, derived from robustness theory (see for example Hampel *et al.*, 1986).

GREG: Generalized regression estimator proposed by Cassel *et al.*, (1976).

HT: Horvitz-Thompson estimator $\sum_s d_k y_k$, where $d_k = \pi_k^{-1}$.

QR: Wright (1983) estimators, in the form $T_x'\hat{B}_q + \sum_s r_k e_k$.

RQR: Generalization of the Wright (1983) estimators, obtained using a general metric as well as constraints on the weights.

Appendix C

List of the principal constants

c_k: Factor capable of accounting for heteroscedasticity problems.

d_k: Sampling weights.

g_k: g-weight defined by w_k/d_k.

h_k: Quantity used to reduce the influence of outlier auxiliary information in $\hat{B}_q$.

π_k, π_{kl}: Inclusion probabilities of first and second order, respectively.

q_k, r_k: Quantities defining an estimator QR. The q_k are used to build the regression coefficients involved in the first part, $T_x'\hat{B}_q$ the r_k are used for the second part, $\sum_s r_k e_k$.

u_k, $\hat{u}_k$: Weights used to build $\hat{B}_q$ in a robust way.

u_k^*: Weights used to consider a robust correction factor $\sum_s r_k e_k$.

w_k: Calibrated weight attributed to y_k to form $\sum_s w_k y_k$.

References

Bershad, M.A. (1960). Some Observations on Outliers. Unpublished memorandum, Statistical Research Division, U.S. Bureau of the Census.

Bolfarine, H., and Zacks, S. (1992). *Prediction Theory for Finite Population*. New York: Springer-Verlag.

Brewer, K.R.W. (1994). Survey sampling inference: Some past perspectives and present prospects. *Pakistan Journal of Statistics*, 10, 213-233.

Bruce, A.G. (1991). Robust Estimation and Diagnostics for Repeated Sample Surveys. Mathematical Statistics Working Paper 1991/1, Statistics New Zealand.

Cassel, C.M., Särndal, C.-E. and Wretman, J.H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite population. *Biometrika*, 63, 615-620.

Chambers, R.L. (1982). Robust Finite Population Estimation. Ph. D. thesis, Johns Hopkins University, Dept. of Biostatistics.

Chambers, R.L. (1986). Outlier robust finite population estimation. *Journal of the American Statistical Association*, 81, 1063-1069.

Chambers, R.L., and Kokic, P.N. (1993). An integrated approach for the treatment of outliers in sub-annual surveys. *Proceedings on the 49th Session, International Statistical Institute*.

Coakley, C.W., and Hettmansperger, T.P. (1993). A bounded influence, high breakdown, efficient regression estimator. *Journal of the American Statistical Association*, 88, 872-880.

Dalén, J. (1987). Practical Estimators of a Population Total Which Reduce the Impact of Large Observations. R & D Report, Statistics Sweden.

Deville, J.-C., and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.

Donoho, D.L., and Huber, P.J. (1983). The notion of breakdown point. In *A Festschrift for Erich Lehmann*, (Eds. P.J. Bickel, K.A. Doksum and J.L. Hodges). Belmont, CA: Wadsworth.

Fuller, W.A., Loughin, M.M. and Baker, H.D. (1994). Regression weighting in the presence of nonresponse with application to the 1987-1988 Nationwide Food Consumption Survey. *Survey Methodology*, 20, 75-85.

Gambino, J. (1987). Dealing With Outliers: A Look at Some Methods Used at Statistics Canada. Technical report, Business Survey Division, Statistics Canada.

Gross, W.F., Bode, G., Taylor, J.M. and Lloyd-Smith, C.W. (1986). Some finite population estimators which reduce the contribution of outliers. *Proceedings of the Pacific Statistical Congress. Aucklaud*, New Zealand, 20-24 May 1985.

Gwet, J.-P., and Rivest, L.-P. (1992). Outlier resistant alternatives to the ratio estimator. *Journal of the American Statistical Association*, 87, 1174-1182.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics*: *The Approach Based on Influence Functions*. New York: John Wiley & Sons, Inc.

Hidiroglou, M.A., and Srinath, K.P. (1981). Some estimators of the population total from simple random samples containing large units. *Journal of the American Statistical Association*, 76, 690-695.

Huber, P.J. (1981). *Robust Statistics*. New York: John Wiley & Sons, Inc.

Hulliger, B. (1995). Outlier robust Horvitz-Thompson estimators. *Survey Methodology*, 21, 79-87.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons, Inc.

Lee, H. (1991). Model-based estimators that are robust to outliers. *Proceedings of the 1991 Annual Research Conference*. U.S. Bureau of Census.

Lee, H. (1995). Outliers in business surveys. In *Business Survey Methods*, (Eds. B.G. Cox, D.A. Binder, B.N. Chinnappa, A. Christianson, M.J. Colledge and P.S. Kott). New York: John Wiley & Sons, Inc.

Lee, H., Ghangurde, P.D., Mach, L. and Yung, W. (1992). Outliers in Sample Surveys. Methodology Branch Working Paper BSMD-92-008E, Statistics Canada.

Mosteller, F., and Tukey, J.W. (1977). *Data Analysis and Regression, A Second Course in Statistics*. Redding, MA: Addison-Wesley.

Rivest, L.P., and Rouillard, E. (1991). M-estimators and outlier resistant alternatives to the ratio estimator. *Proceedings*: *Symposium 90, Measurement and Improvement of Data Quality*, Statistics Canada, 245-257.

Rousseeuw, P.J., and Leroy, A.M. (1987). *Robust Regression and Outlier Detection*. New York: John Wiley & Sons, Inc.

Royall, R.M. (1970). On finite population sampling under certain linear regression models. *Biometrika*, 57, 377-387.

Särndal, C.-E. (1996). Efficient estimators with simple variance in unequal probability sampling. *Journal of the American Statistical Association*, 91, 1289-1300.

Särndal, C.-E., Swensson, B. and Wretman, J.H. (1989). The weighted residual technique for estimating the variance of the general regression estimator of the finite population total. *Biometrika*, 76, 527-537.

Särndal, C.-E., Swensson, B. and Wretman, J.H. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Simpson, D.G., and Chang, Y.-C.I. (1997). Reweighted approximate GM-estimators: Asymptotics and residual-based graphics. *Journal of Statistical Planning and Inference*, 57, 273-293.

Singh, A.C., and Mohl, C.A. (1996). Understanding calibration estimators in survey sampling. *Survey Methodology*, 22, 107-115.

Singh, D., and Chaudhary, F.S. (1986). *Theory and Analysis of Sample Survey Designs*. New York: John Wiley & Sons, Inc.

Statistical Sciences, INC. (1991). *S-PLUS Reference Manual*. Seattle: Statistical Science, Inc.

Stukel, D.M., Hidiroglou, M.A. and Särndal, C.-E. (1996). Variance estimation for calibration estimators: A comparison of jackknifing versus Taylor linearizarion. *Survey Methodology*, 22, 117-125.

Tambay, J.-L. (1988). An integrated approach for the treatment of outliers in sub-annual surveys. *Proceedings on the Section on Survey Research Methods*, American Statistical Association, 229-234.

Welsh, A.H., and Ronchetti, E. (1998). Bias-calibrated estimation from sample surveys containing outliers. *Journal of the Royal Statistical Society*, Series B, 60, 413-428.

Wright, R.L. (1983). Finite population sampling with multivariate auxiliary information. *Journal of the American Statistical Association*, 78, 879-884.