SSC Annual Meeting, July 1995
Proceedings of the Survey Methods Section

A COMPOSITE ESTIMATOR FOR PROVINCIAL UNDERCOVERAGE IN THE CANADIAN CENSUS

L.-P. Rivest[1]

ABSTRACT

A composite estimator for provincial undercoverage in the Canadian census is proposed. This estimator is shown to be optimal for both provincial totals and provincial shares. The efficiency of the proposed composite estimator with respect to a full adjustment is estimated under various bias scenarios.

RÉSUMÉ

Un estimateur composite des taux de sous-dénombrement provinciaux dans le recensement canadien est suggéré. On montre que cet estimateur est optimal pour la répartition et les totaux provinciaux. L'efficacité de l'estimateur est calculée sous différents scénarios de biais.

1. INTRODUCTION

The census does not enumerate all the inhabitants that should fill a census form on Census Day. In the 1991 Canadian census, it is estimated that about 3% of the population failed to properly register. Thus the census needs to be adjusted for undercoverage in order to properly represent the demographic picture of the country on Census day. In 1991, several surveys were used in order to adjust the census (Germain and Julien, 1995). The undercount was measured using the Reverse Record Check (RRC); surveys were also carried out to estimate overcoverage. The net undercounts were obtained by subtracting the overcount estimates from the RRC undercount estimates. In 1991, the overcount estimates were relatively small so that most of the net undercoverage came from the RRC.

The RRC is a peculiar survey (Burgess, 1988). Its frames aim to cover all the Canadian residents to be counted in 1991. This includes people counted in the 1986 Census, people missed in 1986, the newborn for 1986-1991, the immigrants that entered the country between '86 and '91 and the non-permanent residents. (These frames do not cover all the persons who should be counted in the 1991 census. They miss the Canadian residents who were out of the country in 1986 and who are back in 1991). Samples are taken from each of these frames. The whereabouts of a sampled person are known only up to its registration into the administrative file from which it has been sampled. Thus, to determine whether a sampled person has been enumerated, the person has to be traced up to 1991. This is done in several steps. First, there is a matching of each sampled person to the Tax Data file in order to obtain a more recent address. Second, the Regional Office staff must contact each sampled person using the administrative address provided. If the sampled person cannot be contacted with the information provided, the staff must trace the person using publicly available data sources and information from relatives and former neighbours. Once contacted, the sampled person completes a questionnaire on which the Census day address is reported as well as demographic information, like date of birth and sex, on him/herself and on other members of the household. Third, at Head Office, the questionnaire corresponding to the Census day address is searched to classify the sampled person as counted or not.

The sampled persons have to be traced on relatively long periods. This causes some difficulties. First, the non-response is higher than the undercount; in 1991 the non-response is estimated at 4.8% (Germain and Julien, 1995). The adjustment for non-response is an important component of the estimated undercount. In 1991, it led to an implied undercoverage rate of 12.2% for non-respondents; it increased the national undercoverage rate by 0.33% (Germain and Julien, 1995). Second, a sampled person who has been counted in 1991 is well classified by the RRC if the administrative file matching generates its census address. Failures to find census addresses will lead to bias. The potential importance of this **tracing bias** can be ascertained by comparing the preliminary results for the 1991 RRC published in November 1992 to the final results of May 1993. Using additional administrative files to generate new addresses for the RRC searches led to reclassifying several sampled persons from missed to enumerated. This revision

[1] Louis-Paul Rivest, Département de mathématiques et de statistique, Université Laval, Ste-Foy, Québec, G1K 7P4.

resulted in a 10% drop of the national undercoverage rate. These features of the RRC should be borne in mind when selecting an estimator for adjusting the census for undercoverage.

The selection of an estimator for undercoverage is based on statistical criteria. Fellegi (1980) suggested to adjust for undercoverage only when adjusting reduces the mean squared error of the estimator. Royce (1992) developed a preliminary test estimator that adjusts the census according to the results of the coverage studies only when a statistic measuring the improvement in mean squared error is large enough. He also considered composite estimators shrinking the net undercount estimates towards 0.

"Compositing" is a method to improve the accuracy of an estimator by shrinking it towards some suitable statistics. When making the decision to adjust the census for undercoverage, shrinking towards 0 is the option to consider since it is always possible to improve on preliminary test estimators such as the one considered by Royce by "compositing". For the 1991 census, most provincial undercoverage estimates are much larger than their standard errors, therefore shrinking towards 0 does not yield much improvement. On the other hand, since the discrepancies between the provincial undercount rates are small, shrinking the provincial undercoverage estimates towards the national estimate might be appropriate. A composite estimator implementing this strategy is proposed in this paper. Its properties for estimating population shares and population totals are investigated. Efficiency calculations for comparing the proposed composite estimator to a full adjustment based on RRC results are presented. The impact of various bias scenarios on the efficiency is also given.

2. A COMPOSITE ESTIMATOR SHRINKING TOWARDS THE NATIONAL UNDERCOVERAGE RATE

This section presents a new composite estimator and derives the optimal shrinking factor for estimating provincial totals and provincial shares. Following Royce (1992), define:

- $Y = (Y_1, Y_2, \ldots Y_{10})$ as the vector of the provincial census counts;
- $\hat{U}$ as the vector of the provincial net undercount estimates;
- p as the vector of provincial shares, *i.e.*, $p_i = Y_i / (Y_1 + \ldots + Y_{10})$.

The composite estimator under study is defined by

$$\hat{V} = (1-\alpha)\ \hat{U}'1p + \alpha\hat{U}$$

where 1 is a 10×1 vector of 1's. It gives full undercount adjustments when α is equal to 1 while when $\alpha=0$, it yields synthetic adjustments equal to the provincial shares of the national undercoverage.

To estimate α and to evaluate the precision of $\hat{V}$, following Royce (1992), we suggest using the weighted mean squared errors (WMSE) defined by:

$$WMSE(g(Y+\hat{V})) = \sum_{i=1}^{10} w_i (g_i(Y+\hat{V}) - g_i(T))^2 \quad (1)$$

where:

- T is the vector of the unknown true provincial totals;
- $g_i(T)$ is a function of the provincial totals; this work focusses on provincial totals $(g_i(T)=T_i)$ and provincial shares $(g_i(T)=T_i/(T_1+\ldots+T_{10}))$;
- w_i is the weight of the ith province, which as in Royce (1992) is set equal to $1/p_i$.

Expanding (1) into a Taylor series eases the calculations. This yields:

$$WMSE(g(Y+\hat{V})) \approx (\hat{V}-U)'\Omega(\hat{V}-U)$$

where U is the unknown vector of the true provincial undercounts and Ω is a 10×10 matrix depending on the functions g_i. For provincial totals, $\Omega=\text{diag}(1/p_i)$ (this is a 10x10 diagonal matrix whose (i,i) term is given by $1/p_i$) while for provincial shares, up to a proportionality constant, $\Omega=\text{diag}(1/p_i)-11'$ (see Table 1 in Royce, 1992). Efficiency calculations are based on the expectation with respect to the distribution of $\hat{U}$ of the approximate WMSE,

$$EWMSE(g(Y+\hat{V})) = E[(\hat{V}-U)'\Omega(\hat{V}-U)].$$

The distribution of $\hat{U}$ is taken to be $N_{10}(U+b_u, \text{diag}(\sigma_i^2))$ where b_u represents the vector of the undercount biases, $N_{10}(\mu, \text{diag}(\sigma_i^2))$ stands for the 10-dimensional normal distribution with mean vector μ and variance-covariance matrix $\text{diag}(\sigma_i^2)$, and σ_i^2 stands for the RRC variance of the undercount estimate in province i. In this paper the σ_i^2 are assumed to be known.

The next proposition presents the derivation of the optimal shrinking factors for provincial shares and provincial totals.

Proposition 1

For the composite estimator $\hat{V}$, the expected weighted mean squared errors for totals and for shares are given by

$$EWMSE_{totals}(g(Y+\hat{V})) = \alpha^2\sum(1/p_i - 1)\sigma_i^2$$

$$+(\alpha b_u - (1-\alpha)U)'\Omega(\alpha b_u - (1-\alpha)U) + \sum(\sigma_i^2 + b_i^2)$$

$$EWMSE_{\text{shares}}(g(Y+\hat{V})) = \alpha^2 \sum (1/p_i - 1)\sigma_i^2$$
$$+(\alpha b_u - (1-\alpha)U)'\Omega\,(\alpha b_u - (1-\alpha)U),$$

where $\Omega = \text{diag}(1/p_i) - \mathbf{1}\mathbf{1}'$.

The shrinking factor minimizing the expected mean squared error for both totals and shares is given by:

$$\alpha = \frac{U'\Omega(U+b_u)}{\sum (1/p_i - 1)\sigma_i^2 + (U+b_u)'\Omega(U+b_u)}.$$

Proof: Let tr(.) denote the trace operator. The general form for the expected weighted mean squared error is

$$EWMSE(g(Y+\hat{V})) = E(([(1-\alpha)\hat{U}'\mathbf{1}p$$
$$+\alpha\hat{U} - U]'\Omega\,[(1-\alpha)\hat{U}'\mathbf{1}p + \alpha\hat{U} - U]))$$

For provincial shares, $\Omega = \text{diag}(1/p_i) - \mathbf{1}\mathbf{1}'$, thus $\Omega p = 0$ and

$$EWMSE(g(Y+\hat{V})) = EWMSE(g(Y+\alpha\hat{U}))$$
$$= tr(\Omega E((\alpha\hat{U} - U)(\alpha\hat{U} - U)'))$$
$$= tr(\Omega\,[Var(\alpha\,\hat{U}) + E(\alpha\,\hat{U} - U)E'(\alpha\,\hat{U} - U)])$$
$$= \alpha^2 \sum (1/p_i - 1)\sigma_i^2 + (\alpha\, b_u - (1-\alpha)\,U)'\Omega(\alpha b_u - (1-\alpha)U).$$

For totals, the expected weighted mean squared error is given by:

$$E([(1-\alpha)\hat{U}'\mathbf{1}p + \alpha\hat{U} - U]'(\text{diag}(1/p_i) - \mathbf{1}\mathbf{1}' + \mathbf{1}\mathbf{1}')$$
$$[(1-\alpha)\hat{U}'\mathbf{1}p + \alpha\hat{U} - U])$$
$$= \alpha^2 \sum (1/p_i - 1)\sigma_i^2 + (\alpha b_u + (1-\alpha)U)'$$
$$\Omega(\alpha b_u + (1-\alpha)U) + E((\sum \hat{U}_i - \sum U_i)^2).$$

Since the expected weighted mean squared errors are quadratic functions of α, the optimal value of the shrinking parameter is easily found by setting the derivatives with respect to α of the EWMSE's equal to 0.

Q.E.D.

The above formula for the optimal value of α is the same as that derived by Royce (1992) for estimating shares with composite estimator $\alpha\hat{U}$. In the presence of a known bias b_u, a plug-in estimator for α is given by

$$\hat{\alpha} = \frac{(\hat{U} - b_u)'\Omega\,\hat{U}}{\sum_{i=1}^{10} (1/p_i - 1)\sigma_i^2 + \hat{U}'\Omega\hat{U}}. \qquad (2)$$

3. EFFICIENCY COMPARISONS

This section derives an estimate of the expected weighted mean squared error of composite estimator $\hat{V}$, when the shrinking parameter is estimated by (2), and when there is no bias, *i.e.*, when $b_u = 0$. Proposition 1 shows that the outcome of this investigation should be similar for both totals and shares. Therefore our analysis focusses on shares. In order to simplify the calculations it is convenient to let A be a 10x9 matrix such that $AA' = \Omega$ (this is possible since the rank of Ω is 9). Let $\hat{W} = A'\hat{U}$ and $\Sigma_W = \text{Var}(\hat{W}) = A'\,\text{diag}(\sigma_i^2)A$; note that Σ_W is a 9x9 positive definite matrix. With this notation, the shrinking factor estimated in the absence of bias is given by

$$\hat{\alpha} = 1 - \frac{tr(\Sigma_W)}{\hat{W}'\hat{W} + tr(\Sigma_W)}.$$

Since there is no bias, $E(\hat{W}) = A'E(\hat{U}) = A'U = W$, say. This section compares the expected weighted mean squared error of the composite estimator with that of the estimator making a full adjustment for undercoverage (*i.e.*, with $\alpha = 1$).

The expected weighted mean squared error of the full adjustment estimator is given by $EWMSE(g(Y+\hat{U})) = tr(\Sigma_W)$ while that of the composite estimator with estimating shrinkage parameter is equal to:

$$\text{EWMSE}(g(Y+\hat{V})) = \text{EWMSE}(g(Y+\hat{\alpha}\hat{U}))$$
$$= E((\hat{\alpha}\hat{W} - W)'(\hat{\alpha}\hat{W} - W)).$$

The aim of this section is to find an approximately unbiased estimator of this quantity. Observe that

$$(\hat{\alpha}\hat{W} - W)'(\hat{\alpha}\hat{W} - W) = (\hat{\alpha} - 1)^2\hat{W}'\hat{W}$$
$$+2(\hat{\alpha} - 1)\hat{W}'(\hat{W} - W) + (\hat{W}'\hat{W} - W'W).$$

Straightforward manipulations allow to rewrite the expected weighted mean squared error of $\hat{V}$ as:

$$\text{EWMSE}(g(Y+\hat{V})) = tr(\Sigma_W)$$
$$\{1 + E[\hat{\alpha}(1-\hat{\alpha})] - 2E[\frac{\hat{W}'(\hat{W} - W)}{\hat{W}'\hat{W} + tr(\Sigma_W)}]\} \qquad (3)$$

Since $\hat{\alpha}(1-\hat{\alpha})$ is an unbiased estimator of its expectation, only the estimation of the last term of the right hand side of the above expression needs special attention. An estimator for this term is provided in the next proposition which is proved in the appendix.

Proposition 2
Under the assumption that $\hat{W}$ is distributed a $N_9(W, \Sigma_W)$ random vector,

$$E\{\frac{\hat{W}'(\hat{W}-W)}{\hat{W}'\hat{W}+tr(\Sigma_W)}-\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}+2\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}-(\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}})^2-4\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}\}$$

is $o(tr(\Sigma_W)/W'W)^2$.

Replacing, in formula (3), $\hat{W}'(\hat{W}-W)/[\hat{W}'\hat{W}+tr(\Sigma_W)]$ by the approximation derived in Proposition 2 yields an approximately unbiased estimator of the expected weighted mean squared error of the composite estimator. A simple expression for this estimator can be derived by noting that

$$\hat{\alpha}(1-\alpha)=\frac{\hat{W}'\hat{W}tr(\Sigma_W)}{[\hat{W}'\hat{W}+tr(\Sigma_W)]^2}$$

$$=\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}-2(\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}})^2+o_p(\frac{tr(\Sigma_W)}{(W'W)})^2.$$

This leads to the following estimator of the expected weighted mean squared error of the composite estimator,

$$tr(\Sigma_W)\{1-\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}+4\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}-8\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}\}. \quad (4)$$

The efficiency of the composite estimator compared to a full adjustment is, in the absence of bias, easily obtained from formula (4):

$$\{1-\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}+4\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}-8\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}\}^{-1}.$$

The components in (4) have interesting interpretations. First $8tr(\Sigma_W)\hat{W}'\Sigma_W\hat{W}/(\hat{W}'\hat{W})^3$ is $O_p([tr(\Sigma_W)/W'W]^2)$. Removing this term gives a first order approximation to the expected weighted mean squared error. Second $4\hat{W}'\Sigma_W\hat{W}/(\hat{W}'\hat{W})^2$ is related to the cost for estimating α. Finally, if α was known, then $tr(\Sigma_W)[1-tr(\Sigma_W)/\hat{W}'\hat{W}]$ would estimate the expected mean squared error of the composite estimator.

One can rewrite $4\hat{W}'\Sigma_W\hat{W}/(\hat{W}'\hat{W}^2)$ as a function of p_i, $\hat{U}_i$ and σ_i^2:

$$4\frac{\hat{W}'\Sigma_W\hat{W}}{\hat{W}'\hat{W}^2}=\frac{\Sigma_i(\hat{U}_i/p_i-\Sigma_k\hat{U}_k)^2\sigma_i^2}{(\Sigma_i\hat{U}_i^2/p_i-(\Sigma\hat{U}_k)^2)^2}.$$

The numerator of the above fraction is proportional to $\Sigma_i(\hat{r}_i-\hat{r})^2\sigma_i^2$ where $\hat{r}_i$ is the undercoverage rate in the *ith* province and $\hat{r}$ is the national rate. This sheds some light on the effect of estimating α. This estimation increases the expected mean squared error when an undercoverage rate different from the national rate occurs in a province with a large σ_i^2. It is clear that σ_i^2 can be large only in a large province. Thus estimating α is costly when the undercoverage rate is, in some large province, markedly different from the national rate.

In 1991, the undercoverage estimate in Ontario was 3.64%. It was larger than the national undercoverage rate of 2.84%. This showed up on the efficiency of the composite estimator since it dropped from 1.17 to 1.03 when accounting for the estimation of α (this is based on the following 1991 results: $4\hat{W}'\Sigma_W\hat{W}/(\hat{W}'\hat{W}^2)=.17$ and $tr(\Sigma_W)/\hat{W}'\hat{W}=.14$).

4. EFFICIENCY COMPARISONS IN THE PRESENCE OF BIAS

Suppose now that the undercount estimates are biased, *i.e.*, $E(\hat{U})=U+b_u$. In the w-notation $E(\hat{W})=W+b$ where $b=A'b_u$. Let EWMSE$(g(Y+\hat{V});b_u)$ denote the expected weighted mean squared error of the composite estimator in the presence of bias b_u. The aim of this section is to derive an approximately unbiased estimator of EWMSE$(g(Y+\hat{V});b_u)$ for estimating shares. As in Proposition 1,

$$\begin{aligned}\text{EWMSE}(g(Y+\hat{V});b_u)&=E(\hat{\alpha}(\hat{W}-W)'(\hat{\alpha}\hat{W}-W))\\&=E((\hat{\alpha}\hat{W}-W-b)'(\hat{\alpha}\hat{W}-W-b))\\&\quad+2b'E(\hat{\alpha}\hat{W}-W-b)+b'b).\end{aligned}$$

The first term of the right hand side represents the EWMSE in the absence of bias. It can be estimated by (4). Since $E(\hat{W})=W+b$ an unbiased estimator of the second term is $-2b'(1-\hat{\alpha})\hat{W}$. This leads to the following estimator of the expected weighted mean squared error of the composite estimator in the presence of bias,

$$tr(\Sigma_W)\{1-\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}+4\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}-8\frac{tr(\Sigma_W)}{\hat{W}'\hat{W}}\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}-\frac{2b'\hat{W}}{tr(\Sigma_W)+\hat{W}'\hat{W}}\}+b'b.$$

This has to be compared to $tr(\Sigma_W)+b'b$, the EWMSE obtained with a full adjustment. When the components of bias b are large, $b'b$ is the dominant term in the two EWMSE's and the two estimators have approximately the same EWMSE. The bias has an impact on the difference between the two EWMSE estimates only when $b'W$ is non null. It is easy to show that $b'\hat{W}$ is proportional to $\Sigma_i(\hat{r}_i-\hat{r})b_{ui}$, where b_{ui} represents the *ith* component of the bias vector b_u. When bias is positively correlated with undercoverage (*i.e.*, when the raw bias is large in provinces with a large undercount rate) $b'\hat{W}$ is positive and the composite estimator gains in efficiency when compared to a full adjustment.

Given the importance of the changes between the preliminary and the final RRC estimates, one can conjecture that, if additional administrative files had been available for carrying out the RRC searches, further reductions in undercoverage would have been obtained. A likely bias scenario is therefore one that is proportional to the differences between the preliminary RRC results of November 92 from the final results of May 93. Under this scenario $b'\hat{W}$ is positive since Ontario has the largest bias and the largest undercoverage rate. Thus, under this bias scenario, the efficiency of the composite estimator is larger than the 1.03 calculated in Section 3.

Acknowledgements: I am grateful to Normand Thibault and Robert Courtemanche for carrying out Monte Carlo simulation studies of the composite estimator that led to the second order calculations presented in Section 3 and to Don Royce for stimulating discussions.

APPENDIX: PROOF OF PROPOSITION 2

Write $\hat{W}=W+z$ where z is a $N_9(0,\Sigma_W)$ random vector. Then

$$E[\frac{\hat{W}'(\hat{W}-W)}{\hat{W}'\hat{W}+tr(\Sigma_W)}] =$$

$$E[\frac{W'z+z'z}{W'W+2W'z+z'z+tr(\Sigma_W)}].$$

As $tr(\Sigma_W)/W'W$ goes to 0, the above function can be expanded in a Taylor series,

$$(\frac{W'z}{W'W}+\frac{z'z}{W'W})(1-2\frac{W'z}{W'W}-\frac{z'z}{W'W}-\frac{tr(\Sigma_W)}{W'W}$$

$$+4(\frac{W'z}{W'W})^2+(\frac{z'z}{W'W})^2+4\frac{W'z\,z'z}{W'WW'W}$$

$$+4\frac{W'z\,tr(\Sigma_w)}{W'WW'W}+2\frac{z'z\,tr(\Sigma_w)}{W'WW'W}-8(\frac{W'z}{W'W})^3)$$

$$+o_p((\frac{tr(\Sigma_W)}{W'W})^2).$$

To calculate the expectation of the above quantity, the following results concerning high order moments of a normal random vector are useful:

i) $E((W'z)^2)=W'\Sigma_W W$

ii) $E(z'z)=tr(\Sigma_W)$,

iii) $E((z'z)^2)=tr(\Sigma_W)^2+2tr(\Sigma_W^2)$

iv) $E((W'z)^4)=3(W'\Sigma_W W')^2$

v) $E((W'z)^2z'z)=W'\Sigma_W W'tr(\Sigma_W)+2W'\Sigma_W^2 W$

vi) $E((W'z)z'z)=E(W'z)=E((W'z)^3z'z)=0$

Thus,

$$E[\frac{\hat{W}'(\hat{W}-W)}{\hat{W}'\hat{W}+tr(\Sigma_W)}]=\frac{tr(\Sigma_W)}{W'W}-2\frac{W'\Sigma_W W}{(W'W)^2}$$

$$-2\frac{tr(\Sigma_W)^2+tr(\Sigma_W^2)}{(W'W)^2}+12\frac{(W'\Sigma_W W)}{(W'W)^2}\frac{tr(\Sigma_W)}{W'W}$$

$$+16\frac{2W'\Sigma_W^2 W}{(W'W)^3}-24\frac{(W'\Sigma_W W)^2}{(W'W)^4}+o(\frac{tr(\Sigma_W)}{W'W})^2.$$

In a similar way one shows that

$$E[\frac{tr(\Sigma_w)}{\hat{W}'\hat{W}}]$$

$$=E\frac{tr(\Sigma_W)}{W'W+2W'z+z'z}]$$

$$=\frac{tr(\Sigma_W)}{W'W}[1-\frac{tr(\Sigma_W)}{W'W}+4\frac{W'\Sigma_W W}{(W'W)^2}]+o(\frac{tr(\Sigma_W)}{W'W})^2.$$

Furthermore, since

$$\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2}=\frac{W'\Sigma_W W}{(W'W)^2}\left[1+2\frac{W'\Sigma_W z}{W'\Sigma_W W}+\frac{z'\Sigma_W z}{W'\Sigma_W W}\right]$$

$$\left[1-4(\frac{W'z}{W'W})-2\frac{z'z}{W'W}+12(\frac{(W'z)}{W'W})^2+o(\frac{tr(\Sigma_W)}{W'W})\right],$$

the following approximation is easily derived:

$$E(\frac{\hat{W}'\Sigma_W\hat{W}}{(\hat{W}'\hat{W})^2})=\frac{W'\Sigma_W W}{(W'W)^2}$$

$$(1-8\frac{W'\Sigma_W^2 W}{W'\Sigma_W WW'W}-2\frac{tr(\Sigma_w)}{W'W}$$

$$+12\frac{W'\Sigma_W W}{(W'W)^2}+\frac{tr(\Sigma_W^2)}{W'\Sigma_W W}+o(\frac{tr(\Sigma_W)}{W'W})^2.$$

Replacing the expectations considered in Proposition 2 by the above approximations and estimating the expectations of $O(tr(\Sigma_W)/W'W)^2$ terms by substituting W for $\hat{W}$ prove the result.

REFERENCES

Germain, M.-F., and Julien, C. (1995). "Results of the 1991 census coverage error measurement program", Statistics Canada Paper.

Burgess, R.D. (1988). "Evaluation of reverse record check estimates of undercoverage in the Canadian Census of Population", *Survey Methodology*, 14, 137-156.

Fellegi, I.P. (1980). "Should the census count be adjusted for allocation purposes?", *Proceedings of the 1980 Conference on Census Undercount*, U.S. Bureau of the Census.

Royce, D. (1992). "A comparison of some estimators of a set of population totals", *Survey Methodology*, 18, 109-125.