The Effect of Clustering on the Precision of Estimation

A thesis presented in partial
fulfilment of the requirements

for the degree

of Master of Business Studies
in Marketing at
Massey University

Zhengping Guan

1997

ABSTRACT

The effect of clustering interval on design effect may be important in selection of alternative sampling designs by evaluating the cost-efficiency in the context of face-to-face interview surveys. There has been little work in investigating this effect in New Zealand. This study attempts to investigate this effect by using data from a two-stage sampling face-to-face interview survey. Seventeen stimulated samples are generated. A simple method, $design\ \ effect = \frac{ms_b}{ms}$, is developed to estimate design effects for 81 variables for both the simulated samples and the original sample. These estimated design effects are used to investigate the effect of clustering interval. This study also investigates the effect of cluster size. The results indicate that clustering interval has little influence on design effect but cluster size substantial influence. The evaluation of the cost-efficiency in alternative clustering intervals is discussed. As an improvement in the efficiency of a sample design by an increase in clustering interval can not be justified by the increase in cost, it seems that the sample design with the smallest clustering interval is the best. An alternative method $design\ \ effect \approx mr^2$ is also discussed and tested in estimating design effects. The result indicates that the applicability of $design\ \ effect \approx mr^2$ is the same as that of $design\ \ effect = \frac{ms_b}{ms}$.

ACKNOWLEDGMENT

I would like to thank ACNielsen McNair for providing data from a face-to-face interview survey.

Thanks are also due to Mr Nick Jones, Managing Director of ACNielsen McNair, for his helpfulness and kind cooperation, and due to Mr James Reilly for his kind assistance in preparing the data.

CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

1. INTRODUCTION

Surveys using clustered multi-stage sampling designs are common in research in business and other social sciences. For a given sample size, these sampling designs may reduce the cost of data collection. However, such designs lead to increase in the sampling variances of estimates.

This study investigates the way in which final stage clustering affects sampling variances in face-to-face interview surveys.

In view of the need to make an adjustment to a sampling variance estimate from a complex sample design, Kish (1965) proposed a measurement which he called "design effect" to describe the sampling variance increase due to the complex sample design. He held the position that sample designs affect variance estimation and statistical analysis. However, Skinner, Holt & Smith (1989 chapter 2) argued that it was population structure rather than sample designs that affected variance estimation and statistical analysis. These two positions are often consistent. For a given sample design, population structure may affect variance estimation and statistical analysis, and vice versa.

Skinner et al (1989, p 24) also proposed an alternative measurement which they called "misspecification effect" instead of design effect. That is, the measurement of sample design efficiency is sampling variance of the actual sample design over the expected value of sampling variance of a simple random sample with the same size, rather than sampling variance of the actual sample design over sampling variance of a simple random sample with the same size. However, it is difficult in practice to obtain the expected value of a sampling variance estimate. Thus, design effect is likely to be more applicable in measuring the efficiency of sample designs than misspecification

effect.

Sampling variance increase due to clustering in surveys is caused by similarity of elements within clusters. This similarity is measured by the homogeneity of within-cluster elements.

There is a voluminous body of literature concerning complex sample design, variance estimation, design effect and homogeneity. However, there has been little research into the relation between design effect and intervals of selecting elements within clusters in New Zealand. The need to evaluate the cost-efficiency of the alternative sample designs with different clustering intervals requires to conduct an investigation into the effect of clustering interval on design effect.

Data for this study is from a face-to-face interview survey conducted by ACNielsen-McNair. This is a two-stage sample (see Chapter 4 for specification of the sample). A number of simulated samples are drawn from it to investigate the effect of clustering interval (see Chapter 4 for the detailed discussion in generating simulated samples).

Based on the design effects estimated from both the original sample and the simulated samples, this study investigates the following:

a. *The relation between design effect and clustering interval;*

b. *The relation between design effect and cluster size;*

c. *The applicability of the formula:*

$design\ \ effect \approx mr^2$

(see Chapter 4 for both specification and derivation of this formula);

d. *The effect of clustering interval on cost-efficiency of alternative sample designs.*

The results for both *a* and *b* should be that design effect decreases with either increase in clustering interval or decrease in cluster size. The result for *c* should justify the alternative estimation method for design effect. The result for *d* should provide the guideline for selection of the alternative sample designs with different clustering intervals.

2. METHODS OF ESTIMATING SAMPLING VARIANCES

Methods of estimating sampling variance of complex sample designs can be categorized into **standard (mathematical) methods**, **subsampling methods**, and **modelling methods.** This chapter reviews the construction of these methods in the literature.

2.1 Standard (Mathematical) Methods

These methods have been developed by mathematical derivation in obtaining the expected values of estimates' variances. Such derivation is based on probability theory. These methods have been discussed in the standard sampling theory textbooks, for examples, Hansen, Hurwitz & Madow (1953), Kish (1965) and Cochran (1963 & 1977).

Notation:

K is number of clusters in the population;
k is number of clusters in the sample;
M is population of a cluster;
m is number of sampled elements in a sampled cluster;
N = KM is number of elements in the population;
n = km is number of elements in the sample;
Y_{ji} is the value of the ith element in the jth cluster;
μ_j is mean of the elements within the jth cluster in the population;
$\overline{Y}_j = \sum_{i=1}^{m} \frac{Y_{ji}}{m}$ is sample mean within the jth cluster;
μ_Y is mean of the population;
$\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{m} Y_{ji}}{km} = \frac{\sum_{j=1}^{k} \overline{Y}_j}{k}$ is element mean for the

sample;

$\sigma_Y^2 = \frac{\sum_{i=1}^{N}(Y_i - \mu_Y)^2}{N-1} = \frac{\sum_{j=1}^{K}\sum_{i=1}^{M}(Y_{ji} - \mu_Y)^2}{KM-1}$ is the population variance;

$\sigma_B^2 = \frac{\sum_{j=1}^{K}(\mu_j - \mu_Y)^2}{K-1}$ is between-cluster variance in the population;

$\sigma_W^2 = \frac{\sum_{j=1}^{K}\sum_{i=1}^{M}(Y_{ji} - \mu_j)^2}{N-K} = \frac{\sum_{j=1}^{K}\sum_{i=1}^{M}(Y_{ji} - \mu_j)^2}{K(M-1)}$ is within-cluster variance in the population;

$SS = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y})^2$ is total sum of squares in the sample;

$SS_B = \sum_{j=1}^{k} m(\bar{Y}_j - \bar{Y})^2$ is sum of squares between clusters in the sample;

$SS_W = \sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y}_j)^2$ is sum of squares within clusters in the sample;

$S^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y})^2}{km-1} = \frac{SS}{km-1}$ is the sample variance (also called mean squares in the sample);

$S_B^2 = \frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k-1} = \frac{SS_B}{m(k-1)}$ is between-cluster variance in the sample or variance among the means of sampling units selected in the first stage of two-stage sampling;

$S_W^2 = \frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y}_j)^2}{k(m-1)} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y}_j)^2}{n-k} = \frac{SS_W}{k(m-1)}$ is within-cluster variance in the sample (also called

mean squares within clusters in the sample) or variance in the second stage of two-stage sampling.

One common standard method is to estimate sampling variance of an estimate in a two-stage sample with equal subsample size and epsm (equal probability selection method). Sampling variance is contributed by both the first stage and the second stage of the sample. That is,

$$\sigma_{\bar{y}}^2 = (1-f_1)\frac{\sigma_B^2}{k} + (1-f_2)\frac{\sigma_W^2}{km} \qquad 2.0$$

where:

$f_1 = \frac{k}{K}$ *is sampling fraction at the first stage;*

$f_2 = \frac{m}{M}$ *is sampling fraction at the second stage;*

$1-f_1$ *and* $1-f_2$ *are finite population corrections, corresponding to the first-stage and the second-stage respectively.*

This formula is mathematically derived by considering the expected values of estimates for a parameter from both the first stage and the second stage of the sample. The detailed derivation is in Cochran (1977, chapter 10).

However, as it is difficult in practice to obtain both between-cluster variance in the population and within-cluster variance in the population, these two variances can be estimated by between-cluster variance in the sample and within-cluster variance in the sample respectively. That is, σ_B^2 can be estimated by S_B^2 and σ_W^2 by S_W^2. As the second stage sampling is made within the clusters selected by the first stage sampling, both σ_B^2 and σ_W^2 can not directly be replaced by S_B^2 and S_W^2 respectively in the

formula. The sampling fraction at the first stage f_1 must be taken into account in estimating the sampling variance at the second stage (i.e., within-cluster sampling variance).

Thus, sampling variance in a two-stage sample can be estimated by:

$$\hat{\sigma}^2_{\bar{Y}} = (1-f_1)\frac{S_B^2}{k} + f_1(1-f_2)\frac{S_W^2}{km}. \qquad 2.1$$

This formula is mathematically proved by seeking the expected values of the estimate variances of a parameter from both the first stage and the second stage of the sample. The proof is in Cochran (1977, p 278).

As $f_1 \rightarrow 0$, the second term in this formula is negligible and $1-f_1 \rightarrow 1$. This estimator can therefore be replaced by:

$$\hat{\sigma}^2_{\bar{Y}} = \frac{S_B^2}{k} = \frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k(k-1)}. \qquad 2.2$$

Sudman (1976) suggested that finite population correction could be taken as 1 if sampling fraction was less than 0.02. This estimator may overestimate sampling variance even if f_1 is not very small. Such overestimation leads to a conservative sampling variance estimate.

Formula 2.2 will produce conservative sampling variance estimates if $f_1\frac{S_B^2}{k}$ is larger than $f_1(1-f_2)\frac{S_W^2}{km}$. That is, that $f_1\frac{S_B^2}{k}$ is larger than $f_1(1-f_2)\frac{S_W^2}{km}$ implies that the estimates of formula 2.2 are larger than those of formula 2.1.

If $\frac{S_B^2}{k} > \frac{S^2}{n}$, that is, the sampling variance contributed by the first stage sampling is larger than the entire sample sampling variance estimated by using the method of simple random samples, $f_1 \frac{S_B^2}{k}$ will be larger than $f_1(1-f_2)\frac{S_W^2}{km}$, and formula 2.2 will produce conservative estimates.

Proof:

$$f_1 \frac{S_B^2}{k} - f_1(1-f_2)\frac{S_W^2}{km} > 0 \rightarrow$$

$$\frac{S_B^2}{k} - (1-f_2)\frac{S_W^2}{km} > 0 \rightarrow$$

$$S_B^2 - (1-f_2)\frac{S_W^2}{m} > 0 \rightarrow$$

$$S_B^2 - \frac{S_W^2}{m} > 0 \rightarrow$$

$$mS_B^2 > S_W^2$$

That is, if $mS_B^2 > S_W^2$, $f_1 \frac{S_B^2}{k}$ *is larger than* $f_1(1-f_2)\frac{S_W^2}{km}$.

And,

$$\because SS = SS_B + SS_W$$

$$SS_B = (k-1)mS_B^2$$

$$SS_W = k(m-1)S_W^2$$

$$\therefore mS_B^2 > S_W^2 \rightarrow$$

$$m(k-1)S_B^2 > (k-1)S_W^2 \rightarrow$$

$$m(k-1)S_B^2 = SS_B > (k-1)S_W^2 = \frac{SS_W}{k(m-1)}(k-1) = \frac{SS - SS_B}{k(m-1)}(k-1) \rightarrow$$

$$SS_B\left(1+\frac{k-1}{k(m-1)}\right) > \frac{k-1}{k(m-1)}SS \rightarrow$$

$$SS_B\frac{km-k+k-1}{k(m-1)} = \frac{km-1}{k(m-1)}SS_B > \frac{k-1}{k(m-1)}SS \rightarrow$$

$$(km-1)SS_B > (k-1)SS \rightarrow$$

$$\frac{SS_B}{k-1} > \frac{SS}{km-1} \rightarrow$$

$$\frac{SS_B}{km(k-1)} > \frac{SS}{km(km-1)}$$

$$\therefore \frac{S_B^2}{k} > \frac{S^2}{km} = \frac{S^2}{n}$$

There is often $\frac{S_B^2}{k} > \frac{S^2}{n}$ in such clustered samples as two-stage samples and one stage cluster samples.

Alsagoff, Esslemont & Gendall (1986) used formula 2.0 to estimate sampling variances for two variables (i.e., mean household income, and mean expenses on groceries). As the sample clusters (subsamples) were not of equal size, they used weights of cluster sizes to adjust sampling variance estimates.

Another method of estimating sampling variance in two-stage simple random sampling is to use the method for one-stage simple random cluster samples. That is, subsamples

from the second stage are taken as clusters. These clusters are also called *ultimate clusters* (Hansen et al 1953 & Kalton 1979). In a one-stage cluster sample, sampling variance of an estimate of mean is:

$$\sigma_{\bar{Y}}^2 \approx \frac{\sigma_Y^{\,2}}{kM}\left[1+(M-1)\rho\right] \qquad 2.3$$

where $\rho = \dfrac{2\sum_{j}^{K}\sum_{i<h}^{M}(Y_{ji}-\mu_Y)(Y_{jh}-\mu_Y)}{(M-1)(KM-1)\sigma_Y^{\,2}}$ *is the measurement of homogeneity (Cochran 1977, p 241).*

This formula is mathematically derived by deriving sampling variance from cluster totals to elements. The detailed derivation is in Kish (1965) and Cochran (1963 & 1977).

In two-stage sampling, subsamples are formed within the clusters selected in the one-stage cluster sample. These subsamples are taken as clusters. Thus, sampling variance can be estimated by:

$$\hat{\sigma}_{\bar{Y}}^2 = \frac{S^2}{km}\left[1+(m-1)\hat{\rho}\right] \qquad 2.4$$

where $\hat{\rho} = \dfrac{2\sum_{j}^{k}\sum_{i<h}^{m}(Y_{ji}-\bar{Y})(Y_{ih}-\bar{Y})}{(m-1)(km-1)S^2}$.

Formula 2.2 is equal to formula 2.4. From these two formulae, a method of estimating design effect can be constructed.

2.2 Subsampling Methods

The idea of variance estimation through replication of subsamples is to split at random a single sample into a number of subsamples. Estimates of a parameter are obtained from each subsample. It can be assumed that these estimates are from simple random samples. The variance of the entire sample estimate for the parameter can be estimated from variability between the subsample estimates.

Several variance estimation methods have been developed. These methods are **random group** (sometimes called **ultimate cluster), balanced repeated replication (BRR), jackknife, and bootstrap**. Good introductions to these methods are in Lehtonen & Pahkinen (1995), Särndal, Swensson & Wretman (1992), and Thompson (1997). Different calculation methods or different subsamplings in estimating sampling variances are used in these variance estimation methods.

In multi-stage (including two-stage) sampling, these methods are applied on the basis of the primary sampling units (PSUs) selected in the first stage of multi-stage sampling without any attention paid to the subsampling within the PSUs (Lee, Forthofer & Lorimor 1986; Rao & Wu 1988 and Sitter 1992). If PSUs are self-representing, then these PSUs are taken as strata. The sampling units selected at the first stage of subsampling within these PSUs are sampled into subsamples for variance estimation by replication methods.

2.2.1 Random Group Methods

The idea of these methods was originally suggested by Mahalanobis (1939, 1944 & 1946). Deming (1956), Hansen et al (1953), Kish (1965) and Sudman (1976) also discussed these methods. These methods split at random a single sample into a number of subsamples with the same sampling

design as that of the entire sample. Estimates of a parameter are obtained from each subsample. Then, the following formula is applied to estimate the variance of the entire sample estimate for the parameter.

$$v(\hat{\theta}) = \frac{\sum_{j=1}^{k} (\hat{\theta}_j - \hat{\theta})^2}{k(k-1)}$$

where:

$v(\hat{\theta})$ *is the sampling variance estimate of the entire sample estimate for a parameter* θ

$\hat{\theta}$ *is an estimate of* θ *from the entire sample;*

$\hat{\theta}_j$ *is an estimate of* θ *from the jth subsample;*

k *is number of subsamples;*

$\hat{\theta}$ *in the right side of the formula is sometimes replaced by* $\sum_{j=1}^{k} \frac{\hat{\theta}_j}{k}$ *(the average of the estimates from k subsamples).*

2.2.2 Balanced Repeated Replication Methods (BRR)

These methods are also called *balanced half-sample methods*. Based on Plackett & Burman (1946), McCarthy (1966 & 1969) developed and applied the idea of balanced repeated replication for a sample comprising only two sampling units from each stratum, as random group methods have poor stability of variance estimation in such sample design. Kish & Frankel (1970) applied a balanced repeated replication method for estimating sampling variance of regression coefficients.

The idea of these methods can be described as follows:

First, random selection of one sampling unit is made from a pair of sampling units in each stratum, whenever a sample is stratified or post-stratified. These sampling units selected form a subsample (i.e., a half-sample). The remainder also forms a complementary subsample. This procedure is repeated a number of times. Thus, a number of pairs of subsamples are obtained.

Second, pairs of estimates of a parameter are obtained from each pair of subsamples.

Third, the difference between a subsample estimate and either its complementary subsample estimate or the average of this pair of sample estimates is squared. Then, the average of such squared differences for all pairs of subsamples is the variance estimate desired.

This can be expressed in terms of the formula:

$$v(\hat{\theta}) = \frac{\sum_{j}^{k} (\hat{\theta}_j - \hat{\bar{\theta}}_j)^2}{k}, \quad or \quad v(\hat{\theta}) = \frac{\sum_{j}^{k} (\hat{\theta}_j - \hat{\theta}_j')^2}{4k}$$

where:

$v(\hat{\theta})$ *is the sampling variance estimate of the entire sample estimate for a parameter* θ;

$\hat{\theta}$ *is the estimate of* θ *from the entire sample;*

$\hat{\theta}_j$ *is an estimate of* θ *from the jth subsample;*

$\hat{\theta}_j'$ *is an estimate of* θ *from the jth complementary subsample;*

$\hat{\bar{\theta}}_j$ *is the average of both* $\hat{\theta}_j$ *and* $\hat{\theta}_j'$;

k is the number of pairs of subsamples.

2.2.3 Jackknife Methods

The idea of these methods was originally developed by Quenouille (1949 & 1956) for reducing the bias of an estimate, rather than estimating variance. Tukey (1958) suggested using this idea for estimating sampling variances. Durbin (1959) applied a jackknife method to variance estimation with a ratio estimator.

This idea can be described as follows:

> *First, a single sample is randomly split into k groups. The jth subsample is formed by random reduction of the jth group from these k groups.*
>
> *Second, k estimates can be obtained from these k subsamples. The variance estimate desired can be calculated by the formulae:*

$$v_1(\hat{\theta}) = \frac{\sum_j^k (\hat{\theta}_j - \hat{\bar{\theta}})^2}{k(k-1)} \quad or \quad v_2(\hat{\theta}) = \frac{\sum_j^k (\hat{\theta}_j - \hat{\theta})^2}{k(k-1)}$$

> *where:*
>
> $v_1(\hat{\theta})$ *and* $v_2(\hat{\theta})$ *are sampling variance estimates of the entire sample estimate for a parameter* θ;
>
> $\hat{\theta}$ *is the estimate of* θ *from the entire sample;*
>
> $\hat{\bar{\theta}} = \sum_{j=1}^{k} \frac{\hat{\theta}_j}{k}$;
>
> $\hat{\theta}_j$ *is the estimate of a parameter from the jth subsample;*
>
> k *is the number of subsamples.*

$v_2(\hat{\theta})$ is more conservative than $v_1(\hat{\theta})$, as $v_2(\hat{\theta}) = v_1(\hat{\theta}) + (\hat{\theta} - \hat{\bar{\theta}})^2 / (k-1)$ (Wolter 1985).

2.2.4 Bootstrap Methods

The idea of these methods was originally suggested by Efron (1979, 1981 & 1982). He used it to generate subsampling distributions as a means of obtaining approximate variance estimates and confidence intervals. Subsamples (also called bootstrap samples) are formed by drawing sampling units with replacement from a single sample and the subsample size is the same as this sample size. The same formula as used for **random group methods** is applied to calculate sampling variance for a parameter estimate with this sample design.

2.3 Modelling Methods

2.3.1 The Taylor Linearization Method

This is an old and well-known method. It is applied in the context of a nonlinear estimator of a survey statistic of interest. Typical examples of such statistics are ratios, differences of ratios, correlation coefficients and regression coefficients. The idea of this method is to approximate the nonlinear estimator by a linear function of the observations. The linear function is produced by the Taylor series expansion. Then, other variance estimation methods are used to estimate the variance of the estimator. That is, the Taylor linearization method per se does not produce a variance estimator (Wolter 1985 & Woodruff 1971).

The Taylor series expansion expresses a nonlinear function in terms of linear terms plus remainder terms. This expression is produced by programming derivatives of the nonlinear function with respect to the observations. If the nonlinear function is of order 2 continuous derivatives, the Taylor series expansion can be expressed as linear terms plus a remainder term (Fuller 1996 and Wolter 1985). That is,

$$\hat{\theta} = \theta + \sum_{i=1}^{k} \frac{\partial f(Y)}{\partial y_i}(\hat{Y}_i - Y_i) + R_n(\hat{Y}, Y),$$

where:

$R_n(\hat{Y},Y)=\sum_{i=1}^{k}\sum_{j=1}^{k}(1/2!)\frac{\partial^2 f(\ddot{Y})}{\partial y_i \partial y_j}(\hat{Y}_i-Y_i)(\hat{Y}_j-Y_j)$ *is the remainder term;*

$\hat{\theta}=f(\hat{Y})$ *is the nonlinear estimator of a parameter of interest* $\theta=f(Y)$ *at* $y=Y$;

$\hat{Y}$ *is the estimator of* Y *which is a sequence of population totals or means for k parameters* $Y_1,\ldots,Y_k$;

$\hat{Y}_1,\ldots,\hat{Y}_k$ *is the estimators of* $Y_1,\ldots,Y_k$;

$\hat{Y}$ *and* $\hat{Y}_1,\ldots,\hat{Y}_k$ *depends on sample design and sample size;*

$\hat{\theta} \rightarrow \theta$ and the remainder term $\rightarrow 0$ *as* $n \rightarrow \infty$;

y_i *and* y_j *are the observations corresponding to* $Y_1,\ldots,Y_k$;

$\ddot{Y}$ *is between* $\hat{Y}$ *and* Y.

If the nonlinear function has order s (i.e., positive integer) continuous derivatives, the Taylor series expansion can express it as a polynomial of up to order s-1 plus the remainder term (Fuller 1996, pp 224-245). That is,

$$\hat{\theta} = \theta + f^{(1)}(Y)(\hat{Y} - Y) + \cdots + \frac{1}{(s-1)} f^{(s-1)}(\hat{Y} - Y)^{(s-1)}(Y) + R_n(\hat{Y}, Y)$$

where:

$f^{(i)}(Y)$ *is the ith derivative of* $f(y)$ *at* $y=Y$;

$i = 1, 2, \ldots, s-1;$

and $\hat{\theta} \rightarrow \theta$ as $n \rightarrow \infty$.

The linear function is produced by only retaining the linear terms in the function of the estimator. Thus, for a nonlinear parameter of interest θ with order 2 continuous derivatives, the linear variance estimator of $\hat{\theta}$ of the parameter = $Var\left\{\sum_{i=1}^{k}\frac{\partial f(\hat{Y})}{\partial y_i}(\hat{Y}_i - Y_i)\right\} = \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{\partial^2 f(\hat{Y})}{\partial y_i \partial y_j}Cov\left\{\hat{Y}_i, \hat{Y}_j\right\}$, where $Cov\left\{\hat{Y}_i, \hat{Y}_j\right\}$ is covariance between $\hat{Y}_i$ and $\hat{Y}_j$. This is the Taylor first order approximation to the variance estimator of a nonlinear estimator. As the Taylor first order approximation often produces satisfactory results, the first order expansion rather than higher order expansion is often used whenever the Taylor linearization method is applied (Särndal et al 1992 & Wolter 1985).

2.3.2 The Generalized Variance Function Method

The generalized variance function method is to build a mathematical relation connecting the variance of an estimator of a survey statistic to the expectation of the estimator through estimating the parameters of the model from past survey data or a small subset of data (e.g., a pilot sample survey data). Then, the model can produce the variance of the estimator by entering the estimate of the statistic into the model, rather than directly and individually computing the variance (Wolter 1985).

2.4 Discussion of Variance Estimation Methods

Among the foregoing methods of estimating sampling variance for an estimate of a parameter, **standard (mathematical) methods** are mathematically derived by applying probability theory. Thus, **standard (mathematical)**

methods can produce the most accurate and reliable results if data satisfies their requirements. However, these methods are only applied to estimate design effect or sampling variance for the simplest statistics - means.

Subsampling methods and modelling methods are used if standard methods are inapplicable. The following discussion is concerned with subsampling methods and modelling methods.

Random group methods require that the sampling design of subsamples is the same as that of their parent sample. This leads to formation of random groups before data collection, that is, all random groups have to be designed in the sample scheme before the survey starts. This tends to reduce the degree of stratification, and thus the efficiency of the sampling scheme. Due to this drawback, actual formation of random groups is seldom used in practice (Lee et al 1986 & Särndal et al 1992).

As random groups are often formed after data collection has been completed, the restriction that sampling design of subsamples is the same as that of their parent sample must be abandoned, that is, subsamples are not independent. Such non-independence causes underestimation of sampling variance, as there is covariance between estimates from random groups (subsamples). However, in many large-scale surveys, this covariance is not important (Wolter 1985, p 34).

On the other hand, the precision of estimates of sampling variances by random group methods depends on the number of subsamples. The number should be larger for more precise estimation. However, the number of subsamples is limited by the size of both the entire sample and the subsamples. If subsample size is too small, the subsample estimates may be of high variability. With a given number of subsamples, the high variability of the subsample estimates will lead to a larger sampling variance estimate

of the entire sample estimator for a parameter. That is, the small subsample size may lead to overestimation of the sampling variance of the entire sample estimator. A large number of subsamples may also lead to more costly and cumbersome calculation of the sampling variance estimate.

Thus, sampling variance estimation faces the decision of the appropriate number of subsamples within a given size of a single sample. As there are no laws to follow in making such decision, researchers have to determine the number of subsamples on their own experiences. There are several examples of determining number of subsamples. Mahalanobis (1939, 1944 & 1946) preferred to use four subsamples. Deming (1960) suggested that 10 subsamples be appropriate. Sudman (1976) suggested that the number of subsamples is limited to four.

Norlen & Waller (1979) suggested replication as a solution to overcome the difficulty in deciding the appropriate number of subsamples and reduce the variability of sampling variance estimates arising from different numbers of subsamples. That is, sampling variance estimates for the entire sample estimator of a parameter are obtained by repeating using a random group method where the subsample size is large enough and the number of subsamples small, and then the mean of these sampling variance estimates is the sampling variance estimate desired. This solution can be expressed in the formula:

$$v(\hat{\theta}) = \frac{1}{r}\sum_{i=1}^{r} v(\hat{\theta})_i$$

where:

$v(\hat{\theta})$ *is the sampling variance estimate of the entire sample estimate for a parameter* θ*;*

r *is the number of repeated replication;*

$v(\hat{\theta})_i$ *is the ith sampling variance estimate obtained by the formula 2.5.*

Random group methods are not useful if the number of PSUs is small at the first stage of multi-stage sampling (Hansen et al 1953).

Random group methods also have poor stability of variance estimation, if there are only two sampling units in each stratum in the case of stratified sampling designs. This leads to the production of balanced repeated replication methods.

Balanced repeated replication methods require that there is only a pair of sampling units within each stratum. These methods may not be applicable if it is difficult to stratify the sample into only two units in each stratum. The number of pairs of subsamples is also required to be integer and larger than the number of strata. However, the calculation with these methods is intensive if the number of strata is large. Thus, the appropriate number of subsamples is one of the decisions for application of balanced repeated replication methods.

The main interest of balanced repeated replication methods lies in their application to variance estimation for more complex statistics, for examples, correlation coefficients and regression coefficients, as they can produce more accurate results for such complex statistics (Särndal et al 1992). Balanced repeated replication methods can also lead to valid inference for both smooth and non-smooth statistics such as median and quantiles in complex sampling designs (Rao 1997).

Jackknife methods are more applicable than balanced repeated replication methods, if there is a large number of sampling units within each stratum or it is difficult to stratify the sample.

However, as subsamples are formed by omitting one group (unit), these methods perform poorly with non-smooth statistics such as median and quantiles for which balanced repeated replication methods may have valid performance. This is supported by some literature with empirical evidence. For examples, Sitter (1992) found that the performance of jackknife methods with the median was poor in comparison with the performance of five other methods in stratified sampling, that is, a jackknife estimator is of asymptotic inconsistency for estimating variance of the median. Similarly, Kovar, Rao & Wu (1988) empirically found that jackknife methods performed poorly for estimating variance of the quantiles.

Moreover, like random group methods, these methods have difficulty in deciding the appropriate number of subsamples.

Bootstrap methods have both advantages and disadvantages over the other methods (Kovar et al 1988; Rao & Wu 1988; Särndal et al 1992; & Sitter 1992).

Rao & Wu (1988) and Sitter (1992) found that bootstrap methods might be better in estimating confidence interval with one-tailed error rate than the other replication and the Taylor linearization methods. Similarly, Kovar et al (1988) investigated the application of bootstrap methods in stratified random sampling with replacement. They found that bootstrap methods tended to produce better confidence interval for ratios and correlation coefficients than the Taylor linearization method and jackknife methods.

However, the performance of bootstrap methods is less stable than those of the Taylor linearization method and jackknife methods (Kovar et al 1988; Rao & Wu 1988 & Särndal et al 1992).

Moreover, like random group methods and jackknife methods,

these methods have difficulty in deciding the appropriate number of subsamples.

The validity of **the Taylor linearization method** depends on that of the Taylor series expansion. The validity of the Taylor series expansion depends on two conditions. The first is the validity of the nonlinear function estimating a parameter of interest from observations. The second is the validity of a linear function approximating the nonlinear function. Thus, the validity of the Taylor linearization method depends on that of both the generation of nonlinear function from observations and the approximation of a linear function to this nonlinear function.

The Taylor linearization method also requires that the sample is large enough. The larger is the sample, the more precise the estimator with this method. Otherwise, it has a tendency to lead underestimation of sampling variance (Särndal et al 1992 & Woodruff 1971). The variance estimator may be unreliable in the case of highly skewed population distribution (Wolter 1985).

The validity of **the generalized variance function method** also depends on that of the variance functions estimating the parameters of interest from observations. This method is usually applied for large-scale multi-stage sample surveys and large number of statistics where statistics are grouped as it is much more convenient and economical than the others (Wolter 1985). An example of application of this method is the Current Population Survey conducted by the US Bureau of the Census.

Obviously, each sampling variance estimation method has both advantages and disadvantages. One common drawback of **subsampling methods** is the difficulty of deciding the appropriate number of subsamples, and the key in applying **modelling methods** is the validity of the nonlinear function of describing a parameter generated from

observations. When subsampling methods and modelling methods are applicable, which of them is the best? None of them is the best, according to such criteria as the accuracy of estimation, flexibility of application, the cost of calculation and convenience of administration. This is a common finding in the literature.

Kish and Frankel (1974) investigated the performance of the Taylor linearization method, balanced repeated replication methods and jackknife methods by using these methods to estimate sampling variance of correlation coefficients, ratio means, and regression coefficients. They found that there was the variability of sampling variance estimation among these methods for a given statistic. Kish (1987) further discussed the application of these methods. He found out that none of these three methods was clearly superior.

Wolter (1985) used the results from five empirical studies to investigate the performance of random group methods, balanced repeated replication methods, jackknife methods and the Taylor linearization method in estimating sampling variances. He also investigated the performance of generalized function method. He found that all these methods had advantages and disadvantages. He pointed out none of these methods was best.

Similarly, Lehtonen & Pahkinen (1995) found that there was variability of performance with the Taylor linearization method, balanced repeated replication methods, jackknife methods and bootstrap methods in estimating sampling variances.

The finding that none of subsampling methods and modelling methods is clearly superior to the others implies that it is difficult to make the selection among these alternative methods of estimating sampling variances when standard methods are inapplicable.

3. DESIGN EFFECT

3.1 Introduction

The methods discussed in the previous chapter are often used to estimate the sampling variance or design effect of an estimate from a complex design sample.

Kish (1965) proposed "design effect" as a measurement of the efficiency of complex sample designs. This proposition is based on Hansen et al (1953) who pointed out that in a one-stage simple random cluster sample, sampling variance based on clusters tends to be larger than that based on elements. This sampling variance increase is caused by similarity of elements within clusters, which can be measured by within-cluster homogeneity ρ.

Since then, there have been some common findings in the literature. These findings are:

1. Design effect depends on sampling designs, that is, design effect varies from one sampling design to another;

2. Design effect depends on the nature of variables measured for a given statistic, that is, design effect varies from one variable to another for a given sampling design;

3. Design effect depends on statistics for a given variable, that is, design effect varies from one statistic to another for a given sampling design.

This chapter reviews these findings in the literature.

3.2 Design Effects for Different Statistics and Variables

3.2.1 Design Effect for Different Statistics

> *For a given variable and a given sampling design, the design effect of one statistic is different from that of another.*

This is a common finding in the literature. Within a given sampling design, a number of authors have described design effect on various statistics for a given variable.

Kish (1965) found that design effects for differences between subclass means were less than those for the corresponding subclass means. He pointed out that the covariance between subclass means was positive and this implied that design effects for subclass means were larger than those for differences between subclass means. This is consistent with the empirical result of Verma & Lê (1996). Verma & Lê (1996) used data from 48 nationally representative surveys under the Demographic and Health Surveys Programmes to investigate design effects for subclass means and those for differences between these subclass means. They found that design effects for differences between subclass means were usually lower than those for the corresponding subclass means.

Kish & Frankel (1970) & Frankel (1971) further provided an empirical investigation on design effects for overall means and other statistics such as domain means, correlation coefficients and regression coefficients. They found that design effects for overall means are larger than those for domain means. The reason is that in a clustered sample, homogeneity ρ is nearly equal for both the original sample clusters and the crossclusters that are formed by subclasses across the original sample clusters, but the sizes of crossclusters are reduced. Such reductions in crosscluster sizes lead to decreases in

design effects for subclass means. They also found that within given variables or subclasses, the design effects for regression coefficient and correlation coefficient are less than those for the corresponding means.

From these empirical findings, Kish & Frankel (1974) conjectured that the design effects for simple statistics such as means tend to be larger than those for more complex statistics such as difference between means, correlation coefficients and regression coefficients. Kish (1987) further provided detailed discussion of these findings and conjectures. He pointed out that relation among design effects for complex statistics was difficult to predict or conjecture. These conjectures about the relation between simple statistics and complex statistics are supported by Bebbitington & Smith (1977), Campbell (1977) and Scott & Holt (1982).

Bebbitington & Smith (1977) found that design effects for means were larger than those for correlation coefficients.

Campbell (1977) and Scott & Holt (1982) investigated the effect of clustering on ordinary least squares regression analysis. In order to explain the observation that design effects for regression coefficients of independent variables are smaller than that for the sample mean of the dependent variable, they analysed design effect for a regression coefficient in a simple ordinary least squares regression analysis (i.e., *a simple regression equation y=bx+e, where y is the dependent variable; b is the regression coefficient; and e is error*) with the assumption that the regression coefficient is the same in all clusters. They found that design effect for the regression coefficient was $1 + (m-1)\rho_e\rho_x$ (intracluster correlation coefficient of residual and independent variable respectively). As the product of ρ_e and ρ_x is often smaller than ρ (intracluster correlation coefficient of dependent variable), the design effect for the

regression coefficient is smaller than that for sample mean of the dependent variable.

3.2.2 Design Effect for Different Variables

> *For a given statistic and a given sampling design, the design effect of one variable is different from that of another.*

In clustered samples, the homogeneity of elements within clusters varies from one variable to another. This leads to variability of design effect among variables. Such variability is a common finding in the literature (Ferringo, Valli, Groenerveld, Buch & Coetzee 1992; Kish 1965 & 1987; Kish & Frankel 1974 and Verma & Lê 1996).

3.3 Design Effect and Stratification

Stratification may reduce design effect and thus improve the efficiency of sampling designs. Such improvement depends on both the within-stratum variance in a sample and the entire sample variance. If the latter is larger than the former, design effect is less than one. The efficiency of the sample design is improved. However, such improvement is usually small if the sample is geographically and proportionally stratified (Bebbitington & Smith 1977; Hansen et al 1953 & Kish 1965). This is also supported by the calculated result from Alsagoff et al (1986).

3.4 Design Effect and Clustering

The effect of clustering on design effect has been discussed in the literature for many years. The literature is directly related to what this study is concerned with. Face-to-face interview surveys often use multi-stage

sampling (clustering) designs, as such designs are much more economical than one-stage simple random sampling designs. Each stage of multi-stage sampling affects both survey cost and design effect. However, the final stage clustering designs are crucial in influencing survey cost.

There are two components in clustering: cluster size and the ways of selecting elements to form clusters. Usually, there are three ways: **selecting consecutive elements** from a random start point (the randomly selected element), **systematically selecting elements with some interval** from a random start point, and **simple random selection of elements**, to form clusters within given sampling areas in final stage of multi-stage (included two-stage) sampling. These sampling areas are often city blocks or enumeration districts (i.e., "mesh blocks" in New Zealand).

Simple random sampling is rarely used in practice to select elements to form clusters, as its application is not economical in face-to-face interview surveys. On the other hand, selecting consecutive elements to form clusters is the same as systematically selecting every-first elements to form clusters. Thus, the decisions facing the final stage sampling designs of surveys are the cluster size and clustering interval. These decisions will affect both the design effect and the cost of data collection for a sample with given size.

3.4.1 Design Effect and Cluster Size

> *Design effect increases with increase in cluster size, no matter how clusters are formed.*

Increase or decrease in design effect is caused by cluster size m and homogeneity ρ. Hansen et al (1953) called this the effect of clustering on sampling variance in terms of $1+(m-1)\rho$. Kish (1965) expressed it as design effect. That

is,

$$\text{Design effect} \approx 1+(m-1)\rho. \qquad 3.0$$

This effect is obtained by mathematically deriving from sampling variance based on cluster totals to sampling variance based on elements. The detailed derivation is in Hansen et al (1953), Kish (1965) and Cochran (1963 & 1977).

This formula can be used to calculate either design effect or homogeneity if cluster size and either homogeneity or design effect is known. This formula is often used to calculate homogeneity in practice.

Hansen et al (1953) empirically investigated the effect of clustering on homogeneity in the final stage of multi-stage sampling. They used data from the 1940 Census of the Population in USA to form simulated clustered samples with fixed cluster number and various cluster sizes: 3, 9, 27, and 62 (or 252 for rural areas) for their study. These clusters are formed by selecting consecutive households within city blocks or enumeration districts.

They found that homogeneity decreased with the increase of cluster size, but the rate of cluster size increase was much larger than the rate of homogeneity decrease. Such homogeneity decrease rate is so small that homogeneity can in practice be assumed to be unchanged with increase of cluster size (Sudman 1976). Thus, design effect increases with increase in cluster size.

Similarly, Laniel & Mohl (1994) used data from the Canadian Labour Force Survey in Ottawa to investigate the effect of cluster size on the efficiency of sampling designs. They also created simulated samples for their study. With given entire sample size, three average sample cluster sizes: 4.7, 8.2 and 16.4, were formed within the

sampling areas with different sizes: 50, 100, 150, 200 and 250. That is, there were three simulated samples with average sample cluster sizes: 4.7, 8.2 and 16.4 respectively for each sampling area. This led to the formation of 15 simulated samples. They found that the efficiency of sample design is improved with the decrease of cluster size for given entire sample size. However, they did not consider the effect of variation in number of clusters, though such effect may be small. For given entire sample size, reduction in cluster size leads to increase in number of clusters. This increase in number of clusters leads to a reduction in the sampling variance of an estimate from this larger-cluster-size sample while the sampling variance of the estimate from a simple random sample is the same, as the entire sample size is the same in different cluster sizes. This leads to underestimation of the effect of cluster size.

3.4.2 Design Effect and Clustering Interval

> *For given sample cluster size, design effect decreases with the increase of clustering interval.*

Although this position is popular in the literature, for examples, Hansen et al (1953), Kish (1965) and Sudman (1976), there is little empirical investigation of it.

Laniel & Mohl (1994) used data from the Canadian Labour Force Survey in Ottawa to investigate the effect of clustering interval on the efficiency of sampling designs. They formed 15 simulated samples for their study. As clustering interval is the size of sample areas over that of sample clusters, the minimum clustering interval is (50 / 16.4) ≈ 3, and the maximum clustering interval (250 / 4.7) ≈ 53.2. They found that the efficiency of sample design is improved with increase in clustering interval.

They insisted that this effect was substantial. However, they did not take variation in number of clusters into account in estimating sampling variance from the simulated samples. For a given sample size and a given sampling area size, increase in clustering interval leads to both decrease in cluster size and increase in number of clusters. Both decrease in cluster size and increase in number of clusters reduce the sampling variance of an estimate from this larger-clustering-interval sample while the sampling variance of the estimate from a simple random sample is the same, as the entire sample size is the same in different clustering intervals. Thus, the effect of clustering interval is overestimated.

Laniel & Mohl (1994) also found that the curve of the combination of the effect of cluster size and that of clustering interval is nearly linear.

4. METHOD

4.1 Procedure

In the first stage, a search of relevant literature was undertaken to identify variance estimation methods, homogeneity of elements within cluster, variability of design effects among both variables and statistics, the effect of stratification on design effect, and the effect of both cluster size and clustering interval on design effect.

In the second stage, ACNielsen McNair was asked to provide data from a face-to-face interview survey with a clustered sample design for this study. The data included the following information:

- *The details of the coding;*
- *Demographic information;*
- *The method for selecting clusters;*
- *A cluster id;*
- *A description of the cluster sampling scheme.*

There are two reasons for obtaining this secondary data rather than the first hand data. First, the data for this study had to be from face-to-face interview surveys with clustered sample designs and should be large. This leads to the impossibility of conducting such surveys as resources were limited for this study. Second, it is unnecessary to conduct a face-to-face interview survey to collect the data for this study as the first hand data is not required.

4.2 Samples

Data is used from a two-stage sample rather than a multi-stage sample in this study. This is for the simplicity and accuracy of estimating sampling variances of estimates. If data is from a multi-stage sample, the estimation of sampling variances of estimates may be more complicated and less accurate. On the other hand, data from two-stage samples can satisfy the requirements of data for this study.

4.2.1 Original Sample

This is a two-stage sample for a face-to-face interview survey with the following design:

> *The sampling elements (i.e., ultimate sampling units) are households. The sample is proportionally geographically stratified into urban and rural areas. 117 sampling areas are selected at the first stage sampling by simple random sampling. Eight households for each sampling area are selected by systematic sampling method with every third household from the random starting point. These eight households form a cluster. One adult is randomly selected from each household for the interview.*

In the data collection, interviewers are required to do callbacks to make up 8 responses for each cluster. If there are still less than 8 responses in a cluster after these callbacks, the interviewer is required to interview extra households until eight responses are included in the cluster (see table 1). This sample includes 936 households. Almost all these households have responses except that some of them refuse for some variables (see table 2).

In table 1 and table 2, there are:

The maximum number in a cluster of households is 50;

The number of 90 % of households selected is between 1 and 21;

A few variables have missing values;

And the maximum number of missing cases is 45 out of 936, that is, less than 5%.

Very low missing value rate of this data leads to little effect of nonresponse on variance estimation and statistical analysis. As these clusters are of equal size, weights are not required to adjust the effect due to variation in cluster size on variance estimation. That distribution of selected elements is between the random start point and the 21st household may reduce the effect of geographical size on variance estimation. All these indicate that data of this sample is nearly perfect for this study.

A large number of variables are included, involving newspaper and magazine readership, news preferences, social attitudes and activities, cigarette smoking, and mortgage on houses. 81 variables are selected for this study. The definition of these variables is Appendix A.

Table 1. Frequency of Households Interviewed

Household-number	Frequency	Percent	Cumulative Percent
1	60	6.4	6.4
2	52	5.6	12.0
3	47	5.0	17.0
4	55	5.9	22.9
5	48	5.1	28.0
6	48	5.1	33.1
7	51	5.4	38.6
8	54	5.8	44.3
9	47	5.0	49.4
10	51	5.4	54.8
11	52	5.6	60.4
12	47	5.0	65.4
13	43	4.6	70.0
14	36	3.8	73.8
15	31	3.3	77.1
16	26	2.8	79.9
17	23	2.5	82.4
18	24	2.6	84.9
19	20	2.1	87.1
20	18	1.9	89.0
21	17	1.8	90.8
22	15	1.6	92.4
23	12	1.3	93.7
24	12	1.3	95.0
25	10	1.1	96.0
26	4	.4	96.5
27	4	.4	96.9
28	7	.7	97.6
29	4	.4	98.1
30	1	.1	98.2
31	3	.3	98.5
32	2	.2	98.7
33	4	.4	99.1
34	1	.1	99.3
36	1	.1	99.4
39	1	.1	99.5
43	1	.1	99.6
44	2	.2	99.8
47	1	.1	99.9
50	1	.1	100.0
Total	936	100.0	

Table 2. Response Rate for Designed Sample Size 936

Response Number	Response Rate	Variable Number	Proportion of Variable Number
891	891/936 = 95.2%	4	5.0
925	925/936 = 98.8%	1	1.2
932	932/936 = 99.6%	1	1.2
933	933/936 = 99.7%	9	11.1
934	934/936 = 99.8%	9	11.1
935	935/936 = 99.9%	16	19.8
936	936/936 = 100%	41	50.6
Total		81	100.0

4.2.2 Simulated Samples

Based on the original sample, 17 simulated samples were formed. The detailed procedures to form these simulated samples are in Appendix B. These simulated samples have a given number of clusters while cluster size and clustering interval vary. This is different from the formation of simulated samples of Laniel & Mohl (1994). The same cluster number avoids the effect of variation in number of clusters on design effect. Some of 17 stimulated samples and the original sample are used to investigate the effect of cluster size, and some other samples to investigate the effect of clustering interval.

There are 4 cluster sizes in these samples and the original sample, that is, 2, 4, 6, and 8. The original sample is symbolized with **CS8_1,** and the simulated samples with **S61, S62, S63, S64** and **S65** for cluster size 6, **CS4_2, S41, S42, S43**, and **S44** for cluster size 4, **CS2_1, CS2_2, CS2_3, CS2_4, CS2_5, CS2_6,** and **CS2_7** for cluster size 2.

The samples **CS8_1**, **S61, S41** and **CS2_1** are used to investigate the effect of cluster size. The clusters of **S61, S41** and **CS2_1** are also called "pseudo-compact" clusters as they are formed by selecting consecutive elements within the original sample's clusters. Thus, these four samples have different cluster sizes and the same clustering interval.

The samples **S41, CS4_2, CS2_1, CS2_2, CS2_3, CS2_4, CS2_5, CS2_6** and **CS2_7** are used to investigate the effect of clustering interval. The clustering intervals of these samples are multiples of the smallest clustering interval. That is, the clustering interval of **CS4_2** is twice as much as that of **S41**, and those of **CS2_2, CS2_3, CS2_4, CS2_5, CS2_6** and **CS2_7** are multiples of the clustering interval of **CS2_1**.

The samples **S62, S63, S64, S65, S42, S43** and **S44** are used to help investigate the effect of clustering interval. The clustering intervals of these samples are not a multiple of the smallest clustering interval. But the mean clustering interval of each sample is larger than the smallest clustering interval. That is, in **S62, S63, S64** and **S65,** each mean clustering interval is larger than the clustering interval of **S61**, and in **S42, S43** and **S44,** each mean clustering interval larger than the clustering interval of **S41**.

As it is unknown whether there are differences between these mean clustering intervals or not, the investigation on the effect of the clustering interval will be made by comparing the design effect of the smallest clustering interval with each design effect from **S62, S63, S64, S65, S42, S43** and **S44.** That is, the design effects of **S62, S63, S64,** and **S65** are compared with that of **S61**, and the design effects of **CS4_2, S42, S43**, and **S44** with that of **S41.**

4.3 Estimation for Design Effect

4.3.1 Considerations of Simplicity

Simplicity should be considered in estimating sampling variance estimates or design effect estimates, as it reduces cost of calculation.

First, sampling variance estimates or design effect estimates will be calculated over strata rather than in each stratum. That is, the effect of stratification is ignored. Thus, it is simpler to estimate sampling variance or design effect. The main reason for ignoring the effect of stratification is that this study is concerned with the effect of clustering. Another reason is that proportional and geographical stratification has little influence on design effect. The effect of stratification usually improves the efficiency of sample designs. Ignoring it may lead to overestimation of design effect. Moreover, the data used in this study is geographically and proportionally stratified into two strata: rural areas and urban areas. Such low degree of stratification may be of little effect on design effect.

Second, this study only investigates design effect for the mean, the simplest descriptive statistic, rather than complex statistics. This leads to a larger number of alternative methods for selection in estimating sampling variance or design effect. On the other hand, investigating design effect for the mean is much simpler and less expensive than investigating those for complex statistics. Moreover, for a given sample and a given variable, the design effect for mean is larger than those for complex statistics. This leads to a conservative result if the design effect for mean is used, rather than those for complex statistics, to adjust these complex statistics in statistical analysis.

4.3.2 Estimation Method for Design Effect

Although the literature suggested a great number of methods for the estimation of sampling variance or design effect, this study uses a simple standard method to estimate design effects for means if data satisfies the requirements of the method. This method is to estimate design effect for mean by the ratio of between-cluster mean squares in the sample and mean squares in the sample, that is, $design\ \ effect = \frac{ms_b}{ms}$. It requires that the samples have clusters (subsamples) with equal size and both clusters and elements are selected with equal probability. The detailed derivation of $design\ \ effect = \frac{ms_b}{ms}$ is in Appendix C. In practice, the elements are by systematic sampling rather than simple random sampling. Thus, elements selected by systematic sampling are taken as the same as those selected with equal probability.

The method $design\ \ effect = \frac{ms_b}{ms}$ seems more reliable and simpler than other methods in estimating design effect for means. This is discussed in the following.

First, subsampling and modelling methods of estimating sampling variances discussed in Chapter 2 are approximate. They are often used if standard mathematical methods are not applicable or the cost of applying these standard methods is tremendous. The chances to use these estimation methods are often those to estimate variance for complex statistics in complex designs. On the other hand, these approximate variance estimation methods are of variable performance.

Second, data for this study is from a two-stage sample with the clusters of a given size selected by epsm (equal probability selection method) and with the elements

selected by systematic sampling from the random start point. Very high response rate is in the data collection. Thus, this data set satisfies the requirements of $design \quad effect = \frac{ms_b}{ms}$.

With such data, the estimation of design effects for means can be carried out without attention paid to adjustments to the variation in sampling variance estimates arising from the effect of variation in both cluster size and probability of selection.

Moreover, this method can use a single sample to estimate sampling variance for mean without concerning either the decision for the optimal number of subsamples in the subsampling methods or the validity of the functions of describing the parameters of interest from observations in the modelling methods. Both the decision for the optimal number of subsamples and the validity of the functions of describing the parameters of interest from observations affect the reliability and simplicity of estimating sampling variance or design effect.

Therefore, the performance of this method is likely to be the most reliable and simplest, though it is also approximate.

4.3.3 An Alternative Method of Estimating Design Effect

One alternative method of estimating design effect for means is to build connection of the design effect with $r^2 = \frac{SS_B}{SS}$ (i.e., ratio of between-cluster sum of squares and total sum squares in the sample) and m (i.e., cluster size), that is, $design \quad effect \approx mr^2$. This idea is proposed by Don Esslemont (the adviser of this study). It can

mathematically be derived from $design\ \ effect = \frac{ms_b}{ms}$. The detailed derivation is in the following:

$$Design\ \ effect = \frac{ms_b}{ms}$$

$$= \frac{\frac{S_B^2}{k}}{\frac{S^2}{n}}$$

$$= \frac{\frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k(k-1)}}{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n(n-1)}}$$

$$= \frac{\frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k(k-1)}}{\frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y})^2}{km(km-1)}}$$

$$= \frac{\frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k-1}}{\frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y})^2}{km-1}} * \frac{km}{k}$$

$$= \frac{\frac{\sum_{j=1}^{k} m(\bar{Y}_j - \bar{Y})^2}{k-1}}{\frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y})^2}{km-1}}$$

$$= \frac{\frac{SS_B}{k-1}}{\frac{SS}{km-1}}$$

$$= \frac{SS_B}{SS} * \frac{(km-1)}{(k-1)}$$

$$= \frac{SS_B}{SS} * \frac{(n-1)}{(k-1)}$$

$$\approx r^2 * \frac{n}{k}$$

$$\approx mr^2.$$

For testing the applicability of this alternative method in estimating design effect, the data is required to have not only a fixed number of clusters but also no missing values within clusters, in order to avoid the effects on design effect of the variation in both cluster size and cluster number. Thus, 41 variables with three samples **CS8_1, S61** and **S41** were selected (see Appendix F).

4.4 Significance Tests

T-tests will be applied to investigate significance of both the effect of clustering interval and the effect of cluster size on design effect. Such significance tests will be done by comparing design effects between either different cluster sizes or different clustering intervals. That is:

> *The null hypothesis is no differences between two separate clustering intervals or cluster sizes;*
> *And the effect hypothesis is significant differences between these two clustering*

intervals or cluster sizes.

4.5 Evaluation of Cost-Efficiency in the Sample Designs with Alternative Clustering Intervals

Evaluation of cost-efficiency in the sample designs with alternative clustering intervals provides the basis for selection of alternative sampling designs with different clustering intervals. As an increase in clustering interval leads to both a decrease in design effect and an increase in the cost of travelling between two interviews, the parameter to evaluate the cost-efficiency of sampling designs is the product of the travel cost between two interviews and the design effect. A smaller product of travel cost and design effect leads to a better sampling design.

Whenever selection is made between two alternative clustering intervals A and B for a sample design, either A or B can be selected if:

(Interview-cost + Travel-cost A) Design effect A =
(Interview-cost + Travel-cost B) Design effect B

where:

Interview-cost is the cost of interviews;
Travel-cost A is the cost of travelling within clusters for sample design A;
Travel-cost B is the cost of travelling within clusters for sample design B.

Clustering interval A is better than clustering interval B, if:

(Interview-cost + Travel-cost A) Design effect A <
(Interview-cost + Travel-cost B) Design effect B.

5. RESULTS

5.1 Design Effects

Design effects for individual variables were estimated from the 18 samples (see Appendix D). Both the effect of cluster size and the effect of clustering interval can be observed from these estimated design effects.

5.1.1 The Effect of Cluster Size

There are several observations on the effect of cluster size. These observations are described in the following.

First, design effects for individual variables tend to increase with an increase in cluster size. This is indicated by table 3. That an individual design effect increases with an increase in cluster size is consistent among different quartiles of variables.

Table 3. Design Effects for the Quartiles of Variables in Different Cluster Sizes

Cluster Size	8	6	4	2
Clustering*	CS8_1	S61	S41	CS2_1
Percentile :				
25	1.08	1.07	1.04	.98
50	1.24	1.16	1.15	1.05
75	1.42	1.36	1.25	1.15
100	3.91	3.05	2.19	1.25

* See Chapter 4: METHOD.

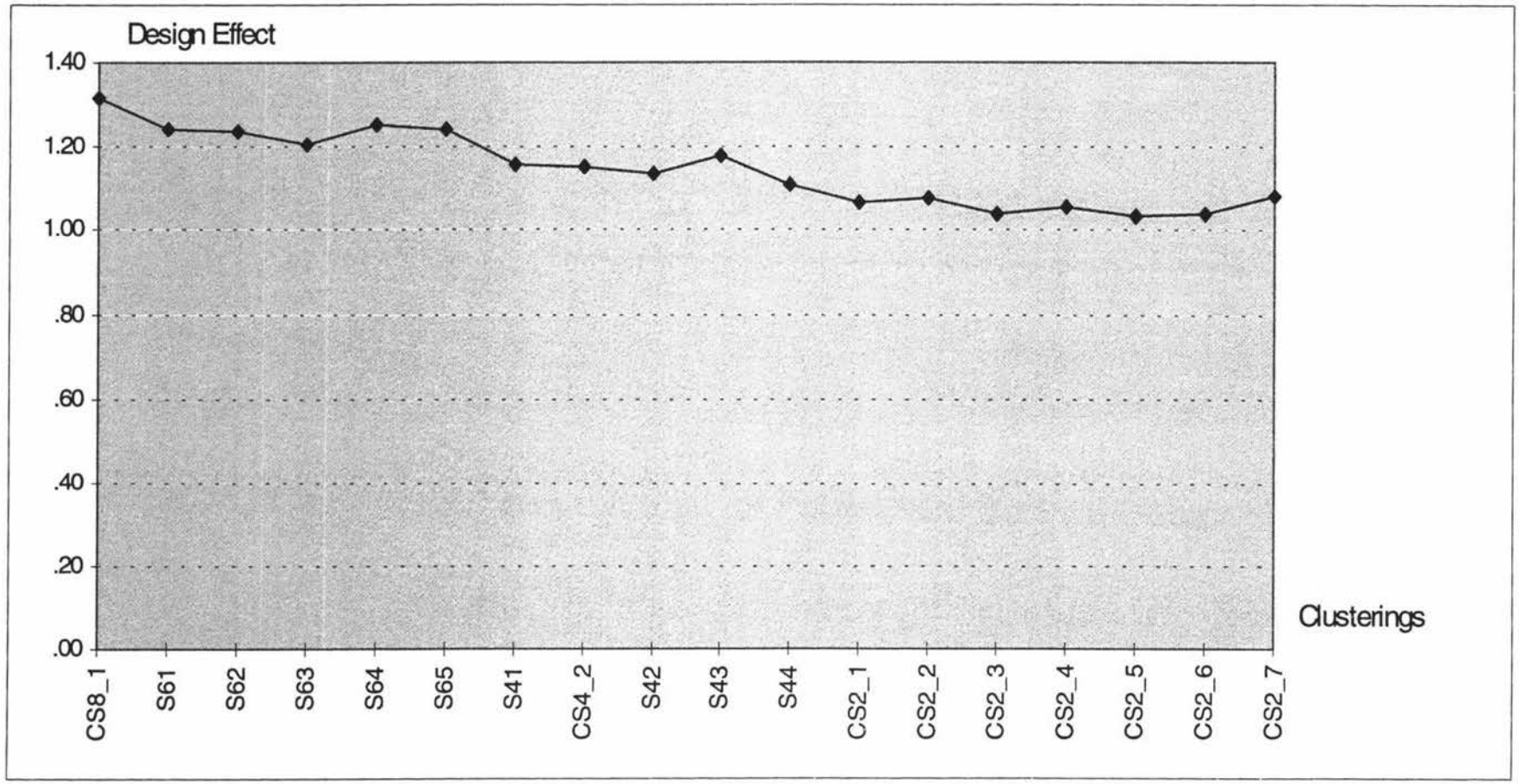

Figure 1. Relation between Design Effect and Clusterings

Second, the curve of mean design effect of variables tends to slope down from a larger cluster size to a smaller size, no matter what the clustering interval is. This is indicated by figure 1. That is,

A mean design effect with cluster size 8 is always larger than those with cluster size 6;

A mean design effect with cluster size 6 is always larger than those with cluster size 4;

A mean design effect with cluster size 4 is always larger than those with cluster size 2.

Third, within a given clustering interval, the curve of mean design effect of variables slopes down from a larger cluster size to a smaller size. This curve is linear. This is indicated by figure 2.

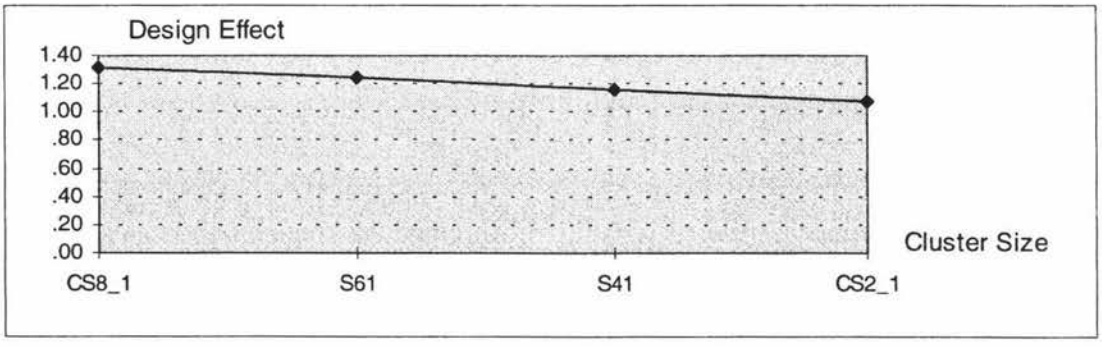

Figure 2. Relation between Design Effect and Cluster Size

Figure 2 and table 4 illustrate that within pseudo-compact clusters with different cluster sizes, mean design effect of variables increases by about 0.04 as cluster size increases by one unit. The detailed observations of such design effect increase are:

1. Mean design effect of variables with cluster size 8 is 0.08 greater than that with cluster size 6;

2. Mean design effect of variables with cluster size 6 is 0.08 greater than that with cluster size 4;

3. Mean design effect of variables with cluster size 4 is 0.09 greater than that with cluster size 2.

Fourth, the variability of individual design effect among variables is stronger with larger cluster sizes. This is reflected in table 4.

The larger *range*, *standard deviation*, and *coefficient of variation* among individual design effects are with a larger cluster size. The smaller ones are with a smaller cluster size.

Table 4. Variability of Design Effect among Variables in different Cluster Sizes

Cluster Sizes	**8**	**6**		**4**		**2**	
Clusterings #	**CS8_1**	**S61**	**S6***	**S41**	**S4***	**CS2_1**	**S2***
Means	1.32	1.24	1.24	1.16	1.15	1.07	1.06
Standard Deviation	0.40	0.31	0.30	0.19	0.20	0.11	0.12
Coefficient of Variation	0.30	0.25	0.24	0.16	0.17	0.10	0.11
Range	3.03	2.24	2.22	1.31	1.40	0.54	0.65

see Chapter 4: METHOD.

* S6 is the average of S61, S62, S63, S64 and S65;
  S4 is the average of S41, S42, S43 and S44;
  S2 is the average of CS2_1, CS2_2, CS2_3, CS2_4, CS2_5, CS2_6 and CS2_7.

Fifth, the results from t-tests for the differences of design effects between cluster sizes indicate that the effect of cluster size on design effect is of high significance level. This is reflected in table 5.

All these results are from t-tests which are applied to

test α level of differences between mean design effect of variables with a larger cluster size and that with a smaller cluster size.

All results from these tests are the same. Their α levels are less than 0.005.

Table 5. t-tests for Differences of Design Effects between Cluster Sizes

Difference of Design Effects between Cluster Sizes*	Correlation Coefficients between Cluster Sizes	t-Values	Significance Level (α)
CS8_1 - S61	0.96	5.16	0.00
CS8_1 - S41	0.87	5.80	0.00
CS8_1 - CS2_1	0.48	6.36	0.00
S61 - S41	0.91	4.82	0.00
S41 - CS2_1	0.57	5.19	0.00

* See Chapter 4: METHOD.

5.1.2 The Effect of Clustering Interval

The observations on the effect of clustering interval are described in the following.

First, within a given cluster size, design effects for individual variables are very similar in different clusterings. This is reflected in table 6.

Individual design effects of the quartiles of variables are very similar for different clusterings within a given cluster size.

Second, mean design effects of variables are very similar for clustering intervals with different multiples of the smallest clustering interval, though there are small

differences among these design effects. This is indicated by figure 3.

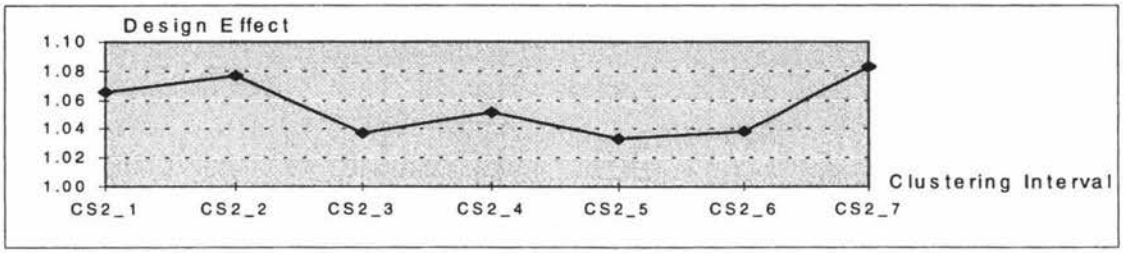

Figure 3. Relation between Design Effect and Clustering Interval with Cluster Size 2

In figure 3, design effect increases with increase in clustering interval from CS2_1 to CS2_2, from CS2_3 to CS2_4 and from CS2_6 to CS2_7, but decreases with increase in clustering interval from CS2_2 to CS2_3 and CS2_4 to CS2_5. These increases and decreases are small. Their mean absolute deviation is 0.02.

Third, design effects with the larger mean clustering intervals are very similar to one with the smallest mean clustering interval. This is reflected in figure 4, figure 5 and table 4.

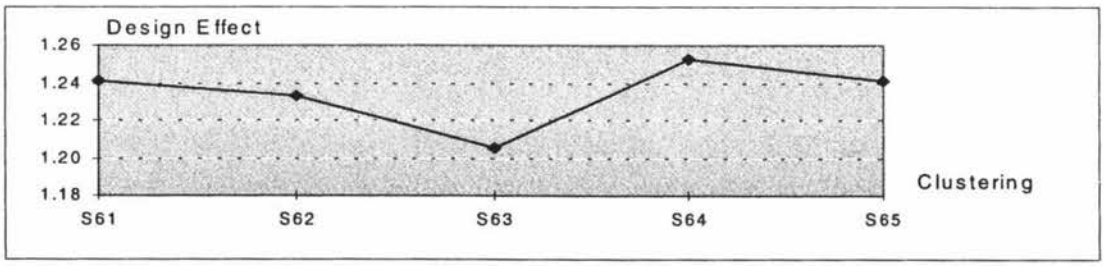

Figure 4. Relation between Design Effect and Clustering Interval with Cluster Size 6

In figure 4, design effects of S62, S63, S64 and S65 are larger or less than that of S61. The differences are small between design effect of S61 and each of the design effects with S62, S63, S64 and S65. Their mean absolute deviation is 0.01.

Table 6. Design Effects for the Quartiles of 81 Variables in Different Clusterings

Cluster Size	8	6					4					2						
Clustering*	CS8_1	S61	S62	S63	S64	S65	S41	CS4_2	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Percentile:																		
25	1.08	1.07	1.07	1.03	1.09	1.07	1.04	1.01	1.00	1.04	.98	.98	.98	.97	.97	.96	.96	.99
50	1.24	1.16	1.17	1.17	1.21	1.21	1.15	1.13	1.12	1.15	1.06	1.05	1.07	1.01	1.03	1.01	.99	1.05
75	1.42	1.36	1.35	1.32	1.34	1.35	1.25	1.25	1.23	1.27	1.19	1.15	1.15	1.10	1.12	1.10	1.11	1.15
100	**3.91**	**3.05**	3.19	3.06	3.03	3.08	**2.19**	2.30	2.20	2.15	2.33	**1.25**	1.41	1.39	1.40	1.44	1.54	1.55

* See Chapter 4: METHOD.

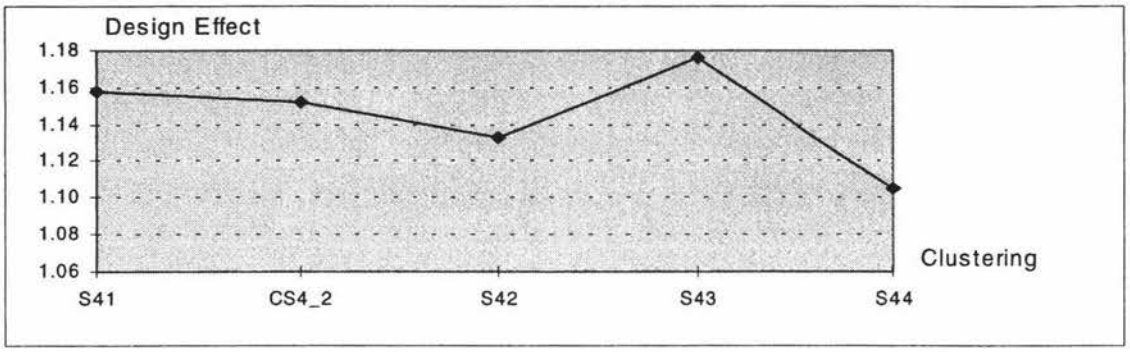

Figure 5. Relation between Design Effect and Clustering Interval with Cluster Size 4

In figure 5, design effects of S42, S43, S44 and CS4_2 are larger or less than that of S41. The differences are also small between design effect of S41 and each of the design effects with S42, S43, S44 and CS4_2. Their mean absolute deviation is 0.02.

In table 4, mean design effect over clustering intervals is the same as that for the smallest clustering interval within a given cluster size. That is, design effects for S6, S4 and S2 are the same as those for S61, S41 and CS2_1 respectively.

Fourth, the variability of individual design effects tends to be the same in the different clustering intervals for a given cluster size. This is indicated by table 7.

In table 7, *range, standard deviation*, and *coefficient of variation* among individual design effects are little different among clustering intervals for a given cluster size.

Table 7. Variability of Design Effect among Variables in Different Clusterings

Clusterings*	CS8_1	S61	S62	S63	S64	S65	S41	CS4_2	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Means	**1.32**	**1.24**	1.24	1.21	1.25	1.24	**1.16**	**1.15**	1.13	1.18	1.11	**1.07**	1.08	1.04	**1.05**	1.03	1.04	**1.08**
Standard Deviation	**0.40**	**0.31**	0.31	0.28	0.29	0.30	**0.19**	**0.20**	0.20	0.20	0.21	**0.11**	0.11	0.10	**0.11**	0.11	0.15	**0.12**
Coefficient of Variation	**0.30**	**0.25**	0.25	0.23	0.23	0.24	**0.16**	**0.17**	0.18	0.17	0.19	**0.10**	0.10	0.10	**0.10**	0.11	0.14	**0.11**
Range	**3.03**	**2.24**	2.32	2.17	2.11	2.25	**1.31**	**1.46**	1.39	1.25	1.60	**0.54**	0.53	0.54	**0.57**	0.62	1.13	**0.63**

* See chapter 4: METHOD.

Table 4 also indicates that the variability of individual design effects tends to be the same in the different clustering intervals for a given cluster size. Within a given cluster size, *mean range*, *mean standard deviation*, and *mean coefficient of variation* over clustering intervals are very similar to *range*, *standard deviation*, and *coefficient of variation* for the smallest clustering interval. That is, *range*, *standard deviation*, and *coefficient of variation* for S6, S4 and S2 are the same as those for S61, S41 and CS2_1 respectively.

Fifth, the results from the t-tests for the differences of design effects between clustering intervals indicate that the effect of clustering interval on design effect tends to be of low significance level within a given cluster size. This is indicated by table 8.

The α level of these t-test results tends to be larger than 0.05. Only three values of α are smaller than 0.05.

Even if the increase in the clustering interval is substantial, the t-test result indicates that the difference of design effects between a larger clustering interval and a smaller one is likely to be of very low significance level. In the results of t-tests for S41 - CS4_2, CS2_1 - CS2_2, CS2_1 - CS2_3, CS2_1 - CS2_4, CS2_1 - CS2_5, CS2_1 - CS2_6 and CS2_1 - CS2_7, only the value of α for CS2_1 - CS2_5 is smaller than 5%.

Table 8. t-tests for Differences of Design Effects between Clustering intervals

Difference of Design Effects between Clustering Intervals*	Correlation Coefficients between Clustering Intervals	t-Values	Significance Level (α)
S61 - S62	0.92	0.58	0.56
S61 - S63	0.89	2.29	0.02
S61 - S64	0.90	0.72	0.47
S61 - S65	0.94	0.01	0.99
S41 - CS4_2	**0.77**	**0.38**	**0.71**
S41 - S42	0.72	1.57	0.12
S41 - S43	0.72	1.13	0.26
S41 - S44	0.73	3.18	*0.00*
CS2_1 - CS2_2	**0.33**	**0.62**	**0.53**
CS2_1 - CS2_3	**0.23**	**2.02**	**0.05**
CS2_1 - CS2_4	**0.37**	**1.04**	**0.30**
CS2_1 - CS2_5	**0.35**	**2.43**	**0.02**
CS2_1 - CS2_6	**0.05**	**1.46**	**0.15**
CS2_1 - CS2_7	**0.27**	**1.00**	**0.32**

* See Chapter 4: METHOD.

5.2 Applicability of $design\ \ effect \approx mr^2$

For testing the applicability of $design\ \ effect \approx mr^2$, the design effects estimated by $design\ \ effect \approx mr^2$ are compared with those estimated by the method used in this study. The detailed comparison between these "alternative" design effects and the corresponding design effects by the method used in this study is in Appendix F.

These results indicate both methods are consistent. The design effects by the alternative method are virtually the

same as those by the method used in this study. This is reflected in table 9.

For almost all variables, the absolute difference between the alternative design effects and the used design effects within these three samples is less than or equal to 0.01. The mean absolute differences are also very small. They are between 0.008 and 0.011. Correlation coefficients of the alternative design effects and the used design effects for these three samples are 1.00. Thus, these two methods are consistent in estimating design effects.

Table 9. Comparison of Two Design Effect Estimation Methods

Samples		CS8_1	S61	S41
Mean Absolute Difference between Two Methods		0.011	0.009	0.008
Correlation Coefficient of Two Methods		1.00	1.00	1.00
Number of Variables with differences between Two Methods	Absolute Difference between Two Methods:			
	0.00	2	9	8
	0.01	35	29	33
	0.02	3	3	
	0.03	1		

5.3 The Effect of Clustering Interval on Cost-Efficiency of Sample Designs.

The differences between design effects of different clustering intervals are small. Even in different significant clustering intervals, such differences of design effects are also small (see figure 3). Thus, it is difficult to use these samples for investigating the effect of clustering intervals on cost-efficiency of sampling designs.

Although this difficulty exists, the attempt in investigating the effect of clustering interval on cost-efficiency of sampling designs is required as it is of interest in selection of the alternative sampling designs with different clustering intervals. On the other hand, such investigation is also of interest in investigating the effect of clustering interval on design effect. Two samples **S41** and **CS4_2** are used to investigate such effect. From section 4.5, if:

*(travel-cost **CS4_2**)(design effect **CS4_2**)-(travel-cost **S41**)(design effect **S41**) > (interview-cost)(design effect **S41**)-(interview-cost)(design effect **CS4_2**),*

the sampling design of **S41** is better than that of **CS4_2**.

As the clustering interval of **CS4_2** is twice as much as that of **S41,** the travel cost of **CS4_2** is supposed to be twice as much as that of **S41**. The interview cost is the same in both designs. Thus, if:

*(travel-cost **S41**) > 0.9% (interview-cost),*

the sampling design of **S41** is better than that of **CS4_2**. That is, if the cost of travelling between two elements in **S41** is larger than 0.9% of the cost of an interview in the survey.

The cost of both travel and interview can be expressed in terms of the time which an interviewer spends on travel and interview. Thus, if the time an interviewer spends on travel from an interviewed household to next is more than 0.9% of the time he or she spends on an interview, **S41** is of a better sampling design.

Usually the travelling time is more than 0.9% of the interview time. Thus, whenever the effect of clustering interval is included in selection of alternative sampling

designs, the smallest clustering interval will be selected on the balance of the cost-efficiency.

6. DISCUSSION

An increase in clustering intervals seems to lead to little reduction in design effect. This is at odds with the finding of Laniel & Mohl (1994) who insisted that an increase in clustering interval led to an improvement in the efficiency of a sample design, that is, a reduction in design effect.

Within a given cluster size, increasing some of intervals between within-cluster elements, that is, an increase in mean clustering interval, does not lead to a reduction in either mean design effect of variables or an individual design effect. Even if the clustering interval is increased by a multiple, either the mean design effect or the individual design effect still tends to be unchanged.

On the other hand, within a given cluster size, the variability among individual design effects tends to be the same with either an increase in mean clustering interval or a multiple of increase in clustering interval.

Moreover, the results of t-tests indicate that differences of design effects between different clustering intervals are of very low significance level. Such low significance level seems to indicate the insignificance of the effect of clustering interval on design effect.

Attempting to improve sampling designs by increasing clustering interval seems pointless.

The result of investigating the effect of clustering interval on evaluating cost-efficiency of alternative sampling designs indicates that an improvement in the efficiency of sampling designs by an increase in clustering interval is so small that it can not be justified by the increase in cost.

Even if there is a larger improvement in the efficiency of

sampling designs by an increase in clustering interval, it seems that this improvement can not be justified by an increase in cost. Suppose a larger improvement in the efficiency of a sampling design is the reduction of 0.05 in design effect by doubling the clustering interval, that is, five times of the improvement in the efficiency of sampling design from **S41** to **CS4_2** in this study. Thus, the "breakeven" time of travelling between two elements is 4.5% of the time of an interview. If the time of an interview is 60 minutes, the breakeven time will be 2.7 minutes. As the time of travelling between two elements is usually larger than 2.7 minutes, the improvement is not justified by the increase in cost.

An increase in cluster size leads to a larger design effect. This is consistent with the finding of Laniel & Mohl (1994) who claimed that the efficiency of a sample design was reduced by an increase in cluster size.

For a given clustering interval, either mean design effect of variables or an individual design effect increases with an increase in cluster size. The mean design effect is likely to increase with an increase in cluster size following a straight line. Even in different clustering intervals, either mean design effect or an individual design effect also increases with an increase in cluster size.

On the other hand, the variability of design effect among individual variables is stronger with a larger cluster size, no matter what the clustering interval is.

Moreover, the results of t-tests indicate that differences between different cluster sizes are of high significance level. This high significance level seems to indicate the significance of the effect of cluster size on design effect.

The alternative method of estimating sampling variance,

that is, $design\ \ effect \approx mr^2$, is consistent with the method used in this study. Both the extremely high correlation coefficients and the extremely small mean absolute differences between design effects by these two methods indicate the validity of this alternative method.

This alternative method $design\ \ effect \approx mr^2$ is applicable for two-stage simple random sampling or one-stage simple random cluster sampling, if sum of squares between clusters and sum of square between elements in the population can be estimated and cluster size is nearly the same. However, like the method used in this study, this method is only used to estimate design effect for means.

7. CONCLUSION

The foregoing discussion leads to the following conclusions:

1. The effect of clustering interval has little influence on design effect. That is, design effect does not sensibly decrease with increase in clustering interval. This effect is consistent on the mean design effect of variables and the individual design effects. Such effect does neither strengthens nor weakens the variability of design effects among individual variables.

2. In face-to-face surveys, a sampling design with a smaller clustering interval is better than one with a larger clustering interval on the balance of the cost of data collection and the design effect.

3. The effect of cluster size has a substantial influence on design effect. This effect increases either the mean design effect of variables or the individual design effects by increasing cluster size, no matter what the clustering interval is. This effect also strengthens the variability of design effects among variables.

4. The formula $design\ \ effect \approx mr^2$ is applicable for estimating design effect for means in both two-stage simple random sampling and one-stage simple random cluster sampling with the same cluster size.

APPENDICES

Appendix A. Definition of Variables Selected

Table 10. Variables Selected (1)

Variables	Definition of variables
AGE	age of respondent.
	House ownership and mortgage.
YEARMOUT	how long ago took out mortgage.
REPAYTM	repayment term.
	How many copies have you read in the last 7 days?
MOR_PAPS	morning daily newspapers.
EVE_PAPS	evening daily newspapers.
	Have you read it in the last 7 days?
NTRUNTVE	weekly newspapers- New truth and TV extra.
SUN_NEWS	weekly news papers.
SUN_STAR	weekly news papers-Sunday star times.
NEW_IDEA	weekly magazines-New idea.
LTV_N_RT	weekly magazines-Listener TV and radio times.
NZ_W_WK	weekly magazines-New Zealand women weekly.
TIME	weekly magazines-Time.
TV_GUIDE	weekly magazines-TV guide.
WOM_DAY	weekly magazines-Woman's day.
E_W_WKLY	weekly magazines-English woman's weekly.

Table 10. Variables Selected (2)

Variables	Definition of variables
	Have you read it in the last four weeks?
AIRNZPW	monthly magazines-Air NZ Pacific Wave.
BOAT_NZ	monthly magazines-Boating NZ.
CLEO	monthly magazines-Cleo.
NZ_BUSI	monthly magazines-NZ Business.
AUS_W_WK	monthly magazines-Australian Woman's Weekly.
GRAPEVIN	monthly magazine-Grapevine.
H_BLUEBK	monthly maganzines-Harcout's Blue Book.
METRO	monthly magazines-Metro.
SHENMORE	monthly magazines-She & More.
NTH_N_SH	monthly magazines-North and South.
R_DIGEST	monthly magazines-Readers Digest.
TEARAWAY	monthly magazines-Tearaway.
	Have you read it in the last two months?
H_N_BUIL	bi-monthly mags.-Home & Building.
LIT_TREA	bi-monthly mags.-Little Treasures.
ADVENTUR	bi-monthly mags.-Adventure Magazines.
CUISINE	bi-monthly mags.-Cuisine.
	Have you read it in the last three months?
FASHIONQ	quarterly mags.-Fashion Quarterly.
NZ_GEOGR	quarterly mags.-NZ Geographic.
STYLE	quarterly mags.-Style.

Table 10. Variables Selected (3)

Variables	Definition of variables
	How much do you like reading it?
B_NEWS	business news.
SP_NEWS	sports news and results.
NEWSCOMM	news and commentary about other parts of the world.
ACCIDENT	news about accidents or crime.
LNEWSCOM	local news and commentary.
POLICTIC	political news.
	How often do you do that in the last 12 months?
R_ACTIVI	attended church or religious activities.
WORKHOME	brought work home to complete.
DINEROUT	dined at a restaurant or brasserie.
DRINKOUT	gone to a hotel, club or bar for a drink.
H_KIDSTD	helped your children with their school work.
P_YOUTH	talked about problems of youth or education.
WINEMEAL	had wine with a meal.
VISITCLU	gone to a club or nightclub.
WOMANISS	talked about women's issue.
VISITGAL	visited an art gallery or museum.
H_IMPROV	have undertaken home alterations or improvements.
CONCERT	gone to live theatre or classical music.
BOOK	bought yourself a paper back or hard back book.

Table 10. Variables Selected (4)

Variables	Definition of variables
	How much do you agree or disagree it?
TOUGHLAW	the law should be tougher on law breakers.
SUCCESS	success is very important to me.
KEEPFIT	I exercise regularly to keep fit.
MARIJUAN	smoking marijuana should be allowed.
RAWDEAL	I generally get a raw deal out of life.
C_WOMAN	it is important that a woman should have a career.
ENDSMEET	I find it hard to make ends meet.
FAMILYTO	as a family we spend a lot of time together.
H_FOOD	I try to avoid foods that are unhealthy.
WELLBRAN	I mostly buy well-known brand names.
NEWIDEAS	I am attracted to new ideas.
RACEPROB	racial problems are getting worse.
MONEYMAT	I find it easy to deal with money matters.
LONELY	I often feel quite lonely.
NEWPRODU	I like to try new and different household products.
NONPOLLU	I try to buy household products that won't pollute the environment.
W_FAIRGO	women do get fair go in NZ.
CHANGING	everything in NZ is changing too fast.
MYOPINIO	other people take my opinions seriously.
SPECIAL	I shop a lot for specials and bargains.
MAORICUL	we should make sure NZ keeps its maori culture.
TRUSTFUL	you don't know who to trust these days.

Table 10. Variables Selected (5)

Variables	Definition of variables
	Do you do that?
MMADECIG	products ever smoked-ready made cigarettes.
OWNCIGAR	products ever smoked-roll your own cigarettes.
CIGARS	products ever smoked-cigars, cigarellos.
PIPE	products ever smoked-pipe.
N_MMCIGA	number of ready made cigarettes you smoke per day.
N_OWNCIG	number of roll your own you smoke per day.

Appendix B. Formation of Simulated Samples

The detailed procedures to form simulated samples are as follows:

Sample with cluster size 8:

CS8_1: this is the original sample. That is, 8 elements each cluster are selected from every 3rd household within sampling areas. These 8 elements are the basis for forming the other simulation samples.

Samples with cluster size 6:

S61: formed by the selection of the first 6 elements within clusters, that is, 6 households with the least number each cluster are included;

S62: formed by dropping the 2nd and 4th of the 8 elements within clusters;

S63: formed by dropping the 3rd and 5th of the 8 elements within clusters;

S64: formed by dropping the 4th and 6th of the 8 elements within clusters;

S65: formed by dropping the 5th and 7th of the 8 elements within clusters.

Samples with cluster size 4:

S41: formed by the selection of the first 4 elements within clusters;

CS4_2: formed by dropping the 2nd, 4th, 6th and 8th of the 8 elements within clusters;

S42: formed by dropping the 1st, 3rd, 5th and 7th of the 8 elements within clusters;

S43: formed by dropping the 3rd, 4th, 6th and 7th of the 8 elements within clusters;

S44: formed by dropping the 1st, 2nd, 5th and 6th of the 8 elements within clusters.

Samples with cluster size 2:

CS2_1, CS2_2, CS2_3, CS2_4, CS2_5, CS2_6, and **CS2_7:** these samples are formed by the first element and each of the others from the 8 elements each cluster.

Appendix C. the Mathematical Derivation of $design\ \ effect = \frac{ms_b}{ms}$

The mathematical derivation of $design\ \ effect = \frac{ms_b}{ms}$ is as follows:

From formula 2.2 and formula 2.4,

$$\hat{\sigma}^2_{\bar{Y}} = \frac{S_B^2}{k} = \frac{S^2}{km}\left[1+(m-1)\hat{\rho}\right] = \frac{S^2}{n}\left[1+(m-1)\hat{\rho}\right].$$

Then,

$$\frac{\frac{S_B^2}{k}}{\frac{S^2}{n}} = \left[1+(m-1)\hat{\rho}\right]$$

$$= \frac{\frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k(k-1)}}{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n(n-1)}}$$

$$= \frac{\frac{\sum_{j=1}^{k}(\bar{Y}_j - \bar{Y})^2}{k(k-1)}}{\frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \bar{Y})^2}{km(km-1)}}$$

$$= \frac{\frac{\sum_{j=1}^{k}(\overline{Y}_j - \overline{Y})^2}{k-1}}{\frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \overline{Y})^2}{km-1}} * \frac{km}{k}$$

$$= \frac{\frac{\sum_{j=1}^{k} m(\overline{Y}_j - \overline{Y})^2}{k-1}}{\frac{\sum_{j=1}^{k}\sum_{i=1}^{m}(Y_{ji} - \overline{Y})^2}{km-1}}$$

$$= \frac{\frac{SS_B}{k-1}}{\frac{SS}{km-1}}$$

$$= \frac{\frac{SS_B}{k-1}}{\frac{SS}{n-1}}$$

$$= \frac{ms_b}{ms}.$$

From formula 3.0, the method of estimating design effect is:

$$Design \quad effect = \left[1 + (m-1)\hat{\rho}\right]$$

$$= \frac{\frac{S_B^2}{k}}{\frac{S^2}{n}}$$

$$= \frac{ms_b}{ms}.$$

Appendix D. Design Effects of Variables in Different Clusterings

Table 11. Design Effects of Variables in Different Clusterings (1)

Cluster Size	8	6						4					2					
Clusterings*	CS8_1	S61	S62	S63	S64	S65	CS4_2	S41	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
STYLE	.88	.92	.90	.90	.92	.90	.96	.96	.94	.94	.96	.98	.98	.98	.98	.97	.98	.97
SHENMORE	.93	.86	1.00	1.06	.92	1.01	.87	.91	1.00	1.01	.91	.96	.96	.96	.94	.95	.97	1.22
TEARAWAY	.96	.80	1.01	1.08	.98	.83	1.01	.90	.93	.90	1.08	.98	.95	.99	.96	.97	.97	.97
LTV_N_RT	.97	.92	.88	1.07	.96	.93	.84	.96	1.02	1.09	1.00	1.03	.97	1.01	1.01	.87	.96	1.02
YEAROUT	.99	.93	1.04	.97	1.05	.96	1.01	.93	1.00	.99	.96	1.00	.97	.97	1.02	1.01	.96	1.00
NEW_IDEA	.99	.99	1.05	.98	1.04	.93	1.25	1.03	.97	.94	1.01	1.04	1.09	1.00	1.02	.97	1.17	1.02
BOAT_NZ	1.00	.93	.93	1.07	1.02	.93	.96	.96	.95	.93	.96	.98	1.00	1.00	.99	1.01	.99	.99
AUS_W_WK	1.00	1.04	.93	1.02	.96	1.10	.92	1.07	1.15	1.05	1.01	.96	1.05	.97	.92	1.03	1.03	1.05
W_FAIRGO	1.01	1.02	1.00	.99	1.09	.97	1.13	.97	1.04	1.11	.81	1.13	1.05	.88	1.02	1.01	.96	1.10
H_N_BUIL	1.01	1.01	1.07	.96	1.09	.95	1.23	.98	.93	1.04	.74	.97	1.01	1.00	1.08	.99	1.15	1.05
R_DIGEST	1.02	1.17	1.06	.93	.99	1.12	1.11	1.15	1.03	.95	.96	1.01	1.11	1.07	.97	1.04	.97	.97
FASHIONQ	1.03	1.11	1.01	.93	.97	.92	.97	1.06	.92	.93	.92	.94	1.22	1.04	.92	.94	.94	.92
SUN_STAR	1.03	.94	1.12	.89	1.05	1.07	.94	.97	.97	1.05	1.12	.86	1.11	.99	.96	.99	.89	1.06
CIGARS	1.04	1.00	.93	1.04	1.09	1.17	.93	1.12	1.13	1.04	.	.96	.	.	.95	.	.	.
ACCIDENT	1.05	1.07	1.07	.93	1.01	.98	.88	.88	.91	.98	1.01	.86	1.00	1.03	.88	.99	.84	1.05
PIPE	1.05	.97	1.03	1.08	1.16	1.09	1.11	1.04	1.13	1.04	.93	.96	.94	.95	.94	.95	1.95	.95
NZ_W_WK	1.07	1.10	1.05	1.03	1.13	1.07	1.25	1.07	.95	1.07	1.00	.95	1.08	1.15	1.10	1.03	1.18	1.11
WOMAN_WK	1.07	.95	.92	1.11	1.11	1.13	1.05	1.11	1.23	1.18	1.06	1.02	.98	1.05	.96	1.04	.99	1.02
MONEYMAT	1.07	1.11	.87	1.16	1.00	1.16	.99	1.09	1.21	1.10	.84	1.00	.90	.97	.98	1.15	1.12	1.10
E_W_WKLY	1.08	1.09	1.09	1.02	1.04	.91	.96	.95	.93	1.15	.93	.98	1.00	.99	.99	1.00	.98	.99
H_KIDSTD	1.09	1.07	1.20	.97	1.14	1.11	1.10	1.10	1.01	1.06	.90	1.08	1.13	.87	1.00	1.13	.95	1.12

Table 11. Design Effects of Variables in Different Clusterings (2)

Cluster Size	8	6						4					2					
Clustering*	CS8_1	S61	S62	S63	S64	S65	CS4_2	S41	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
R_ACTIVI	1.10	1.07	1.20	1.01	1.33	1.09	1.20	1.15	.95	1.12	.94	1.06	1.17	.99	1.05	1.02	1.12	.95
WELLBRAN	1.10	1.08	1.00	1.04	1.14	1.23	1.02	1.12	1.24	1.25	1.06	1.18	1.10	.98	1.09	.82	.90	.98
CUISINE	1.10	1.06	.99	1.15	1.03	1.07	1.01	.93	1.06	1.01	.94	.97	.98	.97	.95	.96	1.27	1.32
RAWDEAL	1.10	1.14	1.00	1.01	1.10	1.13	1.06	1.22	.91	.96	1.08	.93	1.05	1.03	.83	.97	1.14	1.15
METRO	1.12	1.12	1.08	1.23	1.06	1.27	.95	1.14	1.13	1.18	.94	1.17	.96	1.13	.95	.96	.96	1.30
POLICTIC	1.13	1.09	1.18	.89	1.15	1.05	1.13	.95	.99	1.03	1.12	.89	1.02	1.02	1.28	.95	.93	1.09
MAORICUL	1.13	1.04	1.26	1.03	1.11	.93	1.17	.95	.93	1.06	1.02	1.00	1.05	.92	1.15	.99	1.07	.92
NZ_BUSI	1.15	.99	1.10	1.17	1.13	1.19	.91	1.05	1.12	1.05	1.10	.96	.97	.97	.97	.98	.96	1.23
AIRNZPW	1.15	1.06	1.10	1.10	1.29	.96	1.08	1.05	1.00	1.35	1.01	1.17	.96	.97	1.11	.96	1.17	.94
SUN_NEWS	1.16	1.15	1.22	1.02	1.21	1.14	1.21	1.09	1.04	1.24	1.04	1.10	1.19	.96	1.19	1.06	1.03	1.01
ADVENTUR	1.17	1.40	1.09	1.17	.93	1.23	.97	1.25	1.17	.95	.96	.98	.99	.98	.99	1.32	.98	.97
WOMANISS	1.20	1.18	1.14	1.07	1.25	1.16	1.20	1.28	1.15	1.18	1.02	1.13	1.16	.93	1.05	.93	.82	1.02
CLEO	1.21	1.16	1.27	.96	1.04	1.13	.99	.99	1.10	.95	1.18	.96	.96	.95	.94	.95	.97	.95
ENDSMEET	1.21	1.12	1.16	1.12	1.20	1.38	1.01	1.18	1.10	1.11	1.08	1.17	1.21	.92	.86	.98	.96	1.17
LONELY	1.21	1.19	1.20	1.19	1.23	1.22	1.27	1.22	1.11	1.01	1.18	1.00	1.09	1.07	1.11	.94	1.16	1.00
NTH_N_SH	1.23	1.29	1.29	1.19	1.02	1.22	1.13	1.01	1.13	1.07	.92	1.03	1.04	.95	1.10	1.13	1.03	1.20
NZ_GEOGR	1.23	1.13	1.21	1.17	1.27	1.21	1.16	1.26	1.18	1.28	.97	1.05	1.18	1.09	1.09	.92	.94	1.15
TV_GUIDE	1.23	1.23	1.08	1.21	1.09	1.22	.96	1.15	1.13	1.02	1.20	1.01	.93	1.14	1.05	1.15	.90	.98
FAMILYTO	1.24	1.15	1.18	1.22	1.08	1.29	1.19	1.12	1.33	.99	1.11	.78	1.15	.96	1.10	.97	1.04	1.12
N_MMCIGA	1.24	1.24	1.09	1.24	1.15	1.32	1.07	1.25	1.25	1.14	1.02	1.10	1.12	1.01	1.02	1.20	.97	.99
C_WOMAN	1.25	1.16	1.11	1.17	1.26	1.18	1.09	1.05	1.14	1.25	1.01	1.12	1.22	.95	1.09	1.03	.98	1.01

Table 11. Design Effects of Variables in Different Clusterings (3)

Cluster Size	8	6						4					2					
Clustering*	CS8_1	S61	S62	S63	S64	S65	CS4_2	S41	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
NTRUNTVE	1.25	1.28	1.17	.99	1.12	1.02	1.24	1.17	.93	1.12	1.36	.99	.98	.97	1.27	.98	.97	.97
NEWPRODU	1.26	1.11	1.27	1.09	1.26	.98	1.12	1.01	.95	1.23	1.16	1.09	1.05	.91	1.05	.92	1.01	1.03
MMADECIG	1.27	1.36	1.16	1.19	1.30	1.30	1.17	1.35	1.08	1.27	.98	1.21	1.13	1.10	1.13	1.18	1.06	.96
LIT_TREA	1.28	1.12	.88	1.28	1.23	1.27	.89	.98	1.49	1.27	1.06	.93	.96	.95	.94	.92	.95	1.07
H_IMPROV	1.29	1.13	1.31	1.31	1.36	1.13	1.13	1.10	1.07	1.39	.89	1.27	.97	1.00	1.01	1.08	1.09	1.23
NEWSCOMM	1.29	1.16	1.14	.99	1.12	1.26	1.06	1.11	1.10	1.05	1.28	1.00	1.06	.97	.91	1.03	.86	1.08
TIME	1.29	1.22	1.16	1.18	1.23	1.22	1.12	1.19	1.11	1.20	.99	1.15	1.08	1.09	1.03	.95	1.17	1.03
SP_NEWS	1.30	1.23	1.37	1.00	1.38	1.01	1.51	1.07	.81	1.20	1.10	1.06	1.22	.91	1.13	1.12	1.16	1.15
OWNCIGAR	1.31	1.09	1.15	1.38	1.15	1.27	1.01	1.09	1.30	1.13	1.23	1.00	.93	1.22	.92	.83	.93	1.27
BOOK	1.32	1.18	1.30	1.14	1.29	1.16	1.27	1.16	.99	1.35	1.05	1.22	1.10	1.05	1.15	1.01	.97	1.15
SPECIAL	1.32	1.19	1.14	1.28	1.25	1.21	1.10	1.22	1.09	1.16	1.36	1.28	.97	.93	.99	.93	1.05	1.04
NONPOLLU	1.32	1.37	1.23	1.10	1.31	1.07	1.24	1.22	.85	1.40	.98	1.14	1.11	1.15	1.19	1.06	1.07	1.19
KEEPFIT	1.33	1.33	1.10	1.30	1.15	1.33	1.10	1.34	1.13	1.02	1.16	1.07	1.07	1.06	.95	1.05	1.11	1.15
N_OWNCIG	1.34	1.06	1.17	1.44	1.19	1.28	.99	1.09	1.26	1.22	1.25	1.03	.92	1.23	.90	.89	.97	1.36
SUCCESS	1.34	1.23	1.34	1.23	1.46	1.34	1.12	1.21	1.09	1.24	1.25	1.04	1.09	1.12	1.17	1.01	1.04	1.08
REPAYTM	1.38	1.39	1.38	1.34	1.29	1.49	1.13	1.32	1.23	1.29	1.02	1.32	1.21	1.06	1.00	1.15	.95	1.23
LNEWSCOM	1.40	1.21	1.20	1.36	1.28	1.26	1.21	1.19	1.12	1.09	1.28	1.03	.94	1.11	1.08	1.22	1.16	1.02
MYOPINIO	1.41	1.28	1.30	1.39	1.24	1.36	1.12	1.07	1.40	1.19	1.12	1.07	1.00	1.12	1.03	1.03	1.11	.97
CHANGING	1.42	1.18	1.40	1.18	1.50	1.26	1.24	1.03	1.08	1.34	1.20	1.21	1.13	.93	1.03	1.05	.98	1.10
MARIJUAN	1.43	1.28	1.22	1.31	1.34	1.26	1.28	1.23	1.27	.95	1.32	.91	.98	1.08	.98	.81	1.09	.94
RACEPROB	1.44	1.29	1.36	1.26	1.38	1.19	1.20	1.19	1.05	1.15	1.17	1.08	1.17	1.14	1.03	1.05	.98	1.02

Table 11. Design Effects of Variables in Different Clusterings (4)

Cluster Size	8	6						4					2					
Clustering*	CS8_1	S61	S62	S63	S64	S65	CS4_2	S41	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
H_FOOD	1.46	1.43	1.44	1.21	1.31	1.38	1.25	1.24	1.14	1.18	1.07	1.07	1.10	1.11	1.17	1.05	1.07	1.09
NEWIDEAS	1.47	1.34	1.37	1.31	1.30	1.23	1.27	1.16	1.07	1.22	1.04	1.12	1.25	1.10	1.00	1.01	.96	1.08
B_NEWS	1.53	1.41	1.35	1.25	1.37	1.35	1.31	1.20	1.14	1.25	1.05	1.11	1.11	1.03	1.13	1.10	1.15	1.10
DRINKOUT	1.55	1.51	1.39	1.32	1.33	1.45	1.29	1.29	1.25	1.17	1.36	1.15	1.07	1.23	1.11	.98	.95	1.04
P_YOUTH	1.55	1.42	1.44	1.37	1.32	1.47	1.17	1.26	1.21	1.15	1.29	1.17	1.12	.98	1.01	1.14	1.00	.97
AGE	1.55	1.57	1.33	1.34	1.54	1.51	1.42	1.62	1.16	1.20	1.17	1.28	1.19	1.20	1.17	.94	1.13	1.03
WORKHOME	1.57	1.54	1.27	1.42	1.36	1.52	1.17	1.33	1.38	1.41	1.18	1.20	1.05	1.06	1.09	.98	.97	1.25
VISITGAL	1.61	1.52	1.47	1.38	1.44	1.38	1.31	1.26	1.13	1.44	1.06	1.25	1.15	1.15	1.15	1.19	1.16	.99
GRAPEVIN	1.66	1.41	1.59	1.40	1.65	1.56	1.47	1.29	1.39	1.32	1.24	.95	1.36	1.04	1.16	1.11	1.07	1.08
DINEROUT	1.68	1.45	1.63	1.36	1.54	1.49	1.14	1.21	1.23	1.54	1.18	1.19	1.01	1.05	1.17	1.05	1.00	1.23
TRUSTFUL	1.73	1.73	1.46	1.48	1.51	1.57	1.33	1.40	1.29	1.44	1.19	1.25	1.16	1.12	1.18	1.20	1.01	1.23
TOUGHLAW	1.73	1.55	1.56	1.43	1.56	1.43	1.37	1.34	1.29	1.27	1.33	1.12	1.26	1.00	1.07	1.04	1.00	1.04
CONCERT	1.74	1.52	1.59	1.44	1.56	1.48	1.29	1.25	1.21	1.44	1.27	1.11	1.04	1.23	1.11	1.08	.99	1.23
VISITCLU	1.75	1.69	1.53	1.50	1.53	1.52	1.54	1.27	1.49	1.34	1.23	1.02	1.18	1.11	1.23	1.12	1.14	.99
H_BLUEBK	1.90	1.92	1.78	1.54	1.69	1.96	1.30	1.46	1.42	1.52	1.16	1.15	1.44	1.04	1.16	1.40	.93	1.28
WINEMEAL	1.95	1.64	1.72	1.64	1.78	1.59	1.41	1.40	1.28	1.64	1.44	1.24	1.05	1.12	1.26	1.13	1.13	1.20
MOR_PAPS	2.42	2.12	1.91	1.90	1.99	2.04	1.48	1.67	1.56	1.61	1.59	1.23	1.13	1.14	1.18	1.20	1.01	1.11
EVE_PAPS	3.91	3.05	3.19	3.06	3.03	3.08	2.30	2.19	2.20	2.15	2.33	1.25	1.41	1.39	1.40	1.44	1.54	1.55
Average	**1.32**	**1.24**	**1.23**	**1.21**	**1.25**	**1.24**	**1.15**	**1.16**	**1.13**	**1.18**	**1.11**	**1.07**	**1.08**	**1.04**	**1.05**	**1.03**	**1.04**	**1.08**

* See Chapter 4: METHOD.

Appendix E. Homogeneity

1. Calculation Method of Homogeneity

It is cumbersome to calculate the measurement of homogeneity ρ by its definition. In practice, the definition of ρ is never used to calculate it. Thus, it is necessary to search for a simple solution to calculating ρ. The most common solution is to estimate ρ if cluster size is known and design effect can be estimated. That is, ρ is estimated by:

$$\hat{\rho} = (\text{design effect} - 1) / (m - 1),$$

where m is cluster size (average cluster size).

This formula is derived from formula 3.0. From formula 3.0,

$$\text{Design effect} - 1 = (m - 1)\ \hat{\rho}.$$

Then,

$$\hat{\rho} = (\text{design effect} - 1) / (m - 1).$$

2. Results

All values of homogeneity for each variable estimated from 18 samples are displayed in table 12. These values are between -0.215 and 1.000. The observations from these values are described in the following.

First, the average homogeneity over variables tends to be the same among clusterings, though there is variability. This is reflected in figure 6 from which the curve of the average homogeneity values over variables across these 18 samples varies around a horizontal line. This is also reflected from table 12. The average homogeneity values for each cluster size are 0.045 for cluster size 8, 0.047 for cluster size 6, 0.048 for cluster size 4, and 0.055 for cluster size 2.

Table 12. Homogeneity across Variables and Clusterings (1)

Cluster Size	8	6					4					2						
Clusterings	CS8_1	S61	S62	S63	S64	S65	S41	CS4_2	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
STYLE	-0.017	-0.016	-0.019	-0.020	-0.016	-0.020	-0.015	-0.013	-0.020	-0.020	-0.015	-0.019	-0.019	-0.024	-0.019	-0.026	-0.019	-0.032
SHENMORE	-0.010	-0.028	0.000	0.011	-0.015	0.002	-0.031	-0.043	0.000	0.002	-0.031	-0.036	-0.044	-0.036	-0.064	-0.055	-0.032	0.222
TEARAWAY	-0.006	-0.039	0.003	0.017	-0.004	-0.034	-0.033	0.002	-0.022	-0.033	0.028	-0.019	-0.055	-0.010	-0.041	-0.026	-0.032	-0.026
LTV_N_RT	-0.005	-0.017	-0.024	0.014	-0.008	-0.014	-0.013	-0.054	0.006	0.031	-0.001	0.030	-0.034	0.010	0.010	-0.127	-0.043	0.021
YEAROUT	-0.002	0.004	0.000	-0.003	0.018	-0.005	-0.010	0.033	0.014	0.036	-0.063	0.107	0.052	-0.120	0.016	0.013	-0.070	0.105
NEW_IDEA	-0.002	-0.014	0.008	-0.007	0.010	-0.007	-0.022	0.004	-0.001	-0.005	-0.014	0.003	-0.030	-0.025	0.022	0.007	-0.040	0.003
BOAT_NZ	-0.002	-0.001	0.010	-0.005	0.008	-0.014	0.011	0.083	-0.009	-0.018	0.002	0.042	0.089	0.001	0.020	-0.025	0.169	0.020
AUS_W_WK	0.000	-0.015	-0.015	0.014	0.004	-0.015	-0.015	-0.015	-0.018	-0.024	-0.015	-0.024	-0.001	-0.001	-0.010	0.006	-0.010	-0.010
W_FAIRGO	0.001	0.008	-0.014	0.005	-0.008	0.021	0.022	-0.025	0.049	0.016	0.004	-0.037	0.051	-0.031	-0.082	0.029	0.029	0.051
H_N_BUIL	0.002	0.003	0.015	-0.009	0.018	-0.011	-0.005	0.078	-0.023	0.012	-0.087	-0.034	0.010	0.001	0.076	-0.008	0.154	0.053
R_DIGEST	0.002	0.035	0.013	-0.015	-0.003	0.024	0.050	0.037	0.010	-0.017	-0.013	0.010	0.110	0.074	-0.034	0.042	-0.025	-0.034
FASHIONQ	0.004	0.023	0.002	-0.014	-0.006	-0.017	0.020	-0.011	-0.026	-0.023	-0.026	-0.060	0.222	0.042	-0.084	-0.064	-0.060	-0.080
SUN_STAR	0.004	-0.012	0.024	-0.022	0.011	0.015	-0.009	-0.021	-0.008	0.016	0.039	-0.143	0.110	-0.010	-0.042	-0.010	-0.113	0.064
CIGARS	0.006	0.001	-0.015	0.009	0.020	0.036	0.041	-0.026	0.044	0.015	0.000	-0.043	.	.	-0.061	.	.	.
ACCIDENT	0.007	0.013	0.014	-0.013	0.003	-0.004	-0.039	-0.040	-0.029	-0.007	0.005	-0.142	0.000	0.028	-0.121	-0.007	-0.160	0.053
PIPE	0.008	-0.006	0.007	0.017	0.034	0.020	0.015	0.039	0.044	0.015	-0.024	-0.043	-0.073	-0.058	-0.069	-0.051	1.000	-0.061
NZ_W_WK	0.010	0.019	0.010	0.006	0.025	0.013	0.022	0.082	-0.016	0.022	0.001	-0.051	0.080	0.154	0.101	0.026	0.176	0.108
WOMAN_WK	0.010	-0.009	-0.016	0.022	0.022	0.025	0.036	0.018	0.076	0.059	0.020	0.019	-0.024	0.054	-0.044	0.038	-0.013	0.018
MONEYMAT	0.010	0.014	0.040	-0.006	0.028	0.023	0.033	0.034	0.005	0.019	-0.034	0.093	0.096	-0.154	0.004	0.104	-0.039	0.141
E_W_WKLY	0.010	0.022	-0.025	0.031	0.000	0.031	0.030	-0.004	0.069	0.032	-0.052	-0.002	-0.097	-0.030	-0.023	0.151	0.115	0.100
H_KIDSTD	0.011	0.018	0.018	0.004	0.007	-0.018	-0.018	-0.015	-0.022	0.048	-0.022	-0.019	-0.001	-0.010	-0.009	-0.001	-0.019	-0.010
R_ACTIVI	0.012	0.017	0.001	0.008	0.029	0.046	0.041	0.008	0.079	0.085	0.020	0.171	0.101	-0.015	0.095	-0.180	-0.104	-0.017

Table 12. Homogeneity across Variables and Clusterings (2)

Cluster Size	8	6					4					2						
Clusterings	CS8_1	S61	S62	S63	S64	S65	S41	CS4_2	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
WELLBRAN	0.014	0.011	-0.002	0.030	0.007	0.014	-0.024	0.005	0.021	0.005	-0.020	-0.026	-0.024	-0.026	-0.050	-0.041	0.270	0.323
CUISINE	0.015	0.018	0.036	-0.022	0.031	0.011	-0.016	0.045	-0.002	0.010	0.041	-0.137	0.000	0.000	0.284	-0.062	-0.087	0.076
RAWDEAL	0.016	0.008	0.052	0.006	0.022	-0.014	-0.018	0.049	-0.024	0.020	0.007	0.002	0.038	-0.081	0.145	-0.007	0.045	-0.078
METRO	0.017	0.013	0.039	0.003	0.066	0.018	0.049	0.073	-0.017	0.040	-0.021	0.055	0.181	-0.009	0.045	0.019	0.124	-0.050
POLICTIC	0.017	0.025	0.016	0.046	0.012	0.054	0.048	-0.016	0.042	0.059	-0.019	0.168	-0.041	0.126	-0.055	-0.036	-0.036	0.304
MAORICUL	0.018	0.028	0.000	0.001	0.021	0.026	0.072	0.020	-0.029	-0.014	0.026	-0.027	0.069	0.121	-0.177	-0.074	0.098	0.103
NZ_BUSI	0.021	-0.002	0.021	0.034	0.026	0.039	0.015	-0.028	0.041	0.015	0.033	-0.041	-0.032	-0.026	-0.026	-0.024	-0.036	0.227
AIRNZPW	0.022	0.011	0.019	0.020	0.059	-0.008	0.015	0.028	0.000	0.115	0.005	0.168	-0.036	-0.032	0.108	-0.041	0.168	-0.060
SUN_NEWS	0.022	0.031	0.044	0.005	0.042	0.028	0.028	0.069	0.012	0.079	0.015	0.097	0.188	-0.037	0.195	0.062	0.030	0.009
ADVENTUR	0.024	0.079	0.018	0.034	-0.015	0.045	0.084	-0.010	0.058	-0.018	-0.015	-0.019	-0.010	-0.024	-0.010	0.323	-0.024	-0.032
WOMANISS	0.029	0.032	0.055	-0.008	0.008	0.027	-0.004	-0.004	0.033	-0.016	0.061	-0.041	-0.044	-0.050	-0.060	-0.055	-0.032	-0.050
CLEO	0.030	0.033	0.028	-0.002	0.024	0.052	0.037	0.010	0.034	0.017	0.092	0.003	0.061	-0.029	-0.086	0.005	-0.167	0.057
ENDSMEET	0.030	0.036	0.027	0.015	0.050	0.032	0.093	0.066	0.050	0.059	0.006	0.126	0.156	-0.065	0.053	-0.073	-0.185	0.008
LONELY	0.032	0.024	0.032	0.024	0.040	0.077	0.059	-0.004	0.033	0.036	0.027	0.172	0.196	-0.079	-0.136	-0.024	-0.043	0.172
NTH_N_SH	0.033	0.059	0.059	0.039	0.004	0.043	0.005	0.044	0.042	0.023	-0.026	0.030	0.042	-0.049	0.101	0.130	0.027	0.203
NZ_GEOGR	0.033	0.027	0.042	0.033	0.054	0.043	0.086	0.053	0.060	0.094	-0.009	0.052	0.181	0.092	0.092	-0.080	-0.064	0.145
TV_GUIDE	0.033	0.046	0.017	0.042	0.019	0.044	0.050	-0.014	0.043	0.008	0.067	0.008	-0.066	0.140	0.051	0.147	-0.096	-0.024
FAMILYTO	0.034	0.029	0.036	0.044	0.016	0.057	0.041	0.064	0.109	-0.004	0.036	-0.215	0.153	-0.043	0.100	-0.026	0.038	0.118
N_MMCIGA	0.035	0.047	0.017	0.048	0.030	0.063	0.082	0.025	0.084	0.047	0.005	0.105	0.116	0.014	0.021	0.200	-0.027	-0.006
C_WOMAN	0.035	0.031	0.021	0.035	0.053	0.036	0.017	0.031	0.047	0.085	0.002	0.139	0.217	-0.052	0.093	0.033	-0.020	0.015
NTRUNTVE	0.036	0.056	0.034	-0.002	0.025	0.004	0.058	0.079	-0.022	0.041	0.119	-0.009	-0.019	-0.026	0.270	-0.024	-0.026	-0.026

Table 12. Homogeneity across Variables and Clusterings (3)

Cluster Size	8	6					4					2						
Clustering	CS8_1	S61	S62	S63	S64	S65	S41	CS4_2	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
NEWPRODU	0.037	0.021	0.053	0.018	0.052	-0.004	0.003	0.039	-0.016	0.075	0.053	0.088	0.048	-0.085	0.045	-0.079	0.006	0.027
MMADECIG	0.038	0.037	0.060	0.029	0.058	0.031	0.053	0.085	-0.004	0.115	0.017	0.216	0.062	0.020	0.150	0.009	-0.035	0.145
LIT_TREA	0.038	0.039	0.041	0.038	0.047	0.043	0.075	0.093	0.036	0.004	0.061	-0.032	0.107	0.102	0.098	-0.053	0.178	0.011
H_IMPROV	0.039	0.045	0.074	0.001	0.077	0.001	0.025	0.167	-0.062	0.068	0.033	0.056	0.216	-0.095	0.133	0.107	0.140	0.150
NEWSCOMM	0.040	0.027	0.061	0.063	0.073	0.027	0.033	0.042	0.023	0.129	-0.035	0.265	-0.030	-0.034	0.013	0.081	0.094	-0.040
TIME	0.041	0.025	-0.025	0.056	0.046	0.053	-0.008	-0.038	0.162	0.091	0.021	-0.070	-0.041	-0.050	-0.060	-0.080	-0.050	0.230
SP_NEWS	0.041	0.076	0.034	0.039	0.063	0.064	0.126	0.061	0.030	0.097	-0.007	0.231	0.146	0.109	0.151	0.200	0.064	0.066
OWNCIGAR	0.042	0.045	0.032	0.035	0.046	0.045	0.064	0.040	0.035	0.067	-0.002	0.145	0.079	0.092	0.030	-0.055	0.168	0.030
BOOK	0.046	0.073	0.046	0.020	0.062	0.015	0.072	0.080	-0.050	0.134	-0.006	0.142	0.114	0.151	0.194	0.064	0.069	0.307
SPECIAL	0.046	0.065	0.021	0.060	0.029	0.067	0.112	0.035	0.043	0.008	0.053	0.069	0.070	0.097	-0.051	0.054	0.109	0.187
NONPOLLU	0.047	0.019	0.032	0.080	0.033	0.058	0.031	0.005	0.106	0.047	0.082	-0.004	-0.076	0.250	-0.090	-0.187	-0.080	0.154
KEEPFIT	0.049	0.012	0.035	0.088	0.037	0.057	0.029	-0.005	0.087	0.074	0.083	0.031	-0.079	0.230	-0.099	-0.108	-0.030	0.358
N_OWNCIG	0.051	0.039	0.027	0.056	0.050	0.043	0.075	0.035	0.029	0.054	0.121	0.237	-0.030	-0.019	-0.013	-0.068	0.054	0.048
SUCCESS	0.055	0.078	0.077	0.069	0.058	0.100	0.110	0.045	0.076	0.099	0.006	0.330	0.214	0.058	0.005	0.154	-0.056	0.232
REPAYTM	0.056	0.047	0.068	0.047	0.093	0.069	0.070	0.075	0.029	0.082	0.085	0.054	0.161	0.090	0.187	-0.023	0.054	0.107
LNEWSCOM	0.058	0.041	0.041	0.072	0.055	0.053	0.063	0.072	0.040	0.032	0.095	0.030	-0.065	0.112	0.079	0.224	0.163	0.018
MYOPINIO	0.058	0.036	0.081	0.037	0.100	0.052	0.011	0.077	0.026	0.113	0.068	0.208	0.111	-0.073	0.031	0.050	-0.018	0.095
CHANGING	0.059	0.057	0.061	0.077	0.047	0.072	0.022	0.042	0.132	0.062	0.039	0.080	-0.026	0.103	0.035	0.005	0.128	-0.026
MARIJUAN	0.061	0.058	0.073	0.053	0.076	0.039	0.064	0.068	0.018	0.050	0.058	0.085	0.169	0.139	0.033	0.052	-0.016	-0.008
RACEPROB	0.066	0.086	0.089	0.043	0.061	0.076	0.081	0.084	0.046	0.060	0.024	0.066	0.099	0.109	0.168	0.051	0.068	0.089
H_FOOD	0.066	0.083	0.071	0.051	0.075	0.071	0.066	0.085	0.046	0.083	0.017	0.144	0.119	0.010	0.101	0.108	0.094	0.105

Table 12. Homogeneity across Variables and Clusterings (4)

Cluster Size	8	6					4					2						
Clustering	CS8_1	S61	S62	S63	S64	S65	S41	CS4_2	S42	S43	S44	CS2_1	CS2_2	CS2_3	CS2_4	CS2_5	CS2_6	CS2_7
Variables:																		
NEWIDEAS	0.067	0.057	0.043	0.062	0.068	0.052	0.078	0.103	0.089	-0.015	0.109	-0.101	-0.025	0.113	-0.016	-0.165	0.125	-0.065
B_NEWS	0.070	0.068	0.074	0.062	0.061	0.047	0.055	0.090	0.024	0.074	0.014	0.114	0.247	0.098	-0.002	0.010	-0.042	0.080
DRINKOUT	0.078	0.084	0.088	0.073	0.064	0.094	0.088	0.058	0.071	0.050	0.096	0.166	0.119	-0.020	0.009	0.136	-0.002	-0.027
P_YOUTH	0.079	0.113	0.067	0.069	0.107	0.103	0.207	0.140	0.054	0.067	0.057	0.283	0.193	0.200	0.171	-0.064	0.132	0.025
AGE	0.082	0.102	0.078	0.063	0.066	0.089	0.096	0.098	0.084	0.056	0.119	0.147	0.071	0.228	0.092	-0.016	-0.048	0.031
WORKHOME	0.082	0.108	0.055	0.083	0.071	0.104	0.109	0.058	0.126	0.136	0.059	0.199	0.046	0.063	0.094	-0.021	-0.030	0.249
VISITGAL	0.084	0.105	0.094	0.075	0.089	0.077	0.088	0.104	0.045	0.146	0.019	0.252	0.155	0.078	0.146	0.185	0.158	0.055
GRAPEVIN	0.094	0.082	0.118	0.080	0.129	0.112	0.095	0.158	0.128	0.107	0.080	-0.055	0.364	0.043	0.163	0.108	0.066	0.079
DINEROUT	0.097	0.091	0.126	0.071	0.108	0.098	0.071	0.047	0.077	0.181	0.061	0.191	0.010	0.048	0.168	0.053	0.000	0.227
TRUSTFUL	0.103	0.103	0.117	0.088	0.111	0.096	0.085	0.096	0.071	0.146	0.090	0.108	0.037	0.240	0.106	0.083	-0.012	0.232
TOUGHLAW	0.104	0.146	0.092	0.096	0.102	0.115	0.134	0.108	0.096	0.147	0.065	0.250	0.179	0.122	0.178	0.200	0.010	0.226
CONCERT	0.104	0.109	0.113	0.086	0.111	0.087	0.112	0.124	0.096	0.089	0.109	0.124	0.262	-0.002	0.075	0.041	-0.002	0.036
VISITCLU	0.111	0.137	0.107	0.100	0.106	0.105	0.089	0.178	0.164	0.115	0.078	0.019	0.178	0.136	0.230	0.123	0.145	0.006
H_BLUEBK	0.129	0.184	0.155	0.107	0.138	0.192	0.153	0.100	0.139	0.172	0.053	0.145	0.437	0.042	0.163	0.404	-0.075	0.283
WINEMEAL	0.136	0.128	0.144	0.128	0.156	0.117	0.132	0.137	0.092	0.213	0.147	0.238	0.054	0.125	0.259	0.129	0.131	0.199
MOR_PAPS	0.203	0.223	0.181	0.180	0.198	0.209	0.225	0.159	0.188	0.204	0.196	0.229	0.134	0.137	0.182	0.201	0.012	0.108
EVE_PAPS	0.416	0.410	0.439	0.411	0.407	0.415	0.397	0.433	0.400	0.383	0.445	0.254	0.409	0.392	0.399	0.442	0.542	0.548
Average	**0.045**	**0.048**	**0.047**	**0.041**	**0.051**	**0.048**	**0.053**	**0.051**	**0.044**	**0.059**	**0.035**	**0.066**	**0.075**	**0.038**	**0.051**	**0.031**	**0.037**	**0.083**

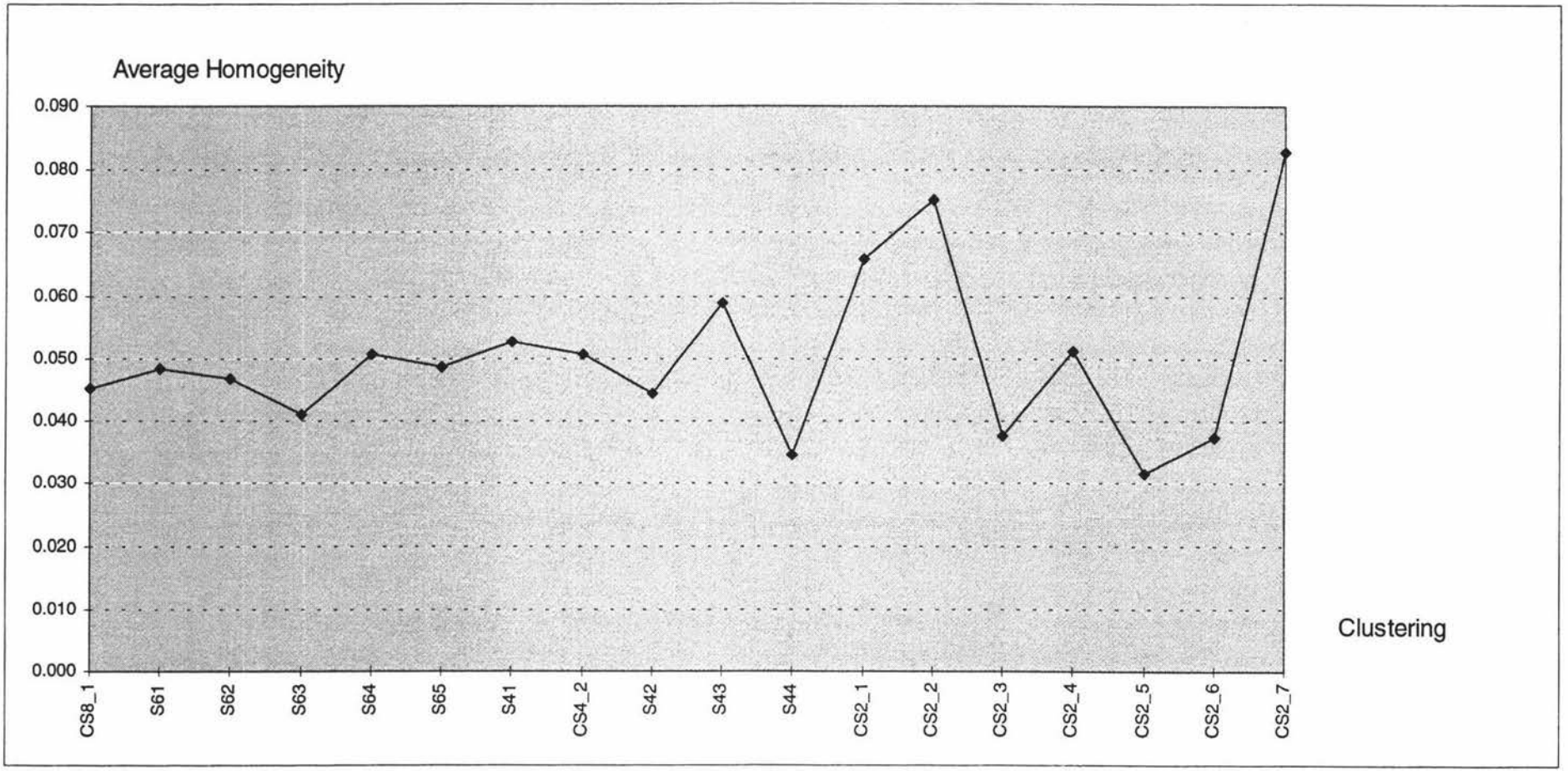

Figure 6. Relation between Homogeneity and Clusterings

Second, the relation between homogeneity and cluster size is reverse to that between design effect and cluster size. This is reflected in figure 2 & 7.

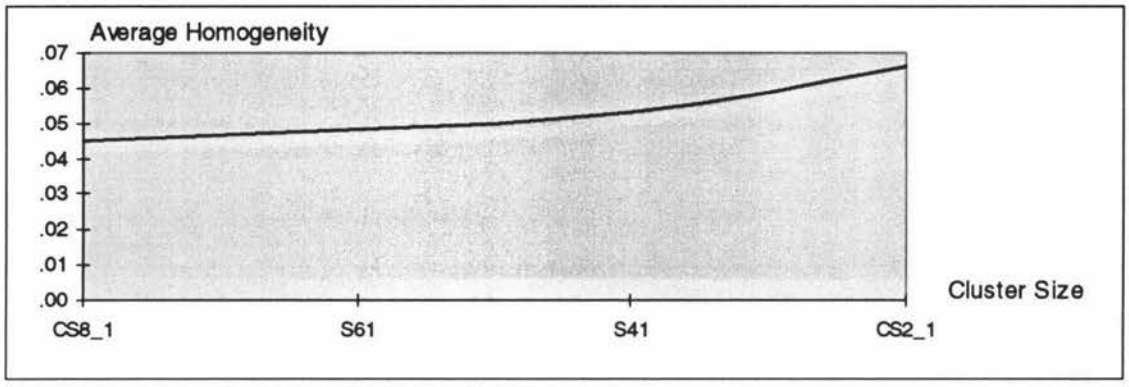

Figure 7. Relation between Homogeneity and Cluster Size with a Given Clustering Interval

Third, the rate of homogeneity increases with decrease of cluster size. This increase rate is accelerated by decreasing cluster size. This is reflected in figure 7 and table 12.

In table 12, within a given clustering, the increase rate of homogeneity is 0.003 from cluster size 8 to cluster size 6, 0.005 from cluster size 6 to cluster size 4, and 0.013 from cluster size 4 to cluster size 2.

Fourth, the variability of the average homogeneity values over variables across clustering within a larger cluster tends to be weaker than that within a smaller cluster. This is reflected in figure 8, 9 & 10.

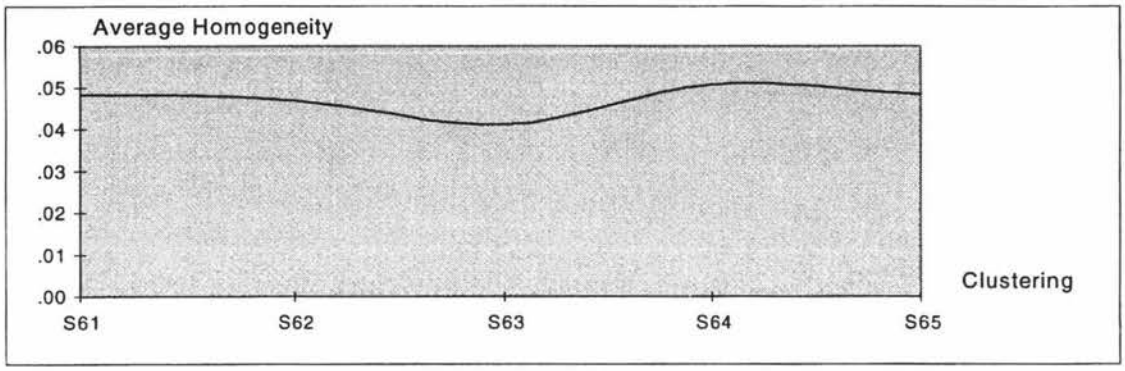

Figure 8. Relation between Homogeneity and Clustering with Cluster Size 6

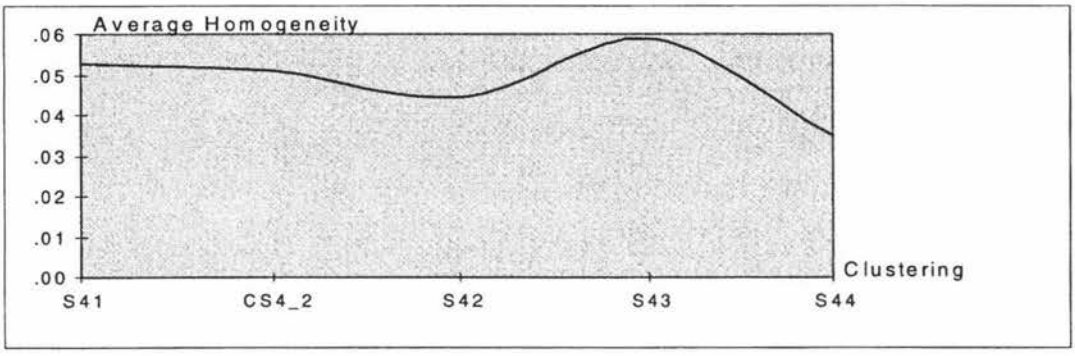

Figure 9. Relation between Homogeneity and Clustering with Cluster Size 4

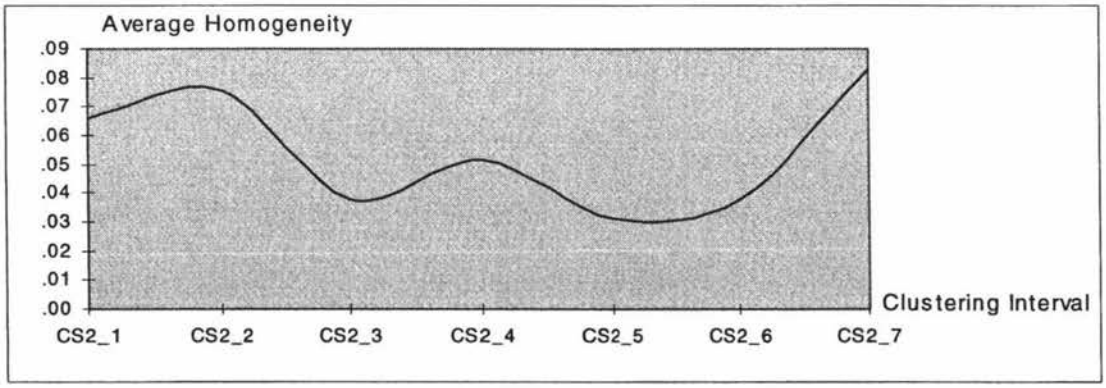

Figure 10. Relation between Homogeneity and Clustering with Cluster Size 2

Appendix F. Comparison of Two Variance Estimation Methods.

Table 13. Comparison of Two Design Effect Estimation Methods with 41 Variables (1)

Samples	CS8_1		S61		S41	
Methods*	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2
Variables						
AIRNZPW	1.15	1.14	1.06	1.05	1.05	1.04
ACCIDENT	1.05	1.04	1.07	1.06	.88	.88
P_YOUTH	1.55	1.54	1.57	1.56	1.62	1.61
MOR_PAPS	2.42	2.40	2.12	2.10	1.67	1.66
EVE_PAPS	3.91	3.88	3.05	3.03	2.19	2.18
NTRUNTVE	1.25	1.24	1.28	1.27	1.17	1.17
BOAT_NZ	.99	.98	.99	.99	1.03	1.03
LTV_N_RT	.97	.96	.92	.91	.96	.95
NZ_W_WK	1.07	1.06	1.10	1.09	1.07	1.06
OWNCIGAR	1.29	1.28	1.22	1.21	1.19	1.18
TV_GUIDE	1.23	1.22	1.23	1.22	1.15	1.14
WOMAN_WK	1.07	1.06	.95	.95	1.11	1.10
H_KIDSTD	1.08	1.07	1.09	1.08	.95	.94
AUS_W_WK	1.00	.99	.93	.92	.96	.95
WOMANISS	1.21	1.20	1.16	1.15	.99	.98

Table 13. Comparison of Two Design Effect Estimation Methods with 41 Variables (2)

Samples	CS8_1		S61		S41	
Methods*	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2
Variables						
NZ_BUSI	1.15	1.14	.99	.98	1.05	1.04
W_FAIRGO	1.00	1.00	1.04	1.03	1.07	1.06
GRAPEVIN	1.66	1.64	1.41	1.40	1.29	1.28
H_BLUEBK	1.90	1.89	1.92	1.91	1.46	1.45
POLICTIC	1.12	1.11	1.12	1.12	1.14	1.14
SHENMORE	.93	.92	.86	.86	.91	.90
R_DIGEST	1.02	1.01	1.17	1.16	1.15	1.14
TEARAWAY	.96	.95	.80	.80	.90	.89
H_N_BUIL	1.01	1.01	1.01	1.01	.98	.98
SP_NEWS	1.28	1.27	1.12	1.12	.98	.97
ADVENTUR	1.17	1.16	1.40	1.39	1.25	1.24
WELLBRAN	1.10	1.09	1.06	1.05	.93	.92
FASHIONQ	1.03	1.02	1.11	1.11	1.06	1.05
NZ_GEOGR	1.23	1.22	1.13	1.13	1.26	1.25

Table 13. Comparison of Two Design Effect Estimation Methods with 41 Variables (3)

Samples	CS8_1		S61		S41	
Methods*	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2
Variables						
STYLE	.88	.87	.92	.91	.96	.95
WORKHOME	1.57	1.56	1.54	1.53	1.33	1.32
DINEROUT	1.68	1.66	1.45	1.44	1.21	1.21
DRINKOUT	1.55	1.54	1.42	1.41	1.26	1.25
WINEMEAL	1.95	1.94	1.64	1.63	1.40	1.39
CONCERT	1.73	1.72	1.55	1.53	1.34	1.33
FAMILYTO	1.24	1.23	1.15	1.14	1.12	1.12
RACEPROB	1.46	1.45	1.43	1.42	1.24	1.23
E_W_WKLY	1.07	1.06	1.11	1.10	1.09	1.08
NEWPRODU	1.26	1.25	1.11	1.10	1.01	1.00
SPECIAL	1.32	1.31	1.37	1.36	1.22	1.21
NEW_IDEA	.99	.98	.93	.92	.93	.93

* **Method 1**: the method used in this study, that is, $design\ \ effect = \frac{ms_b}{ms}$;

Method 2: the alternative method, that is, $design\ \ effect \approx mr^2$.

REFERENCES

Alsagoff, S A; Esslemont, D H B & Gendall, P J (1986). *The precision of omnibus survey estimates in New Zealand.* Research Report No.46, Market Research Centre, Massey University.

Bebbington, A C & Smith, T M F (1977). The effect of survey design on multivariate analysis. In C. A. O'Muircheartaigh & C. Payne (eds) *The analysis of survey data Vol.**2** model fitting.* John Wiley & Sons Inc.

Campbell, C (1977). Properties of ordinary and weighted least squares estimators of regression coefficients for two-stage samples. *Proceedings of the American Statistical Association, Social Statistics Section,* 800-805.

Cochran, W G (1963). *Sampling techniques, 2nd ed.* John Wiley & Sons Inc.

Cochran, W G (1977). *Sampling techniques, 3rd ed.* John Wiley & Sons Inc.

Deming, W E (1956). On simplifications of sampling design through replication with equal probabilities and without stages. *Journal of the American Statistical Association* ***51***, 24-53.

Deming, W E (1960). *Sampling design in business research.* John Wiley & Sons Inc.

Durbin , J (1959). A note on the application of Quenouille's method of bias reduction to the estimation of ratios. *Biometrika* ***46***, 477-480.

Efron, B (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics* **7**, 1-26.

Efron, B (1981). Nonparametric standard errors and confidence intervals. *The Canadian Journal of Statistics* ***9***, 139-172.

Efron, B (1982). *The jackknife, the bootstrap and other resampling plans.* Philadelphia: Society for Industrial and Applied Mathematics **monograph no. *38***.

Ferringo, P; Valli, A; Groenerveld, T; Buch, E & Coetzee,

D (1992). The effect of cluster sampling in an African urban setting. *Central African Journal of Medicine* ***38***(8), 324-330.

Frankel, M R (1971). *Inference from Survey Samples: An Empirical Investigation*. Ann. Arbor., MI: Institute for Social Research.

Fuller, W A (1996). *Introduction to statistical times series (2nd ed)*. John Wiley & Sons Inc.

Hansen, M H; Hurwitz, W N & Madow, W G (1953). *Sample survey methods and theory* ***1 & 2***. John Wiley & Sons Inc.

Kalton, G (1979). Ultimate cluster sampling. *Journal of the Royal Statistical Society* ***A142***, 210-222.

Kish, L (1965). *Survey sampling*. John Wiley & Sons Inc.

Kish, L (1987). *Statistical design for research*. John Wiley & Sons Inc.

Kish, L & Frankel, M R (1970). Balanced repeated replication for standard errors. *Journal of the American Statistical Association* ***65***(331), 1071-1094.

Kish, L & Frankel, M R (1974). Inference from complex samples. *Journal of the Royal Statistical Society* ***B36***, 1-37.

Kovar, J G, Rao, J N K & Wu, C F J (1988). Bootstrap and other methods to measure errors in survey estimates. *The Canadian Journal of Statistics,* ***16***(Supplement), 25-45.

Laniel, N & Mohl, C (1994). Analysis of urban cluster size in the Canadian labour force survey. *Proceedings of the Survey Research Methods Section, American Statistical Association* ***2***, 931-936.

Lee, E S; Forthofer, R N & Lorimor, R J (1986). Analysis of complex sample survey data: problems and strategies. *Sociological Methods & Research* ***15***, 69-100.

Lehtonen, R & Pahkinen, J (1995). *Practical methods for design and analysis of complex surveys*. John Wiley & Sons Inc.

McCarthy, P J (1966). Replication: an approach to the analysis of data from complex surveys. *Vital & Health*

Statistics , Series ***2***, No. 14, Washington, DC: National Centre for Health Statistics, Public Health Service.

McCarthy, P J (1969). Pseudo-replication: half-samples. *Review of the International Statistical Institute* ***37***, 239-264.

Mahalanobis, P C (1939). A sample survey of the acreage under jute in Bengal. *Sankhya* **4**, 511-531.

Mahalanobis, P C (1944). On large-scale sample surveys. *Philosophical Transactions of the Royal Society of London* ***B231***, 329-451.

Mahalanobis, P C (1946). Recent experiments in statistical sampling in the Indian Statistical Institute. *Journal of the Royal Statistical Society* **109**, 325-370.

Norlen, U & Waller, T (1979). Estimation in a complex survey - experiences from a survey of buildings with regard to energy usage. *Statistisk Tidskrift* ***17***, 109-124.

Plackett, R L & Burmand, P J (1946). The design of optimum multifactorial experiments. *Biometrika* ***33***, 305-325.

Quenouille, M H (1949). Problems in plane sampling. *Annals of Mathematical Statistics* ***20***, 355-375.

Quenouille, M H (1956). Notes on bias in estimation. *Biometrika* ***43***, 353-360.

Roa, J N K (1997). Developments in sample survey theory: an appraisal. *The Canadian Journal of Statistics* ***25****(1)*, 1-21.

Roa, J N K & Wu, C F J (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association* ***83***(401), 231-241.

Särndal, C E; Swensson, B & Wretman, J (1992). *Model assisted survey sampling*. New York: Springer-Verlag.

Scott, A J & Holt, D (1982). The effect of two-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association* ***77***, 848-854.

Sitter, R R (1992). A resampling procedure for complex survey data. *Journal of the American Statistical Association* ***87***(419), 755-765.

Skinner, C J; Holt, D & Smith, T M F (eds) (1989).

Analysis of complex surveys. John Wiley & Sons Inc.

Som, R K (1996). *Practical sampling techniques (2nd ed.)*. New York: Marcel Deckker Inc.

Sudman, S (1970). The multiple uses of primary sampling areas of national probability samples. *Journal of the American Statistical Association* ***65***(329), 61-70.

Sudman, S (1976). *Applied sampling*. New York: Academic Press.

Thompson, M E (1997). *Theory of sampling surveys*. London: Chapman & Hall.

Tukey, J w (1958). Bias and confidence in not quite large samples (abstract). *Annals of Mathematical Statistics* ***29***, 614.

Verma, V & Lê, T (1996). An analysis of sampling errors for the demographic and health surveys. *International Statistical Review* ***3***(64), 265-294.

Wolter, K M (1985). *Introduction to variance estimation*. New York: Springer-Verlag Inc.

Woodruff, R S (1971). A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association* ***66***, 411-414.

BIBLIOGRAPHIES

Arbia, G (1993). The use of GIS in spatial statistical surveys. *International Statistical Review* ***61***(2), 339-359.

Bellhouse, D R & Rao, J N K (1994). Analysis of domain means in complex surveys. *Proceedings of the Survey Research Methods Section, American Statistical Association* ***1***, 29-34.

Bienias, J L; Sweet, E M & Alexander, C H (1990). A model for simulating interviewer travel costs for different cluster sizes. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 20-26.

Cantwell, P J (1990). Equal characteristic clustering. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 231-236.

Chapman, D W (1994). Optimum sample design for personal visit establishment surveys. *Proceedings of the Survey Research Methods Section, American Statistical Association* ***2***, 704-707.

Choi, J W (1989). Variance of intracluster correlation estimator. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 188-192.

Fitch, D J (1987). Estimating variance as function of cluster size. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 809-811.

Frankel, M R (1983). Sampling theory. In P. H. Rossi et al (eds) *Handbook of survey research*. Academic Press.

Frankel, M R & Frankel, L R (1977). Some recent developments in sample survey design. *Journal of Marketing Research* ***August***, 280-293.

Groves, R M (1989). *Survey errors and survey costs*. John Wiley & Sons Inc.

Hahn, G J & Meeker, W Q (1993). Assumptions for statistical inference. *The American Statistician* ***47***(1), 1-11.

Harraway, J (1993). *Introductory statistical methods and the analysis of variance(2nd ed)*. Dunedin, NZ: University of Otago Press.

Harris, P (1977). The Effect of clustering on costs and sampling errors of random samples. *Journal of the Market Research Society* ***19****(3)*, 112-122.

Hendricks, W A (1944). The relative efficiencies of groups of farms as sampling units. *Journal of the American Statistical Association* ***39***, 367-376.

Hinkley, D V (1988). Bootstrap methods. *Journal of the Royal Statistical Society* ***B52***(3), 321-337.

Kish, L (1989). Deffs: why, when and how? a review. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 209-211.

Kleinn, C (1994). Comparison of the performance of line sampling to other forms of cluster sampling. *Forest Ecology & Management* ***68***(213), 365-373.

Laniel, N & Mohl, C (1994). Analysis of urban cluster size in the Canadian labour force survey. *Proceedings of the Survey Research Methods Section, American Statistical Association* ***2***, 931-936.

Lee, K (1972). Partially balanced designs for half sample replication method of variance estimation. *Journal of the American Statistical Association* ***67***(338), 324-334.

Mantel, H; Laniel, N; Duval, M & Marion, J (1994). Cost modelling of alternative sample designs for rural areas in the Canadian labour force survey. *Proceedings of the Survey Research Methods Section, American Statistical Association* ***2***, 925-930.

McVay, F E (1947). Sampling methods applied to estimating numbers of commercial orchards in a commercial peach area. *Journal of the American Statistical Association* ***42***, 533-540.

Pfeffermann, D (1983). On the relative efficiency of four ratio type estimators in cluster sampling. *Sankhyã* ***B45***(3), 376-388.

Proctor, C H (1985). Fitting H F Smith's empirical law to cluster variances for use in designing multi-stage

sample surveys. *Journal of the American Statistical Association* ***80***(390), 294-300.

Proctor, C H (1992). Basic cluster sample design. *Proceedings of the Survey Research Methods Section, American Statistical Association,* 526-532.

Royall, R M (1986). The prediction approach to robust variance estimation in two-stage cluster sampling. *Journal of the American Statistical Association* ***81***(393), 119-123.

Sadooghi-Alvandi, M (1986). The choice of subsample size in two-stage sampling. *Journal of the American Statistical Association* ***81***(394), 555-558.

Singh, M P; Drew, J D; Gambino, J G & Mayda, F (1990). *Methodology of the Canadian labour force survey.* Statistics Canada catalogue 71-526.

Skinner, C J (1986). Design effects of two-stage sampling. *Journal of the Royal Statistical Society* ***B48***,(1), 89-99.

Skinner, C J (1981). Estimation of the variance of a finite population for cluster samples. *Sankhyã* ***B43***(3), 392-398.

Som, R K (1996). *Practical sampling techniques (2nd ed.).* New York: Marcel Deckker Inc.

Stehman, S V (1997). Estimating standard errors of accuracy assessment statistics under cluster sampling. *Remote Sensing of Environment* ***60***, 258-269.

Sudman, S (1970). The multiple uses of primary sampling areas of national probability samples. *Journal of the American Statistical Association* ***65***(329), 61-70.

Sudman, S (1978). Optimum cluster designs within a primary unit using combined telephone screening and face-to-face interviewing. *Journal of the American Statistical Association* ***73***(362), 300-304.

Tam, S M (1995). Optimal and robust strategies for cluster sampling. *Journal of the American Statistical Association* ***90***(429), 379-382.

Tepping, B J (1968). Variance estimation in complex surveys. *Proceedings of the Social Statistics Section, American Statistical Association.*

Thomsen, I; Tesfu, D & Binder, D A (1986). Estimation of design effects and intraclass correlation when using outdated measures of size. *International Statistical Review* ***54***(3), 343-349.

Valliant, R (1995). Limitations of balanced half sampling when strata are grouped. *Proceedings of the Survey Research Methods Section, American Statistical Association* ***1***, 120-125.

Williams, R L; Folsom, R E & LaVange, L M (1983). The implication of sample design on survey data analysis. In T Wright (Eds) *Statistical methods & the improvement of data quality*. Academic Press Inc.

Woodruff, R S & Causey, B D (1976). Computerized method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association* ***71***, 315-321.