The Annals of Statistics
1992, Vol. 20, No. 3, 1571–1593

ASYMPTOTIC PROPERTIES OF THE BALANCED REPEATED REPLICATION METHOD FOR SAMPLE QUANTILES[1]

BY JUN SHAO AND C. F. J. WU

University of Ottawa and University of Waterloo

Inference, including variance estimation, can be made from stratified samples by selecting a balanced set of subsamples. This balanced subsampling method is generically called the balanced repeated replication method in survey data analysis, which includes McCarthy's balanced half-samples method and its extensions for more general stratified designs. We establish the asymptotic consistency of the balanced repeated replication variance estimators when the parameter of interest is the population quantile. The consistency results also hold when balanced subsampling is replaced by random subsampling. As a key technical prerequisite, we prove a Bahadur-type representation for sample quantiles in stratified random sampling.

1. Introduction. An important problem in the analysis of survey data is the estimation of variance of nonlinear statistics from complex survey data. Traditional methods as reviewed in Wolter (1985) include the linearization (Taylor) method, the jackknife, the method of random groups and the balanced half-samples method. More recently bootstrap-like methods have been proposed and studied by Rao and Wu (1988) and Sitter (1992a, b). Gurney and Jewett (1975) extended McCarthy's (1969) balanced half-samples method from two primary sampling units (psu's) per stratum to p psu's per stratum, p being a prime.

Most theoretical results on the asymptotic behavior of these methods are for estimators that are smooth functions of a vector of means, for example, Krewski and Rao (1981), Dippo (1981) and Rao and Wu (1985). When the parameter of interest is the quantile of the finite population, not much is known about the theoretical behavior of various methods for estimating the variance of the quantile estimator. First we review these methods. An obvious method is to substitute the unknown quantities in the variance formula (2.8) by their sample analogs. Its consistency can be established as in Francisco and Fuller (1991) by using a pointwise consistent estimator of the density function at the quantile θ_k. The method is, however, not practical for most survey data because it requires a concentration of observations around θ_k in order for the density at θ_k to be estimated with some precision. To circumvent this problem, there are two realistic alternatives. One is to use Woodruff's (1952) confidence intervals for quantiles to derive new variance estimators [see Francisco and

Received April 1990; revised August 1991.
[1]Supported by Natural Sciences and Engineering Research Council of Canada.
AMS 1980 *subject classifications*. Primary 62D05; secondary 62G05, 62G99.
Key words and phrases. Bahadur representation, balanced half-samples, balanced subsampling, random subsampling, repeated random-group, stratified samples.

Fuller (1991) and Rao and Wu (1987)]. Francisco and Fuller proved the asymptotic properties of these variance estimators. Woodruff's ingenious method of inverting the sample distribution function makes it unnecessary to estimate the density at θ_k. Another alternative for variance estimation is the balanced half-samples method for two psu's per stratum and its extensions to general stratified designs, but little is known about their asymptotic properties. The simulation study of Kovar, Rao and Wu (1988) does not provide any conclusive evidence on the relative events of these two estimators.

The main purpose of this paper is to prove the asymptotic consistency of a general class of variance estimators based on a balanced selection of subsamples, which include as special cases the methods by McCarthy, and Gurney and Jewett. This general method for selecting a balanced set of subsamples is generically called the *balanced repeated replication* method.

First we discuss the framework for our asymptotics. We assume a sequence of superpopulations indexed by k, from which a sequence of finite populations are drawn at random. Formally, let L_k be the number of strata of the kth finite population, $L_k \geq L_{k-1}$, N_{kh} be the size of the hth stratum of the kth population and

$$N_k = \sum_{h=1}^{L_k} N_{kh}$$

and

$$W_{kh} = \frac{N_{kh}}{N_k}.$$

Without loss of generality we assume $N_k \geq ck$ for some c. We do not make any explicit assumption on L_k; that is, L_k can be bounded or goes to infinity. Let $\{H_{kh},\ h = 1, \dots, L_k,\ k = 1, 2, \dots\}$ be an array of distribution functions. Define the kth superpopulation to be

$$F_k = \sum_{h=1}^{L_k} W_{kh} H_{kh}.$$

The kth finite population is $\{X_{h1}, \dots, X_{hN_{kh}},\ h = 1, \dots, L_k\}$, where $X_{h1}, \dots, X_{hN_{kh}}$ are drawn randomly with replacement from H_{kh} and independently across the strata. Note that a subscript k should be included in X_{hi} but is omitted for simplicity. Denote a particular array $\{X_{hi},\ i = 1, \dots, N_{kh},\ h = 1, \dots, L_k,\ k = 1, 2, \dots\}$ by X. For each fixed h, let $\{y_{h1}, \dots, y_{hn_{kh}}\}$ be a simple random sample *without* replacement from the finite population $\{X_{h1}, \dots, X_{hN_{kh}}\}$. Assume that $n_{kh} \geq 2$ for all k and h, but are otherwise arbitrary. Denote the whole sample $\{y_{hi},\ i = 1, \dots, n_{kh},\ h = 1, \dots, L_k\}$ by y. Here $n_k = \sum_h n_{kh}$ is the sample size and $f_k = n_k/N_k$ is the sampling fraction.

Let $0 < p < 1$ be a given constant. The p-quantile of the kth superpopulation is

$$\theta_k = F_k^{-1}(p),$$

where $F^{-1}(t) = \inf\{x: F(x) \geq t\}$ for any distribution function F. Let I_A be the indicator function of the set A,

$$\tilde{F}_{kh}(x) = \frac{1}{N_{kh}} \sum_{i=1}^{N_{kh}} I_{(X_{hi} \leq x)}$$

and

$$\tilde{F}_k(x) = \sum_{h=1}^{L_k} W_{kh} \tilde{F}_{kh}(x) = \frac{1}{N_k} \sum_{h=1}^{L_k} \sum_{i=1}^{N_{kh}} I_{(X_{hi} \leq x)}.$$

The p-quantile of the kth finite population is

$$\tilde{\theta}_k = \tilde{F}_k^{-1}(p).$$

The estimators of $\tilde{F}_k(x)$ and $\tilde{\theta}_k$ (or θ_k) are, respectively,

$$\hat{F}_k(x) = \sum_{h=1}^{L_k} \frac{W_{kh}}{n_{kh}} \sum_{i=1}^{n_{kh}} I_{(y_{hi} \leq x)}$$

and

$$\hat{\theta}_k = \hat{F}_k^{-1}(p).$$

We next consider the resampling plan. Let $\mathbf{s}_h$ be a subset of $\{1, \dots, n_{kh}\}$ with size r_{kh}, where r_{kh} is smaller than n_{kh}, $\mathbf{s} = (\mathbf{s}_1, \dots, \mathbf{s}_{L_k})$, $\mathbf{S}_k$ be the collection of all $\mathbf{s}$ for fixed r_{kh}, $h = 1, \dots, n_{kh}$, $k = 1, 2, \dots$, $d_{kh} = n_{kh} - r_{kh}$, $d_k = \sum_h d_{kh}$ and $r_k = \sum_h r_{kh}$. Note that r_k (and resp. d_k) is the number of units retained (and resp. deleted) in the subsample $\mathbf{s}$. For a particular choice of $\{r_{kh}, h = 1, \dots, L_k, k = 1, 2, \dots\}$, a collection of sets $\mathbf{T}_k \subset \mathbf{S}_k$ is called *balanced* if

$$\pi_{h,ij} = \#\{\mathbf{s} \in \mathbf{T}_k: i \in \mathbf{s}_h, j \in \mathbf{s}_h\} = \text{constant}, \qquad i \neq j, \tag{1.1}$$

$$\pi_{hh'ij} = \#\{\mathbf{s} \in \mathbf{T}_k: i \in \mathbf{s}_h, j \in \mathbf{s}_{h'}\} = \text{constant}, \tag{1.2}$$

where the constant in (1.1) [and resp. (1.2)] is independent of i and j for each h and k (and resp. h, h' and k). It is easy to see that $\mathbf{S}_k$ is balanced. Any plan $\mathbf{T}_k$ satisfying (1.1) and (1.2) is called a *balanced subsampling*. It includes McCarthy's (1969) balanced half-samples and its extensions by Gurney and Jewett (1975). Further discussion is given in Section 4.

For each $\mathbf{s}$, let

$$\hat{F}_k^{\mathbf{s}}(x) = \sum_h \frac{W_{kh}}{r_{kh}} \sum_{i \in \mathbf{s}_k} I_{(y_{hi} \leq x)}$$

and

$$\hat{\theta}_k^{\mathbf{s}} = \left(\hat{F}_k^{\mathbf{s}}\right)^{-1}(p).$$

For a balanced $\mathbf{T}_k$, we estimate the variance of $n^{1/2}(\hat{\theta}_k - \theta_k)$ by the *balanced* *repeated replication* (BRR) *variance estimators*

$$v_R(\mathbf{T}_k) = \frac{n_k}{\lambda_k m_k} \sum_{\mathbf{s} \in \mathbf{T}_k} \left(\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k\right)^2 \tag{1.3}$$

or

$$\tilde{v}_R(\mathbf{T}_k) = \frac{n_k}{\lambda_k m_k} \sum_{\mathbf{s} \in \mathbf{T}_k} \left(\hat{\theta}_k^{\mathbf{s}} - \frac{1}{m_k} \sum_{\mathbf{s} \in \mathbf{T}_k} \hat{\theta}_k^{\mathbf{s}}\right)^2, \tag{1.4}$$

where m_k is the number of subsets in $\mathbf{T}_k$ and λ_k depends on r_k and the choice of $\mathbf{T}_k$. Determination of λ_k will be discussed in Section 4. To prove their asymptotic consistency, we first establish a Bahadur-type representation for $\hat{\theta}_k$ (Theorem 2.1 in Section 2), and the asymptotic normality of $\hat{\theta}_k$ (Theorem 2.2 in Section 2). We then give a representation of v_R and $\tilde{v}_R$ (Theorem 3.1 in Section 3). As applications of Theorem 3.1, we prove in Section 4 the consistency of v_R and $\tilde{v}_R$ for any balanced $\mathbf{T}_k$ with appropriate conditions on the resample size. Since the complete enumeration of $\mathbf{T}_k$ can be prohibitively large, we can use random subsampling to reduce the computations. Consistency of variance estimators based on random subsampling is proved in Section 5.

It should be noted that, when $\{y_{h1}, \dots, y_{hn_{kh}}\}$ is a simple random sample *with* replacement from $\{X_{h1}, \dots, X_{hN_{kh}}\}$, we have results parallel to those for without replacement sampling. Their difference will be highlighted throughout the paper.

Note that the jackknife method also satisfies the balance conditions (1.1) and (1.2). Unlike the BRR, it does not delete unit (or units) simultaneously from *every* stratum. It is known that the jackknife variance estimator for the sample quantiles is inconsistent when the total number of units deleted from the sample remains bounded [see, e.g., Shao and Wu (1989)].

The following conditions will be used in establishing the results.

CONDITION 1. There are constants θ_L and θ_U such that

$$\theta_L \le \theta_k \le \theta_U \quad \text{for all } k$$

and there are positive constants b, b_1 and b_2 such that

$$b_1 \le F_k'(x) \le b_2 \quad \text{for all } k \text{ and } x \in [\theta_L - b, \theta_U + b], \tag{1.5}$$

where F' is the derivative of F.

CONDITION 2. As $k \to \infty$, $n_k \to \infty$ and there is a positive constant b_0 such that

$$\max_h \frac{W_{kh} n_k}{n_{kh}} \le b_0 \quad \text{for all } k.$$

CONDITION 3. $\{F_k', \ k = 1, 2, \dots\}$ are equicontinuous on $[\theta_L - b, \theta_U + b]$, where θ_L, θ_U and b are given in Condition 1.

A sufficient condition for (1.5) is that $b_1 \le H'_{kh}(x) \le b_2$ for all h, k and $x \in [\theta_L - b, \theta_U + b]$. Similarly, a sufficient condition for the equicontinuity of F'_k is the equicontinuity of H'_{kh}. Note that if F''_k exist and are uniformly bounded on $[\theta_L - b, \theta_U + b]$, then Condition 3 is satisfied.

2. A Bahadur-type representation for stratified samples. Throughout the paper we use $\mathbf{P}$ to denote the unconditional probability corresponding to the random element X and P to denote the conditional probability (given X) corresponding to the random element y. The expectation and variance taken under P are denoted by E and Var, respectively.

PROPOSITION 2.1. *Under Condition* 1, *we have*

$$\mathbf{P}\{|\tilde{\theta}_k - \theta_k| \ge t_k\} \le 2\exp\{-2N_k t_k^2 b_1^2\} \quad \textit{for all } k,$$

where $\{t_k,\ k = 1, 2, \dots\}$ *is a sequence of constants satisfying* $0 < t_k \le b$.

PROOF. Note that

$$\begin{aligned}
&\mathbf{P}\{\tilde{\theta}_k \ge \theta_k + t_k\}\\
&\quad = \mathbf{P}\{p \ge \tilde{F}_k(\theta_k + t_k)\} = \mathbf{P}\Big\{N_k p \ge \sum_h \sum_i I_{(X_{hi} \le \theta_k + t_k)}\Big\}\\
&\quad = \mathbf{P}\Big\{\sum_h \sum_i I_{(X_{hi} > \theta_k + t_k)} - N_k[1 - F_k(\theta_k + t_k)] \ge N_k[F_k(\theta_k + t_k) - p]\Big\}\\
&\quad \le \exp\Big\{-2N_k[F_k(\theta_k + t_k) - p]^2\Big\} \le \exp\{-2N_k t_k^2 b_1^2\},
\end{aligned}$$

where the first inequality follows from Hoeffding's (1963) lemma and the second inequality follows from Condition 1. Similarly, we have

$$\mathbf{P}\{\tilde{\theta}_k \le \theta_k - t_k\} \le \exp\{-2N_k t_k^2 b_1^2\}$$

and thus the result. □

COROLLARY 2.1. *Under the conditions of Proposition* 2.1,

$$\tilde{\theta}_k - \theta_k = O_p\big(N_k^{-1/2}\big) \quad \textit{with respect to } \mathbf{P},$$

and

$$\tilde{\theta}_k - \theta_k = O\big(N_k^{-1/2}(\log N_k)^{1/2}\big) \quad a.s.\ \mathbf{P}.$$

PROOF. Let $t_k = cN_k^{-1/2}(\log N_k)^{1/2}$ with $c^2 > (2b_1^2)^{-1}$. Then apply the result in Proposition 2.1 and the Borel–Cantelli lemma. □

These results imply that $\tilde{\theta}_k$ and θ_k are close when k is large. If the sampling fraction $f_k \to 0$, then the estimation of $\tilde{\theta}_k$ is asymptotically equivalent to the estimation of θ_k since $\tilde{\theta}_k - \theta_k = o_p(n_k^{-1/2})$. In the sequel we shall focus on the estimation of θ_k. The same results for the estimation of $\tilde{\theta}_k$ can be established similarly.

The following lemma extends a result of Bahadur [see Serfling (1980), page 97].

LEMMA 2.1. *Assume Condition* 1. *Let* $a_k = c_0 N_k^{-1/2}(\log N_k)^q$, $q \geq 1/2$, *and*

$$B_k = \sup_{|x| \leq a_k} \left| \left[\tilde{F}_k(\theta_k + x) - \tilde{F}_k(\theta_k) \right] - \left[F_k(\theta_k + x) - p \right] \right|.$$

Then as $k \to \infty$,

$$B_k = O\left(N_k^{-3/4} (\log N_k)^{(q+1)/2} \right) \quad a.s.\, \mathbf{P}.$$

PROOF. Let c_k be the integer part of $c_0 N_k^{1/4}(\log N_k)^q$. For $l = -c_k, \dots, c_k$, let $\eta_{kl} = \theta_k + a_k c_k^{-1} l$, $\alpha_{kl} = F_k(\eta_{k(l-1)}) - F_k(\eta_{kl})$ and

$$G_{kl} = \left| \left[\tilde{F}_k(\eta_{kl}) - \tilde{F}_k(\theta_k) \right] - \left[F_k(\eta_{kl}) - p \right] \right|.$$

Then

$$|\alpha_{kl}| \leq N_k^{-3/4} \sup_{|x| \leq a_k} F_k'(\theta_k + x).$$

Let $\gamma_k = d_0 N_k^{-3/4}(\log N_k)^{(q+1)/2}$ for a constant $d_0 > 0$. Then by Bernstein's inequality,

$$\mathbf{P}\{G_{kl} \geq \gamma_k\} \leq 2 \exp\left\{ -\frac{N_k \gamma_k^2}{2\left(\gamma_k + |F_k(\eta_{kl}) - p| \right)} \right\}.$$

Following the same proof as in Serfling (1980, pages 98–99), we can show the result if there is a constant $\alpha > 0$ such that

$$\sup_{|x| \leq a_k} F_k'(\theta_k + x) \leq \alpha \quad \text{for sufficiently large } k,$$

which is ensured by Condition 1. □

LEMMA 2.2. *Let t be a constant. Under Condition* 1,

$$n_k^{1/2} \left| \tilde{F}_k\left(\theta_k + t n_k^{-1/2} \right) - \tilde{F}_k(\theta_k) - F_k\left(\theta_k + t n_k^{-1/2} \right) + p \right| \to 0 \quad a.s.\, \mathbf{P}, \tag{2.1}$$

and

$$n_k^{1/2} \left| \tilde{F}_k\left(\tilde{\theta}_k + t n_k^{-1/2} \right) - \tilde{F}_k\left(\tilde{\theta}_k \right) - F_k\left(\tilde{\theta}_k + t n_k^{-1/2} \right) + F_k\left(\tilde{\theta}_k \right) \right| \to 0 \quad a.s.\, \mathbf{P}. \tag{2.2}$$

PROOF. We prove the case of $t > 0$. The other case is similar. Since we can use a subsequence argument, we need only to show (2.1)–(2.2) in the following two cases.

CASE 1. $\lim_{k\to\infty} f_k \log N_k = 0$. From Bretagnolle's inequality [see Shorack and Wellner (1986)], there is a constant $C > 0$ such that for any $z > 0$ and k,

$$\mathbf{P}\left\{N_k^{1/2}\left\|\tilde{F}_k - F_k\right\| \geq z\right\} \leq C\exp\{-2z^2\}, \tag{2.3}$$

where $\|\ \|$ is the sup-norm. Note that (2.3) implies that

$$\left\|\tilde{F}_k - F_k\right\| = O\left(N_k^{-1/2}(\log N_k)^{1/2}\right) \quad \text{a.s. } \mathbf{P}. \tag{2.4}$$

Hence (2.1) follows since its left side is bounded above by $2n_k^{1/2}\|\tilde{F}_k - F_k\|$, which goes to zero a.s. $\mathbf{P}$. The proof for (2.2) is the same.

CASE 2. There is a constant $\alpha > 0$ such that $f_k \log N_k \geq \alpha$ for all k. In this case there is a constant c_0 such that

$$tn_k^{-1/2} \leq c_0 N_k^{-1/2}(\log N_k)^{1/2}.$$

Then (2.1) follows from Lemma 2.1 with $q = 1/2$ and (2.2) follows from (2.1), Lemma 2.1 and Corollary 2.1. □

We now prove a Bahadur-type representation for stratified samples.

THEOREM 2.1. *Assume that Conditions* 1–3 *hold. Then*

$$\hat{\theta}_k = \theta_k + \frac{p - \hat{F}_k(\theta_k)}{F_k'(\theta_k)} + o_p\left(n_k^{-1/2}\right) \quad a.s. \mathbf{P} \tag{2.5}$$

in the sense that for almost all array X, the representation (2.5) *holds in the conditional probability P.*

REMARKS.

1. Under the same conditions, we can show a similar result that

$$\hat{\theta}_k = \tilde{\theta}_k + \frac{F_k(\tilde{\theta}_k) - \hat{F}_k(\tilde{\theta}_k)}{F_k'(\tilde{\theta}_k)} + o_p\left(n_k^{-1/2}\right) \quad \text{a.s. } \mathbf{P}.$$

2. Francisco and Fuller (1991) proved (2.5) for stratified clustered samples. In the special case of no clustering, their conditions are neither stronger nor weaker than ours.

PROOF. Let

$$\psi_{k,t} = \theta_k + tn_k^{-1/2}, \qquad G_k = n_k^{1/2}\left[p - \hat{F}_k(\theta_k)\right]/F_k'(\theta_k),$$

$$Z_k(t) = n_k^{1/2}\left[F_k(\psi_{k,t}) - \hat{F}_k(\psi_{k,t})\right]/F_k'(\theta_k),$$

$$U_k(t) = n_k^{1/2}\left[F_k(\psi_{k,t}) - \hat{F}_k(\hat{\theta}_k)\right]/F_k'(\theta_k),$$

and $T_k = n_k^{1/2}(\hat{\theta}_k - \theta_k)$. From Lemma 2.2 and Condition 1,

$$E[Z_k(t) - G_k] = n_k^{1/2}\big[\tilde{F}_k(\theta_k) - \tilde{F}_k(\psi_{k,t}) + F_k(\psi_{k,t}) - p\big]\big/F_k'(\theta_k)$$
$$\to 0 \quad \text{a.s. } \mathbf{P}.$$

For $t > 0$,

$$\begin{aligned}
\operatorname{Var}[Z_k(t) - G_k] &= n_k \frac{\operatorname{Var}\big[\hat{F}_k(\psi_{k,t}) - \hat{F}_k(\theta_k)\big]}{[F_k'(\theta_k)]^2} \\
&= \sum_h \frac{W_{kh}^2 n_k(N_{kh} - n_{kh})}{n_{kh}(N_{kh} - 1)} K_{kh} \frac{1 - K_{kh}}{[F_k'(\theta_k)]^2} \\
&\le b_0 \sum_h \frac{W_{kh} K_{kh}}{b_1^2} = b_0 \frac{\tilde{F}_k(\psi_{k,t}) - \tilde{F}_k(\theta_k)}{b_1^2},
\end{aligned}$$

where $K_{kh} = \tilde{F}_{kh}(\psi_{k,t}) - \tilde{F}_{kh}(\theta_k)$ and the last inequality follows from Conditions 1 and 2. From Condition 1,

$$F_k(\psi_{k,t}) - p \le t n_k^{-1/2} \sup_{\theta_L - b \le x \le \theta_U + b} F_k'(x) \le t n_k^{-1/2} b_2$$

for sufficiently large k. Hence from Lemma 2.2,

$$\operatorname{Var}[Z_k(t) - G_k] \to 0 \quad \text{a.s. } \mathbf{P}.$$

Similarly, for $t < 0$, $\operatorname{Var}[Z_k(t) - G_k] \to 0$ a.s. $\mathbf{P}$. This proves that for any t,

$$Z_k(t) - G_k = o_p(1) \quad \text{a.s. } \mathbf{P}. \tag{2.6}$$

By the mean-value theorem, there is a $\xi_{k,t}$ satisfying $|\xi_{k,t} - \theta_k| \le |t| n_k^{-1/2}$ such that

$$F_k(\psi_{k,t}) - p = t n_k^{-1/2} F_k'(\xi_{k,t}).$$

Then by Condition 3 and $|\theta_k - \xi_{k,t}| \to 0$,

$$n_k^{1/2}[F_k(\psi_{k,t}) - p]/F_k'(\theta_k) = t F_k'(\xi_{k,t})/F_k'(\theta_k) \to t.$$

Also, $|p - \hat{F}_k(\hat{\theta}_k)| \le \max_h W_{kh} n_{kh}^{-1} = O(n_k^{-1})$ under Condition 2. Hence

$$U_k(t) - t = o_p(1) \quad \text{a.s. } \mathbf{P}. \tag{2.7}$$

For any t and $\varepsilon > 0$,

$$\begin{aligned}
P\{T_k \le t, G_k \ge t + \varepsilon\} &= P\{Z_k(t) \le U_k(t), G_k \ge t + \varepsilon\} \\
&\le P\{|G_k - Z_k(t)| \ge \varepsilon/2\} + P\{|U_k(t) - t| \ge \varepsilon/2\} \to 0
\end{aligned}$$

under (2.6)–(2.7). Similarly, $P\{T_k \ge t + \varepsilon, G_k \le t\} \to 0$. Then

$$P\{T_k \le t, G_k \ge t + \varepsilon\} + P\{T_k \le t + \varepsilon, G_k \le t\} = o(1)$$
for all rational t a.s. $\mathbf{P}$.

Using the same argument in the proof of Lemma 1 of Ghosh (1971), we obtain

$$n_k^{1/2}\left[\hat{\theta}_k - \theta_k - \frac{p - \hat{F}_k(\theta_k)}{F_k'(\theta_k)}\right] = T_k - G_k = o_p(1) \quad \text{a.s. } \mathbf{P}. \qquad \square$$

An application of Theorem 2.1 leads to the following useful result.

THEOREM 2.2 (Asymptotic normality). *Assume that Conditions* 1–3 *hold. Let c be a positive constant. For any array X satisfying* (2.5) *and*

$$(2.8) \quad v_k = \frac{n_k}{[F_k'(\theta_k)]^2} \sum_h \frac{W_{kh}^2(N_{kh} - n_{kh})}{n_{kh}(N_{kh} - 1)} \tilde{F}_{kh}(\theta_k)\left[1 - \tilde{F}_{kh}(\theta_k)\right] \geq c,$$

we have

$$(2.9) \qquad n_k^{1/2}(\hat{\theta}_k - \theta_k)/v_k^{1/2} \to_d N(0,1),$$

where $\to_d$ *denotes convergence in distribution under P.*

PROOF. Note that $v_k = n_k \mathrm{Var}[\hat{F}_k(\theta_k)/F_k'(\theta_k)]$. Under (2.8), Theorem 3 of Bickel and Freedman (1984) applies. Hence the result follows from Theorem 2.1. $\square$

REMARKS.

1. If $f_k \to 0$, then v_k in (2.9) can be replaced by

$$(2.10) \qquad v_k^0 = \frac{n_k}{[F_k'(\theta_k)]^2} \sum_h \frac{W_{kh}^2 N_{kh}}{n_{kh}(N_{kh} - 1)} \tilde{F}_{kh}(\theta_k)\left[1 - \tilde{F}_{kh}(\theta_k)\right].$$

This follows from

$$v_k - v_k^0 = -\frac{f_k}{[F_k'(\theta_k)]^2} \sum_h W_{kh} \frac{N_{kh}}{N_{kh} - 1} \tilde{F}_{kh}(\theta_k)\left[1 - \tilde{F}_{kh}(\theta_k)\right].$$

2. If the sampling within each stratum is with replacement, then the result of the theorem holds with v_k replaced by v_k^0.
3. The result still holds if θ_k in (2.8)–(2.9) is replaced by $\tilde{\theta}_k$.
4. For many regular $\hat{\theta}_k$, $v_k - n_k \mathrm{Var}(\hat{\theta}_k') \to 0$ as $k \to \infty$.

3. Representations for BRR variance estimators. From the representation in (2.5), $m_k^{-1}\sum_{\mathbf{s} \in \mathbf{T}_k}(\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k)^2$ has, as its leading term, the left side expression of (3.1). It is straightforward to show that, for any balanced $\mathbf{T}_k$ that satisfies (1.1) and (1.2), the following holds:

$$(3.1) \qquad \frac{1}{m_k} \sum_{\mathbf{s} \in \mathbf{T}_k} \left[\frac{\hat{F}_k^{\mathbf{s}}(\theta_k) - \hat{F}_k(\theta_k)}{F_k'(\theta_k)}\right]^2 = \sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh} r_{kh}} s_{kh}^2,$$

where m_k is the number of elements in $\mathbf{T}_k$ and

$$s_{kh}^2 = \frac{1}{n_{kh}-1}\sum_i \left[I_{(y_{hi}\le\theta_k)} - \frac{1}{n_{kh}}\sum_i I_{(y_{hi}\le\theta_k)}\right]^2 \Big/ [F_k'(\theta_k)]^2.$$

This sets the stage for proving the following theorem on BRR representation, from which the consistency of the BRR variance estimators v_R and $\tilde{v}_R$ in (1.3) and (1.4) naturally follows (see Section 4).

THEOREM 3.1. *Suppose that* $\mathbf{T}_k$ *is a balanced set for each* k. *Assume Conditions* 1–3,

$$n_{kh} \le \tau r_{kh} \quad \textit{for all } k \textit{ and } h \tag{3.2}$$

and there is a constant $c > 0$ *such that*

$$\limsup_k F_k(-c) < p/\tau \quad \textit{and} \quad \liminf_k F_k(c) > 1 - (1-p)/\tau. \tag{3.3}$$

Then

$$\frac{1}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k\right)^2 = \sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh}r_{kh}} s_{kh}^2 + o_p\left(r_k^{-1}\right) \quad a.s.\ \mathbf{P}, \tag{3.4}$$

and

$$\frac{1}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(\hat{\theta}_k - \frac{1}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\hat{\theta}_k^{\mathbf{s}}\right)^2 = \sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh}r_{kh}} s_{kh}^2 + o_p\left(r_k^{-1}\right) \quad a.s.\ \mathbf{P}.$$

REMARKS.

1. Condition (3.2) is the key condition on the resample size. It implies that $n_k \le \tau r_k$ and $r_k \to \infty$ since $n_k \to \infty$ by Condition 2.
2. The first term on the right-hand side of (3.4) is of the order $O_p(r_k^{-1})$ a.s. $\mathbf{P}$ if
$$\sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh}r_{kh}} = O\left(r_k^{-1}\right),$$
which is implied by (3.2) and Condition 2.
3. Condition (3.3) is much weaker than the tightness of F_k, $k = 1, 2, \dots$.

We first prove some preparatory lemmas.

LEMMA 3.1. *Suppose that Condition* 1 *holds. Let* $\rho = b_1 b/2$, *where* b_1 *and* b *are given in Condition* 1.

(i) *If* $\lim_{k\to\infty} f_k \log N_k = 0$, *then for almost all* X, *there is a* $k_X > 0$ *such that for any* t *satisfying* $b < t \le bn_k^{1/2}$ *and* $k \ge k_X$,

$$P\left\{n_k^{1/2}\left|\hat{\theta}_k - \tilde{\theta}_k\right| > t\right\} \le 2\exp\left\{-2\rho^2(tb^{-1}-1)^2\right\}.$$

(ii) *If there is a constant* $\alpha > 0$ *such that* $f_k \log N_k \ge \alpha$ *for all* k, *then for almost all* X, *there is a* $k_X > 0$ *such that for any* t *satisfying* $b < t \le c_0 n_k^{1/2} N_k^{-1/2} (\log N_k)^q$ *and* $k \ge k_X$,

$$P\left\{ n_k^{1/2} \left| \hat{\theta}_k - \tilde{\theta}_k \right| > t \right\} \le 2 \exp\left\{ -2(b_1 t - \rho)^2 \right\},$$

where $c_0 > b$ *and* $q \ge 1/2$ *are fixed constants.*

PROOF. (i) From (2.4) and Corollary 2.1,

$$n_k^{1/2} \left\| \tilde{F}_k - F_k \right\| \to 0 \quad \text{and} \quad n_k^{1/2} \left| \tilde{\theta}_k - \theta_k \right| \to 0 \quad \text{a.s. } \mathbf{P},$$

since $f_k \log N_k \to 0$. Thus, for almost all X, there is a $k_X > 0$ such that

$$n_k^{1/2} \left\| \tilde{F}_k - F_k \right\| < \rho \quad \text{and} \quad n_k^{1/2} \left| \tilde{\theta}_k - \theta_k \right| < b/2 \quad \text{for all } k \ge k_X.$$

Then

$$P\left\{ n_k^{1/2} \left| \hat{\theta}_k - \tilde{\theta}_k \right| > t \right\} \le P\left\{ n_k^{1/2} \left| \hat{\theta}_k - \theta_k \right| > t/2 \right\}.$$

Denote $\theta_k + t n_k^{-1/2}/2$ by $\psi_{k,t/2}$. Note that

$$\begin{aligned}
&P\left\{ \hat{\theta}_k \ge \psi_{k,t/2} \right\} \\
&\quad = P\left\{ p \ge \hat{F}_k(\psi_{k,t/2}) \right\} = P\left\{ \sum_h \frac{W_{kh} n_k}{n_{kh}} \sum_i I_{(y_{hi} > \psi_{k,t/2})} \ge n_k(1-p) \right\} \\
&\quad = P\left\{ \sum_h \frac{W_{kh} n_k}{n_{kh}} \sum_i \left[I_{(y_{hi} > \psi_{k,t/2})} - E\left(I_{(y_{hi} > \psi_{k,t/2})} \right) \right] \ge n_k \left[\tilde{F}_k(\psi_{k,t/2}) - p \right] \right\} \\
&\quad \le \exp\left\{ -2 n_k \left[\tilde{F}_k(\psi_{k,t/2}) - p \right]^2 \right\} \le \exp\left\{ -2\rho^2 (t b^{-1} - 1)^2 \right\},
\end{aligned}$$

where the first inequality follows from Condition 2 and a stratified version of Hoeffding's inequality (1963, Section 6), and the second inequality follows from

$$\begin{aligned}
n_k^{1/2} \left| \tilde{F}_k(\psi_{k,t/2}) - p \right| &\ge n_k^{1/2} \left| F_k(\psi_{k,t/2}) - p \right| - n_k^{1/2} \left| \tilde{F}_k(\psi_{k,t/2}) - F_k(\psi_{k,t/2}) \right| \\
&\ge b_1 t/2 - \rho = \rho(t b^{-1} - 1).
\end{aligned}$$

Similarly, $P\{\hat{\theta}_k \le \theta_k - t n_k^{-1/2}/2\} \le \exp\{-2\rho^2 (t b^{-1} - 1)^2\}$. This proves (i).

(ii) Let $\phi_{k,t} = \tilde{\theta}_k + t n_k^{-1/2}$. Similar to the proof of (i),

$$\begin{aligned}
&P\left\{ \hat{\theta}_k \ge \phi_{k,t} \right\} \\
&\quad = P\left\{ \sum_h \frac{W_{kh} n_k}{n_{kh}} \sum_i I_{(y_{hi} > \phi_{k,t})} \ge n_k(1-p) \right\} \\
&\quad = P\left\{ \sum_h \frac{W_{kh} n_k}{n_{kh}} \sum_i \left[I_{(y_{hi} > \phi_{k,t})} - E\left(I_{(y_{hi} > \phi_{k,t})} \right) \right] \ge n_k \left[\tilde{F}_k(\phi_{k,t}) - p \right] \right\} \\
&\quad \le \exp\left\{ -2 n_k \left[\tilde{F}_k(\phi_{k,t}) - p \right]^2 \right\}.
\end{aligned}$$

From Lemma 2.1 and the fact that $n_k^{1/2} N_k^{-3/4} (\log N_k)^{(q+1)/2} \to 0$,

$$\sup_{b < t \le \zeta_k} n_k^{1/2} \left| \tilde{F}_k(\phi_{k,t}) - \tilde{F}_k(\tilde{\theta}_k) - F_k(\phi_{k,t}) + F_k(\tilde{\theta}_k) \right| \to 0 \quad \text{a.s. } \mathbf{P},$$

where $\zeta_k = c_0 n_k^{-1/2} N_k^{1/2} (\log N_k)^q$. Also, $|\tilde{F}_k(\tilde{\theta}_k) - p| \le N_k^{-1}$. Hence for almost all X, there is a $k_X > 0$ such that for $k \ge k_X$ and $b < t \le \zeta_k$,

$$n_k^{1/2} \left| \tilde{F}_k(\phi_{k,t}) - p - F_k(\phi_{k,t}) + F_k(\tilde{\theta}_k) \right| \le \rho.$$

From $n_k^{1/2}[F_k(\phi_{k,t}) - F_k(\tilde{\theta}_k)] \ge b_1 t > \rho$, we have

$$n_k \left[\tilde{F}_k(\phi_{k,t}) - p \right]^2 \ge (b_1 t - \rho)^2$$

and therefore

$$P\left\{ \hat{\theta}_k \ge \phi_{k,t} \right\} \le \exp\left\{ -2(b_1 t - \rho)^2 \right\}.$$

Similarly, $P\{\hat{\theta}_k \le \tilde{\theta}_k - t n_k^{-1/2}\} \le \exp\{-2(b_1 t - \rho)^2\}$. This completes the proof. □

LEMMA 3.2. *Suppose that Condition 2 holds. For any* $\{t_k,\ k = 1, 2, \dots\}$, *we have*

$$n_k^2 E\left[\hat{F}_k(t_k) - \tilde{F}_k(t_k) \right]^4 = O(1) \quad a.s.\ \mathbf{P}.$$

PROOF. Suppose that $\{x_{hi},\ i = 1, \dots, n_{kh}\}$ is a simple random sample with replacement from $\{X_{h1}, \dots, X_{hN_{kh}}\}$. Let $z_{kh} = n_{kh}^{-1} \sum_i I_{(y_{hi} \le t_k)} - F_{kh}(t_k)$ and $u_{kh} = n_{kh}^{-1} \sum_i I_{(x_{hi} \le t_k)} - F_{kh}(t_k)$. Since $0 \le I_{(x_{hi} \le t_k)} \le 1$, there is a constant $C > 0$ such that

$$E u_{kh}^2 \le C^{1/2} n_{kh}^{-1}, \qquad E u_{kh}^4 \le C n_{kh}^{-2} \quad \text{a.s. } \mathbf{P}.$$

This and Theorem 4 of Hoeffding (1963) imply

$$E z_{kh}^2 \le C^{1/2} n_{kh}^{-1}, \qquad E z_{kh}^4 \le C n_{kh}^{-2} \quad \text{a.s. } \mathbf{P}. \tag{3.5}$$

Note that

$$\begin{aligned} E\left[\hat{F}_k(t_k) - \tilde{F}_k(t_k) \right]^4 &= E\left(\sum_h W_{kh} z_{kh} \right)^4 \\ &\le 2\left[E\left(\sum_h W_{kh}^2 z_{kh}^2 \right)^2 + E\left(\sum_{h \ne g} W_{kh} W_{kg} z_{kh} z_{kg} \right)^2 \right]. \end{aligned}$$

Applying (3.5), we have

$$E\left(\sum_h W_{kh}^2 z_{kh}^2 \right)^2 = E\left(\sum_h W_{kh}^4 z_{kh}^4 \right) + E\left(\sum_{h \ne g} W_{kh}^2 W_{kg}^2 z_{kh}^2 z_{kg}^2 \right) \le C\left(\sum_h W_{kh}^2 n_{kh}^{-1} \right)^2$$

and

$$E\left(\sum_{h\neq g} W_{kh}W_{kg}z_{kh}z_{kg}\right)^2 = \sum_{h\neq g} W_{kh}^2W_{kg}^2 Ez_{kh}^2 Ez_{kg}^2 \le C\left(\sum_h W_{kh}^2 n_{kh}^{-1}\right)^2.$$

Then the result follows from

$$n_k^2\left(\sum_h W_{kh}^2 n_{kh}^{-1}\right)^2 \le b_0^2\left(\sum_h W_{kh}\right)^2 = b_0^2$$

under Condition 2. □

LEMMA 3.3. *Assume* (3.2) *and* (3.3). *Then for almost all* X,

$$P\left\{\max_{\mathbf{s}\in\mathbf{S}_k}\left|\hat\theta_k^{\mathbf{s}} - \hat\theta_k\right| \ge 2c\right\} \to 0,$$

where c *is given in* (3.3).

PROOF. Let $\mathbf{s}_{(1)}$ and $\mathbf{s}_{(2)}$ be the two sets such that $\hat\theta_k^{\mathbf{s}(1)}$ and $\hat\theta_k^{\mathbf{s}(2)}$ are, respectively, the minimum and maximum among the $\hat\theta_k^{\mathbf{s}}$, $\mathbf{s}\in\mathbf{S}_k$. Then

$$\max_{\mathbf{s}\in\mathbf{S}_k}\left|\hat\theta_k^{\mathbf{s}} - \hat\theta_k\right| \le \hat\theta_k^{\mathbf{s}(2)} - \hat\theta_k^{\mathbf{s}(1)}.$$

It suffices to show that for almost all X,

$$P\left\{\hat\theta_k^{\mathbf{s}(2)} \ge c\right\} \to 0 \quad \text{and} \quad P\left\{\hat\theta_k^{\mathbf{s}(1)} \le -c\right\} \to 0.$$

From (3.2), $1 - \hat F_k^{\mathbf{s}(2)}(c) \le \tau[1 - \hat F_k(c)]$. From (2.4), (3.3) and $\hat F_k(c) - \tilde F_k(c) = o_p(1)$ a.s. $\mathbf{P}$,

$$P\left\{\hat\theta_k^{\mathbf{s}(2)} \ge c\right\} = P\left\{1 - \hat F_k^{\mathbf{s}(2)}(c) \ge 1 - p\right\} \le P\left\{1 - \hat F_k(c) \ge (1-p)/\tau\right\} \to 0 \quad \text{a.s. } \mathbf{P}.$$

Similarly we can show $P\{\hat\theta_k^{\mathbf{s}(1)} \le -c\} \to 0$ a.s. $\mathbf{P}$. This completes the proof. □

LEMMA 3.4. *Let* $\mathbf{T}_k$ *be a subset of* $\mathbf{S}_k$ *with size* m_k. *Assume Condition* 1, (3.2) *and* (3.3). *Then*

$$\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(\hat\theta_k^{\mathbf{s}} - \tilde\theta_k\right)^4 = O_p(1) \quad a.s.\,\mathbf{P}. \tag{3.6}$$

PROOF. From Lemma 3.3 and the fact that $\hat\theta_k - \tilde\theta_k = o_p(1)$ a.s. $\mathbf{P}$, for almost all X,

$$P\left\{\max_{\mathbf{s}\in\mathbf{T}_k}\left|\hat\theta_k^{\mathbf{s}} - \hat\theta_k\right| \ge b r_k^{1/2}\right\} \to 0. \tag{3.7}$$

For (3.6), it suffices to show that for almost all X, if $\{k_j\}$ is a subsequence of integers such that $\lim_{j\to\infty} f_{k_j} \log N_{k_j} = a$ (including ∞), then (3.6) holds with k replaced by k_j. For simplicity we omit the subscript j in the following.

CASE 1. $a = 0$. From Lemma 3.1(i), for almost all X, there is a $k_X > 0$ such that for $k \geq k_X$ and any $\mathbf{s} \in \mathbf{S}_k$,

$$(3.8) \quad P\Big\{r_k^{1/2}\big|\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k\big| > t\Big\} \leq 2\exp\Big\{-2\rho^2(tb^{-1} - 1)^2\Big\} \quad \text{for } b < t \leq br_k^{1/2}.$$

From (3.7), the desired result follows if

$$\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\Big(\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\Big)^4 I_{(|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k| \leq br_k^{1/2})} = O_p(1),$$

which holds if

$$(3.9) \qquad r_k^2 E\Big[\Big(\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\Big)^4 I_{(|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k| \leq br_k^{1/2})}\Big] = O(1),$$

since $\hat{\theta}_k^{\mathbf{s}}$ have the same distribution for all $\mathbf{s} \in \mathbf{S}_k$. Note that the left-hand side of (3.9) equals

$$\begin{aligned}
&4\int_0^\infty t^3 P\Big\{r_k^{1/2}\big|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\big| I_{(|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k| \leq br_k^{1/2})} > t\Big\}\,dt \\
&\quad\leq 4\int_0^{br_k} t^3 P\Big\{r_k^{1/2}\big|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\big| > t\Big\}\,dt \\
&\quad\leq 4\int_0^b t^3\,dt + 4\int_b^{br_k^{1/2}} t^3 P\Big\{r_k^{1/2}\big|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\big| > t\Big\}\,dt \\
&\qquad + 4\int_{br_k^{1/2}}^{br_k} t^3 P\Big\{r_k^{1/2}\big|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\big| > t\Big\}\,dt \\
&\quad\leq b^4 + 8\int_b^{br_k^{1/2}} t^3\exp\big\{-2\rho^2(t/b - 1)^2\big\}\,dt + 4P\Big\{\big|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\big| > b\Big\}\int_{br_k^{1/2}}^{br_k} t^3\,dt \\
&\quad\leq b^4 + 8\int_b^\infty t^3\exp\big\{-2\rho^2(t/b - 1)^2\big\}\,dt + 2b^4 r_k^4\exp\Big\{-2\rho^2\big(r_k^{1/2} - 1\big)^2\Big\},
\end{aligned}$$

where the inequalities follow from (3.8). Thus (3.9) holds.

CASE 2. $a > 0$. There is an $\alpha > 0$ such that $r_k \log N_k/N_k \geq \alpha$. From Lemma 3.1(ii) with $q = 2$, for almost all X, there is an $l_X > 0$ such that for $k \geq l_X$ and any $\mathbf{s} \in \mathbf{S}_k$,

$$(3.10) \quad P\Big\{r_k^{1/2}\big|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\big| > t\Big\} \leq 2\exp\big\{-2(b_1 t - \rho)^2\big\} \quad \text{for } b < t \leq c_0 a_k,$$

where $a_k = r_k^{1/2}(\log N_k)^2/N_k^{1/2}$ and c_0 is a constant satisfying $c_0^2\alpha \geq 2$. Similar to the proof of Case 1, we only need to show (3.9). Note that the left-hand side of (3.9) is bounded by

$$4\int_0^b t^3\,dt + 4\int_b^{c_0 a_k} t^3 P\left\{r_k^{1/2}\left|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\right| > t\right\}dt + 4\int_{c_0 a_k}^{b r_k} t^3 P\left\{r_k^{1/2}\left|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\right| > t\right\}dt$$

$$\le b^4 + 8\int_b^\infty t^3 \exp\left\{-2(b_1 t - \rho)^2\right\}dt$$

$$+ 4P\left\{\left|\hat{\theta}_k^{\mathbf{s}} - \tilde{\theta}_k\right| > c_0 N_k^{-1/2}(\log N_k)^2\right\}\int_0^{b r_k} t^3\,dt$$

$$\le b^4 + 8\int_b^\infty t^3 \exp\left\{-2(b_1 t - \rho)^2\right\}dt + 2b^4 r_k^4 \exp\left\{-2(b_1 c_0 a_k - \rho)^2\right\}, \tag{3.11}$$

where the inequalities follow from (3.10). Since $r_k^{1/2}(\log N_k)^2/N_k^{1/2} \to \infty$, there is a k_0 such that for $k \ge k_0$,

$$b_1 c_0 a_k \ge \rho + c_0 r_k^{1/2} \log N_k / N_k^{1/2} \ge c_0 \alpha^{1/2}(\log N_k)^{1/2}.$$

Hence for $k \ge k_0$, the last term in (3.11) is bounded by

$$2b^4 r_k^4 \exp\left\{-2c_0^2 \alpha \log N_k\right\} = 2b^4 r_k^4 N_k^{-2c_0^2\alpha} \le 2b^4 r_k^4 N_k^{-4} \le 2b^4.$$

Thus (3.9) holds for $k \ge k_X = \max(k_0, l_X)$. This completes the proof. □

PROOF OF THEOREM 3.1. We only prove the first assertion. The proof of the second assertion is similar. Note that

$$\frac{1}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k\right)^2 = \frac{1}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left[\frac{\hat{F}_k(\theta_k) - \hat{F}_k^{\mathbf{s}}(\theta_k)}{F_k'(\theta_k)} + R_k^{\mathbf{s}} - R_k\right]^2$$

$$= \sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh} r_{kh}} s_{kh}^2 + \frac{1}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(R_k^{\mathbf{s}} - R_k\right)^2$$

$$+ \text{ cross product terms,}$$

where the last equality follows from (3.1) and

$$R_k^{\mathbf{s}} = \hat{\theta}_k^{\mathbf{s}} - \theta_k - \left[p - \hat{F}_k^{\mathbf{s}}(\theta_k)\right]\Big/F_k'(\theta_k),$$

$$R_k = \hat{\theta}_k - \theta_k - \left[p - \hat{F}_k(\theta_k)\right]\Big/F_k'(\theta_k).$$

Then the result follows from

$$\frac{r_k}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(R_k^{\mathbf{s}} - R_k\right)^2 = o_p(1) \quad \text{a.s. } \mathbf{P}, \tag{3.12}$$

since the cross product terms are $o_p(1)$ a.s. $\mathbf{P}$ under (3.12), (3.2) and Condition 2. Note that

$$\begin{aligned}\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^4 &\le \frac{8r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left(\hat\theta_k^{\mathbf{s}}-\hat\theta_k\right)^4\\ &\quad+\frac{8r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left[\frac{\hat F_k^{\mathbf{s}}(\theta_k)-\hat F_k(\theta_k)}{F_k'(\theta_k)}\right]^4.\end{aligned}\tag{3.13}$$

From Theorem 2.2 and Corollary 2.1,

$$\hat\theta_k-\tilde\theta_k=O_p\left(n_k^{-1/2}\right)\quad\text{a.s. }\mathbf{P}.$$

Then from Lemma 3.4, the first term on the right-hand side of (3.13) is $O_p(1)$ a.s. $\mathbf{P}$. Note that

$$E\left\{\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}\left[\frac{\hat F_k^{\mathbf{s}}(\theta_k)-\hat F_k(\theta_k)}{F_k'(\theta_k)}\right]^4\right\}=r_k^2E\left[\frac{\hat F_k^{\mathbf{s}}(\theta_k)-\hat F_k(\theta_k)}{F_k'(\theta_k)}\right]^4$$

for a fixed $\mathbf{s}\in\mathbf{S}_k$. Hence from Lemma 3.2, the second term on the right-hand side of (3.13) is $O_p(1)$ a.s. $\mathbf{P}$. This proves that for almost all X,

$$\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^4=O_p(1).\tag{3.14}$$

From Theorem 2.1, $r_k(R_k^{\mathbf{s}}-R_k)^2=o_p(1)$ a.s. $\mathbf{P}$ for any fixed $\mathbf{s}$. Let a be any positive rational number. Then

$$r_kE\left[(R_k^{\mathbf{s}}-R_k)^2I_{[r_k(R_k^{\mathbf{s}}-R_k)^2\le a]}\right]\to 0\quad\text{a.s. }\mathbf{P}$$

and therefore for almost all X,

$$\frac{r_k}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^2I_{[r_k(R_k^{\mathbf{s}}-R_k)^2\le a]}=o_p(1)\quad\text{for } all \text{ rational } a.\tag{3.15}$$

For any $\varepsilon>0$, let X be fixed such that (3.14) and (3.15) hold. Then there is a positive rational a depending on ε such that

$$P\left\{\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^4>a\varepsilon\right\}<\varepsilon.$$

Then

$$\begin{aligned}&P\left\{\frac{r_k}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^2I_{[r_k(R_k^{\mathbf{s}}-R_k)^2>a]}>\varepsilon\right\}\\ &\qquad\le P\left\{\frac{r_k^2}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^4I_{[r_k(R_k^{\mathbf{s}}-R_k)^2>a]}>a\varepsilon\right\}<\varepsilon.\end{aligned}$$

Hence

$$(3.16)\qquad \begin{aligned} &P\left\{\frac{r_k}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^2>2\varepsilon\right\}\\ &\quad<\varepsilon+P\left\{\frac{r_k}{m_k}\sum_{\mathbf{s}\in\mathbf{T}_k}(R_k^{\mathbf{s}}-R_k)^2 I_{[r_k(R_k^{\mathbf{s}}-R_k)^2\le a]}>\varepsilon\right\}.\end{aligned}$$

From (3.15), the second term on the right-hand side of (3.16) is smaller than ε for large k. Since ε is arbitrary, this proves (3.12) and thus the result. □

4. Consistency of BRR variance estimators. We shall establish the consistency of the BRR variance estimators v_R and $\tilde{v}_R$ in (1.3) and (1.4) for any balanced $\mathbf{T}_k$ and with proper conditions on r_{kh}, d_{kh} and choice of λ_k.

4.1. *Non random resample sizes.* We first consider the case of $f_k \to 0$ or with-replacement sampling within stratum. From Remarks 1 and 2 after Theorem 2.2, we need to estimate v_k^0 in this case. Note that

$$E\left[n_k\sum_h \frac{W_{kh}^2}{n_{kh}}s_{kh}^2\right]=v_k^0,$$

and

$$\operatorname{Var}\left[n_k\sum_h \frac{W_{kh}^2}{n_{kh}}s_{kh}^2\right]=n_k^2\sum_h\frac{W_{kh}^4}{n_{kh}^2}\operatorname{Var}(s_{kh}^2)\le Cn_k^2\sum_h\frac{W_{kh}^4}{n_{kh}^3}\le Cb_0^3n_k^{-1},$$

where C is a constant and the two inequalities follow from Conditions 1 and 2, respectively. The BRR estimators v_R and $\tilde{v}_R$ are consistent for v_k^0 if

$$(4.1)\qquad \lambda_k^{-1}=O(1),$$

$$(4.2)\qquad \frac{n_k}{\lambda_k}\sum_h\frac{W_{kh}^2}{n_{kh}}\left(\frac{d_{kh}}{r_{kh}}-\lambda_k\right)s_{kh}^2=o_p(1)\quad \text{a.s. }\mathbf{P}$$

and the conditions in Theorem 3.1 are satisfied.

THEOREM 4.1. *Assume that the conditions in Theorem* 3.1 *hold and that*

$$(4.3)\qquad n_k\le cd_k,$$

where c is a constant. Let $\lambda_k=d_k/r_k$ and assume that

$$(4.4)\qquad \sum_h W_{kh}\left|\frac{d_{kh}}{r_{kh}}-\lambda_k\right|\to 0.$$

Then for balanced $\mathbf{T}_k$,

$$v_R(\mathbf{T}_k) - v_k^0 = o_p(1) \quad a.s.\ \mathbf{P},$$

and

$$\tilde{v}_R(\mathbf{T}_k) - v_k^0 = o_p(1) \quad a.s.\ \mathbf{P}.$$

PROOF. We need to check (4.1) and (4.2). Since $\lambda_k = d_k/r_k$, (4.1) follows from (4.3) and (4.2) follows from (4.1), (4.4) and Condition 2. □

In the following we give some examples of r_{kh}, and λ_k that satisfy (4.3)–(4.4). The results are stated as corollaries.

COROLLARY 4.1 (Proportional BRR). *Let* r_k *and* d_k *be integers satisfying* (4.3) *and* $\lambda_k = d_k/r_k$. *Suppose that we can choose* r_{kh} *to be proportional to* n_{kh}, *that is,*

$$\frac{r_{kh}}{n_{kh}} = \frac{r_k}{n_k}, \qquad h = 1, \dots, L_k. \tag{4.5}$$

Then (4.4) *is satisfied and the result in Theorem* 4.1 *holds.*

In Corollary 4.1, n_{kh} is not required to be divisible by r_{kh}, nor does it require the original sampling plan to be proportional allocation. In the important special case of n_k/r_k being an integer, we have the following result.

COROLLARY 4.2. *Suppose that* n_{kh} *for any* h *is a multiple of* p_k, p_k *being an integer. Let* $r_{kh} = n_{kh}/p_k$ *and* $\lambda_k = p_k - 1$. *Then* (4.3) *and* (4.4) *are satisfied and the result in Theorem* 4.1 *holds.*

Regarding the selection of $\mathbf{T}_k$ for Corollary 4.1 or 4.2, one obvious choice is $\mathbf{T}_k = \mathbf{S}_k$, since $\mathbf{S}_k$ is balanced. The number of elements in $\mathbf{S}_k$, $M_k = \prod_h \binom{n_{kh}}{r_{kh}}$, is usually very large and therefore the computation of $v_R(\mathbf{S}_k)$ can be cumbersome. In some situations, a balanced set $\mathbf{T}_k$ with size much smaller than M_k can be found. An outstanding example is the use of Hadamard matrix for constructing $\mathbf{T}_k$ when $n_{kh} = 2$ and $r_{kh} = 1$. The size of $\mathbf{T}_k$ does not exceed $L_k + 4$, where L_k is the number of strata [McCarthy (1969)]. When $n_{kh} = n$ is a prime power and $r_{kh} = 1$, construction of $\mathbf{T}_k$ with economic size is also available [Gurney and Jewett (1975)]. To allow r_{kh} in Corollary 4.2 to be greater than 1, we can use a generalization of orthogonal array [see Brickell (1984)] to construct balanced $\mathbf{T}_k$. Construction of these arrays with economic size needs to be further investigated. Finally we may point out that (4.4) is generally not satisfied by $r_{kh} = 1$ and unequal n_{kh}, and consequently the variance estimators v_R (1.3) and $\tilde{v}_R$ (1.4) may not be consistent. By using a mixed orthogonal array that satisfies the balance conditions (1.1) and (1.2) and employing internal scalings to account for the difference in the n_{kh}'s, Wu (1991) obtained alternative variance estimators that are consistent.

In the following corollary we relax the assumption of constant n_{kh}/r_{kh} required in Corollaries 4.1 and 4.2.

COROLLARY 4.3. *Suppose that r_{kh} satisfies the following conditions for a sequence $\mathbf{U}_k \subset \{1, 2, \ldots, L_k\}$, $k = 1, 2, \ldots$:*

(a) $\max_{h \in \mathbf{U}_k} |d_{kh}/r_{kh} - \lambda| \to 0$, *where $\lambda > 0$ is a constant*;
(b) $\sum_{h \in \mathbf{U}_k^c} W_{kh} \to 0$, *where $\mathbf{U}_k^c$ is the complement of $\mathbf{U}_k$*;
(c) $d_{kh} \le C r_{kh}$ *for $h \in \mathbf{U}_k^c$, where C is a constant*;
(d) $\sum_{h \in \mathbf{U}_k^c} r_{kh}/r_k \to 0$.

Then if $\lambda_k \equiv \lambda$, (4.3) *and* (4.4) *are satisfied and the result in Theorem* 4.1 *holds.*

PROOF. From (a), (c) and (d), $d_k/r_k \to \lambda$ and therefore (4.3) holds. Condition (4.4) follows from (a), (b) and (c). □

We now consider the case of nonnegligible f_k. When f_k does not tend to zero, the asymptotic variance of $n^{1/2}(\hat{\theta}_k - \theta_k)$ is v_k given in (2.8). Similar results to those in Theorem 4.1 and Corollaries 4.1–4.3 can be established. Note that condition (4.2) should be replaced by

$$\frac{n_k}{\lambda_k} \sum_h \frac{W_{kh}^2}{n_{kh}} \left[\frac{d_{kh}}{r_{kh}} - \lambda_k(1 - f_{kh}) \right] s_{kh}^2 = o_p(1) \quad \text{a.s. } \mathbf{P},$$

where $f_{kh} = n_{kh}/N_{kh}$ is the sampling fraction within the hth stratum. The proof of the following theorem is similar to that of Theorem 4.1 and is omitted.

THEOREM 4.2. *Assume the conditions in Theorem* 3.1 *and that r_{kh} and λ_k are chosen so that* (4.3) *holds, $\lambda_k = d_k/(r_k - \sum_h f_{kh} r_{kh})$ and*

$$\sum_h W_{kh} \left| \frac{d_{kh}}{r_{kh}} - \lambda_k(1 - f_{kh}) \right| \to 0. \tag{4.6}$$

Then for balanced $\mathbf{T}_k$,

$$v_R(\mathbf{T}_k) - v_k = o_p(1) \quad a.s.\ \mathbf{P},$$

and

$$\tilde{v}_R(\mathbf{T}_k) - v_k = o_p(1) \quad a.s.\ \mathbf{P}.$$

If the sampling plan is proportional allocation, that is, $f_{kh} \equiv n_k/N_k$, then condition (4.6) is the same as condition (4.4). Hence if r_{kh} are chosen according to the methods described in Corollaries 4.1–4.3, v_R and $\tilde{v}_R$ are consistent for v_k. In particular, choosing r_{kh} according to (4.5) is justifiable.

If f_{kh} are different for different strata, choosing r_{kh} to satisfy (4.6) may not be possible. In the following we consider random resample sizes r_{kh}, which is a relatively easy way to obtain consistent BRR variance estimators.

4.2. *Random resample sizes.* Assume $f_k \to 0$ or with-replacement sampling within each stratum, but allow the resample sizes r_{kh} to be random. This may be useful when a suitable choice of r_{kh} and λ_k satisfying the conditions in Corollaries 4.1–4.3 does not exist. For example, Corollary 4.2 is not applicable if n_{kh} is not a multiple of p_k for any h. Let

$$w_{kh} = n_{kh}/2 \tag{4.7}$$

and $\lambda_k \equiv 1$. Choose r_{kh} according to

$$\begin{aligned} r_{kh} &= [w_{kh}] \quad \text{with probability } p_{kh}, \\ r_{kh} &= 1 + [w_{kh}] \quad \text{with probability } 1 - p_{kh}, \end{aligned} \tag{4.8}$$

where

$$p_{kh} = \frac{(1 + [w_{kh}] - w_{kh})[w_{kh}]}{w_{kh}}$$

and $[x]$ is the integer part of x. Note that if n_{kh} is an even integer, then $p_{kh} = 1$ and $r_{kh} = n_{kh}/2$ is nonrandom. The following result shows that the BRR estimators with this choice of r_{kh} are consistent.

THEOREM 4.3. *Assume the conditions in Theorem* 3.1 *and that* $\mathbf{T}_k$ *are balanced. Let* $\lambda_k \equiv 1$ *and* r_{kh} *be chosen according to* (4.8). *Then the result in Theorem* 4.1 *holds.*

PROOF. Since $n_k/r_k \le n_k/\sum_h [n_{kh}/2] \le 6$, (4.1) holds. It remains to show (4.2). Let E_r and Var_r be the expectation and variance taken under the probability corresponding to the random selection (4.8). Then $E_r(d_{kh}/r_{kh}) = 1$. If n_{kh} is even, $\mathrm{Var}_r(r_{kh}^{-1}) = 0$. For odd n_{kh}, a straightforward calculation shows that $\mathrm{Var}_r(r_{kh}^{-1}) \le c n_{kh}^{-3}$, where c is a constant. Then

$$E_r\left(n_k \sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh} r_{kh}} s_{kh}^2\right) = n_k \sum_h \frac{W_{kh}^2}{n_{kh}} s_{kh}^2,$$

and

$$\begin{aligned} \mathrm{Var}_r\left(n_k \sum_h \frac{W_{kh}^2 d_{kh}}{n_{kh} r_{kh}} s_{kh}^2\right) &= n_k^2 \sum_h \frac{W_{kh}^4}{n_{kh}^2} \mathrm{Var}_r\left(\frac{n_{kh} - r_{kh}}{r_{kh}}\right) s_{kh}^4 \\ &\le c n_k^2 \sum_h \frac{W_{kh}^4}{n_{kh}^3} s_{kh}^4 \le c b_0^3 n_k^{-1} \sum_h W_{kh} s_{kh}^4 = o_p(1) \end{aligned}$$

a.s. $\mathbf{P}$.

Hence (4.2) holds. □

The scheme (4.7)–(4.8) for choosing r_{kh} is especially useful when f_k is nonnegligible. We have a similar result to Theorem 4.3. Its proof is omitted.

THEOREM 4.4. *Assume the conditions in Theorem* 3.1 *and that* $\mathbf{T}_k$ *are balanced. Let* $\lambda_k \equiv 1$ *and* r_{kh} *be chosen according to* (4.7)–(4.8) *with* w_{kh} *replaced by* $n_{kh}(1 - f_{kh})/2$. *Then the result in Theorem* 4.2 *holds.*

5. Random subsampling. As pointed out in Section 4.1, balanced $\mathbf{T}_k$ with economic size are only available in some special cases. In this section we study another method for computational reduction: the random subsampling method. It is useful when a balanced set $\mathbf{T}_k$ with size much smaller than M_k is not available for a fixed k. Even if a balance set $\mathbf{T}_k$ with size much smaller than M_k exists, use of the random subsampling may be simpler than the construction of a balanced set. It can be viewed as a Monte Carlo approximation to the complete enumeration of $\mathbf{S}_k$.

Suppose that

$$\mathbf{S}_k = \bigcup_{j=1}^{v_k} \mathbf{S}_{k,j}, \tag{5.1}$$

where $\mathbf{S}_{k,j}$ are disjoint subsets of $\mathbf{S}_k$ and have equal number of elements l_k. Let $\{\mathbf{T}^*_{k,1}, \dots, \mathbf{T}^*_{k,u_k}\}$ be a simple random sample (with or without replacement) from $\{\mathbf{S}_{k,1}, \dots, \mathbf{S}_{k,v_k}\}$,

$$\mathbf{T}^*_k = \bigcup_{j=1}^{u_k} \mathbf{T}^*_{k,j} \tag{5.2}$$

and m_k be the total number of elements in $\mathbf{T}^*_k$. Note that $M_k = v_k l_k$ and $m_k = u_k l_k$. Usually m_k is chosen to be much smaller than M_k. When each $\mathbf{S}_{k,j}$ in (5.1) contains only one element, $v_k = M_k$ and the method amounts to taking a simple random sample $\{\mathbf{s}_1, \dots, \mathbf{s}_{u_k}\}$ from $\mathbf{S}_k$. This special case may be called a *complete random subsampling* method. For i.i.d. samples, Shao (1989) studied the method and suggested taking m_k to be n_k^δ for some $\delta > 1$. Another special case for $n_{kh} = g$, $r_{kh} = 1$ for all h and k is to define $\mathbf{S}_{k,j}$ in (5.1) to be a collection of g mutually exclusive subsamples, each of which contains L_k units with one from each stratum. Each $\mathbf{S}_{k,j}$ amounts to grouping the gL_k units into g exclusive subsamples. There are $g^{L_k - 1}$ such groupings to make up the $\mathbf{S}_k$ in (5.1) with $v_k = g^{L_k - 1}$. This method is called the *repeated random-group* method and the estimator v_R is studied in Kovar, Rao and Wu (1988).

The following result shows that $v_R(\mathbf{T}^*_k)$ and $\tilde{v}_R(\mathbf{T}^*_k)$ can be used to approximate $v_R(\mathbf{S}_k)$ and $\tilde{v}_R(\mathbf{S}_k)$, respectively. Thus, $v_R(\mathbf{T}^*_k)$ and $\tilde{v}_R(\mathbf{T}^*_k)$ are consistent variance estimators if $v_R(\mathbf{S}_k)$ and $\tilde{v}_R(\mathbf{S}_k)$ are.

THEOREM 5.1. *Let* $\mathbf{T}^*_k$ *and* u_k *be given in* (5.2). *Suppose that Condition* 1 *and* (4.1) *hold and that* $u_k \to \infty$ *as* $k \to \infty$. *Then*

$$v_R(\mathbf{T}^*_k) - v_R(\mathbf{S}_k) = o_p(1),$$

and

$$\tilde{v}_R(\mathbf{T}^*_k) - \tilde{v}_R(\mathbf{S}_k) = o_p(1).$$

PROOF. Let $A_k^* = v_R(\mathbf{T}_k^*) - v_R(\mathbf{S}_k)$, P_* be the probability corresponding to the random selection of $\mathbf{T}_k^*$ and E_* and Var_* be the expectation and variance taken under P_*. Then $E_*(A_k^*) = 0$ and

$$\begin{aligned}\mathrm{Var}_*(A_k^*) &= \frac{v_k - u_k}{u_k v_k (v_k - 1)} \sum_{j=1}^{v_k} \left[\sum_{\mathbf{s} \in \mathbf{S}_{k,j}} \frac{n_k}{\lambda_k l_k} \left(\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k\right)^2 - v_R(\mathbf{S}_k) \right]^2 \\ &\le \frac{n_k^2}{\lambda_k^2 l_k^2 u_k v_k} \sum_{j=1}^{v_k} \left[\sum_{\mathbf{s} \in \mathbf{S}_{k,j}} \left(\hat{\theta}_k^{\mathbf{s}} - \hat{\theta}_k\right)^2 \right]^2 \\ &\le \frac{c^2 r_k^2}{u_k M_k} \sum_{\mathbf{s} \in \mathbf{S}_k} \left(\hat{\theta}_k^{\mathbf{s}} - \theta_k\right)^4 \doteq O_p\left(u_k^{-1}\right) \quad \text{a.s. } \mathbf{P},\end{aligned}$$

where c is given in condition (4.1) and the last equality follows from Lemma 3.4. Hence for any $\varepsilon > 0$,

$$P_*\{|A_k^*| > \varepsilon\} \le \varepsilon^{-2}\,\mathrm{Var}_*(A_k^*) = o_p(1) \quad \text{a.s. } \mathbf{P}.$$

Then the first assertion of the theorem follows from

$$P\{|A_k^*| > \varepsilon\} = E\left[P_*\{|A_k^*| > \varepsilon\}\right] = o(1) \quad \text{a.s. } \mathbf{P}.$$

The proof of the second assertion is similar. □

REFERENCES

BICKEL, P. J. and FREEDMAN, D. A. (1984). Asymptotic normality and the bootstrap in stratified sampling. *Ann. Statist.* **12** 470–482.

BRICKELL, E. F. (1984). A few results in message authentication. *Congressus Numerantium* **43** 141–154.

DIPPO, C. S. (1981). Variance estimation for nonlinear estimators based upon stratified samples from finite populations. Ph.D. dissertation, Dept. Statistics, George Washington Univ.

FRANCISCO, C. A. and FULLER, W. A. (1991). Quantile estimation with a complex survey design. *Ann. Statist.* **19** 454–469.

GHOSH, J. K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Ann. Math. Statist.* **42** 1957–1961.

GURNEY, M. and JEWETT, R. S. (1975). Constructing orthogonal replications for variance estimation. *J. Amer. Statist. Assoc.* **70** 819–821.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

KOVAR, J., RAO, J. N. K. and WU, C. F. J. (1988). Bootstrap and other methods to measure errors in survey estimates. *Canad. J. Statist. Suppl.* **16** 25–45.

KREWSKI, D. and RAO, J. N. K. (1981). Inference from stratified samples: Properties of the linearization jackknife and balanced repeated replication methods. *Ann. Statist.* **9** 1010–1019.

MCCARTHY, P. J. (1969). Pseudoreplication: Half-samples. *Review of the International Statistical Institute* **37** 239–264.

RAO, J. N. K. and WU, C. F. J. (1985). Inference from stratified samples: Second order analysis of three methods for nonlinear statistics. *J. Amer. Statist. Assoc.* **80** 620–630.

RAO, J. N. K. and WU, C. F. J. (1987). Methods for standard errors and confidence intervals from sample survey data: Some recent work. *Bull. Inst. Internat. Statist.* **3** 5–19.

RAO, J. N. K. and WU, C. F. J. (1988). Resampling inference with complex survey data. *J. Amer. Statist. Assoc.* **83** 231–241.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

SHAO. J. (1989). The efficiency and consistency of approximations to the jackknife variance estimators. *J. Amer. Statist. Assoc.* **84** 114–119.

SHAO, J. and WU, C. F. J. (1989). A general theory for jackknife variance estimation. *Ann. Statist.* **17** 1176–1197.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SITTER, R (1992a). A resampling procedure for complex survey data. *J. Amer. Statist. Assoc.* To appear.

SITTER, R. (1992b). Comparing three bootstrap methods for survey data. *Canad. J. Statist.* To appear.

WOLTER, K. M. (1985). *Introduction to Variance Estimation*. Springer, New York.

WOODRUFF, R. S. (1952). Confidence intervals for medians and other position measures. *J. Amer. Statist. Assoc.* **47** 635–646.

WU, C. F. J. (1991). Balanced repeated replications based on mixed orthogonal arrays. *Biometrika* **78** 181–188.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OTTAWA
OTTAWA, ONTARIO
CANADA K1N 6N5

DEPARTMENT OF STATISTICS
AND ACTUARIAL SCIENCE
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
CANADA N2L 3G1