1352.0.55.073

Australian Bureau of Statistics

Research Paper

Decomposition of Movement Estimates as a Diagnostic Tool for Repeated Business Surveys

New Issue

Research Paper

Decomposition of Movement Estimates as a Diagnostic Tool for Repeated Business Surveys

John Preston

Statistical Services Branch

Methodology Advisory Committee

18 November 2005, Canberra

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30 AM (CANBERRA TIME) THU 16 FEB 2006

ABS Catalogue no. 1352.0.55.073
ISBN 0 642 48177 6

Produced by the Australian Bureau of Statistics

INQUIRIES

The ABS welcomes comments on the research presented in this paper.
For further information, please contact Mr John Preston, Statistical Services Branch on Brisbane (07) 3222 6229 or email <john.preston@abs.gov.au>.

CONTENTS

Decomposition of Movement Estimates as a Diagnostic Tool for Repeated Business Surveys

John Preston
Australian Bureau of Statistics
639 Wickham Street
Fortitude Valley QLD 4006
john.preston@abs.gov.au
(07) 3222 6229

1. Introduction

The ABS conducts a number of repeated business surveys. One of the key objectives of these business surveys is to produce reliable estimates of change from one time period to the next. Output editing is usually the final check for major errors in the survey data to ensure that the estimates are reliable. There are a number of tools that can be used to check whether the estimates of change appear correct, such as comparing the estimate of change with external sources, examining the distributions of the survey data, or decomposing the movement estimates into components which lead to a better understanding of the estimates of change.

The estimates of change are driven by a number of different factors, including:

- birthing of new units, deathing of old units and changing characteristics of continuing units in the population of interest;
- rotation of new units into sample and rotation of previously selected units out of sample;
- changes in the values of auxiliary benchmark variables at the unit and aggregate levels; and
- changes in the values of the variable of interest at the unit level.

Some of these factors arise from actual changes in the population of interest, while others arise from a changes in the relationships between the population and the chosen samples.

The ABS regularly updates its business survey frames by adding births, removing deaths and updating stratification variables. These frame changes have a direct impact on the estimates of change, as well as an indirect impact through the rotation of units into and out of the sample. The rotation of units into and out of the sample is also be caused by the ABS policy to rotate all small and medium business out of individual surveys after three years. This policy is upheld using synchronised sampling, a form Permanent Random Number (PRN) sampling procedure, developed by the ABS to minimise or maximise overlap between surveys as well as control rotation within surveys (Brewer, Gross and Lee 1999).

The ABS has recently developed a Generalised Estimation System for processing its large scale annual and sub-annual business surveys, which allows the specification of an estimator from a wide group of estimators produced under generalised linear regression models. For those business surveys which produce estimates using auxiliary information in the form of known auxiliary totals, changes in the values of auxiliary benchmark variables at the unit and aggregate levels will have an impact on the estimates of change.

In order to better understand the influence of these factors on the estimates of change, a methodology has been developed to decompose the movement estimates into a number of factors. Holt and Skinner (1989) present a framework for the decomposition of the net difference in the population means into effects due to net differences of means within domains and effects due to domain composition. The framework is also extended to allow for non-stationary populations with birthing of new units, deathing of old units and changing characteristics of continuing units. This paper presents an more comprehensive alternative methodology for the decomposition of the movement estimates which can be used as a diagnostic tool to assist in the output editing of sample surveys.

Section 2 introduces the generalised regression estimator used in ABS business surveys. Section 3 presents one possible decomposition of the movement estimates into primary and secondary factors. Section 4 describes how the bootstrap variance estimator can be used to produce approximate variances for the various decomposition components. Section 5 measures the biases and mean squared errors of the decomposition components and the bootstrap variance estimator of the decomposition components in a simulation study. Section 6 examines the ability of the proposed decomposition methodology to identify irregularities.

QUESTION 1: Is the proposed decomposition methodology into primary and secondary factors the "best" approach to the decomposition of the movement estimates?

2. Estimation for ABS Business Surveys

2.1 The Generalised Regression Estimator

Consider a finite population U divided into H strata $U = \{U_1, ..., U_h, ..., U_H\}$, where U_h is comprised of N_h units. The objective is to estimate the population total $Y = \sum_h \sum_{i \in U_h} y_{hi}$, where y_{hi} is the value of the variable of interest y for unit i in stratum h. Assume there exists a set of auxiliary variables $\underset{\sim}{x}_i = \{x_{1i}, ..., x_{ki}, ..., x_{Ki}\}$ for which the population totals $\underset{\sim}{X} = \sum_{i \in U} \underset{\sim}{x}$ are known. Suppose stratified simple random samples $\{s_1, ..., s_h, ..., s_H\}$ without replacement (SRSWOR) of sizes $\{n_1, ..., n_h, ..., n_H\}$ are drawn with selection probabilities $\pi_i = n_h/N_h$ from populations $\{U_1, ..., U_k, ..., U_H\}$. The sampling weights $w_i = 1/\pi_i$ are those used in the Horvitz-Thompson estimator $\hat{Y} = \sum_h \sum_{i \in s_h} w_i y_{hi}$ for variable of interest y. The generalised regression estimator is given by (Sarndal, Swensson and Wretman, 1992):

$$\hat{Y} = \sum_{i \in s} w_i y_i + \left(\underset{\sim}{X} - \sum_{i \in s} w_i \underset{\sim}{x}_i \right)' \underset{\sim}{\hat{\beta}} = \sum_{i \in s} w_i g_i y_i \tag{2.1}$$

where g_i is the g-weight for unit i, defined as:

$$g_i = \left(1 + \left(\underset{\sim}{X} - \sum_{i \in s} w_i \underset{\sim}{x}_i \right) \left(\sum_{i \in s} \frac{w_i \underset{\sim}{x}_i \underset{\sim}{x}_i'}{c_i} \right)^{-1} \frac{\underset{\sim}{x}_i}{c_i} \right)$$

$\hat{\underset{\sim}{\beta}}$ is the vector of the linear regression model parameters given by:

$$\hat{\underset{\sim}{\beta}} = \left(\sum_{i \in s} \frac{w_i \underset{\sim}{x}_i \underset{\sim}{x}_i'}{c_i} \right)^{-1} \left(\sum_{i \in s} \frac{w_i \underset{\sim}{x}_i y_i}{c_i} \right)$$

and c_i are specified positive factors that relate to the variance structure of the linear regression model associated with the GREG estimator:

$$y_i = \underset{\sim}{x}_i' \hat{\underset{\sim}{\beta}} + \varepsilon_i$$

where $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = c_i \sigma^2$ and $Cov(\varepsilon_i, \varepsilon_j) = 0$, for all $i \neq j$.

The generalised regression estimator for domain of interest c is given by:

$$\hat{Y}_c = \sum_{i \in s_c} w_i g_i y_i \qquad (2.2)$$

where s_c is the sample of units within domain of interest c.

2.2 The Generalised Regression Estimator for Movement Estimates

Define the population at time t by $U^{(t)} = \{U_1^{(t)}, ..., U_h^{(t)}, ..., U_H^{(t)}\}$, where $U_h^{(t)}$ is comprised of $N_h^{(t)}$ units. The objective is to estimate the difference in the population totals between the two time points, $Y^{(m)} = Y^{(2)} - Y^{(1)} = \sum_h \sum_{i \in U_h^{(2)}} y_{hi}^{(2)} - \sum_h \sum_{i \in U_h^{(1)}} y_{hi}^{(1)}$, where $y_{hi}^{(t)}$ is the value of the variable of interest y at time point t for unit i in stratum h. The generalised regression movement estimator is given by:

$$\hat{Y}^{(m)} = \sum_{i \in s^{(2)}} w_i^{(2)} g_i^{(2)} y_i^{(2)} - \sum_{i \in s^{(1)}} w_i^{(1)} g_i^{(1)} y_i^{(1)} \qquad (2.3)$$

The generalised regression movement estimator for domain of interest c is given by:

$$\hat{Y}_c^{(m)} = \sum_{i \in s_c^{(2)}} w_i^{(2)} g_i^{(2)} y_i^{(2)} - \sum_{i \in s_c^{(1)}} w_i^{(1)} g_i^{(1)} y_i^{(1)} \qquad (2.4)$$

where $s_c^{(t)}$ is the sample of units at time point t within domain of interest c.

2.3 The Movement Estimator as a Function of Four Factors

The movement estimator (2.4) can be written as the difference between a product of four factors:

$$\hat{Y}_c^{(m)} = \sum_h N_h^{(2)} \hat{p}_{hc}^{(2)} \bar{g}_{hc}^{(2)} \bar{y}_{hc}^{(2)} - \sum_h N_h^{(1)} \hat{p}_{hc}^{(1)} \bar{g}_{hc}^{(1)} \bar{y}_{hc}^{(1)} \qquad (2.5)$$

where $\hat{p}_{hc}^{(t)}$ is the estimated proportion of units at time point t in stratum h in domain of interest c:

$$\hat{p}_{hc}^{(t)} = \frac{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)}}{\sum_{i \in s_h^{(t)}} w_i^{(t)}}$$

$\bar{g}_{hc}^{(t)}$ is the weighted average g-weight at time point t in stratum h in domain of interest c:

$$\bar{g}_{hc}^{(t)} = \frac{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)} g_i^{(t)} y_i^{(t)}}{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)} y_i^{(t)}}$$

and $\bar{y}_{hc}^{(t)}$ is the average value of the variable of interest y at time point t in stratum h in domain of interest c:

$$\bar{y}_{hc}^{(t)} = \frac{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)} y_i^{(t)}}{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)}}$$

In this situation the average g-weight is weighted by the value of the variable of interest, which will mean that changes in the value of the variable of interest will impact on the average value of the variable of interest as well as the weighted average g-weight. The main reason for using the weighted average g-weight, rather than an unweighted average g-weight, is to ensure that the product of the four factors sum to the survey estimates. Alternatively, $\bar{g}_{hc}^{(t)}$ could be calculated as the average g-weight at time point t in stratum h in domain of interest c:

$$\bar{g}_{hc}^{(t)} = \frac{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)} g_i^{(t)}}{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)}}$$

and $\bar{y}_{hc}^{(t)}$ could be calculated as the weighted average value of the variable of interest y at time point t in stratum h in domain of interest c:

$$\bar{y}_{hc}^{(t)} = \frac{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)} g_i^{(t)} y_i^{(t)}}{\sum_{i \in s_{hc}^{(t)}} w_i^{(t)} g_i^{(t)}}$$

which would also ensure that the product of the four factors sum to the survey estimates. In this situation the weighted average value of the variable of interest is weighted by the g-weight, which will mean that changes in the g-weight will impact on the average g-weight as well as the weighted average value of the variable of interest. The weighted average g-weight is used in the rest of the paper. It was chosen over the weighted average value of the variable of interest because changes in the estimates are more often driven by changes in the average value of the variable of interest. The results of the decomposition of the movement estimates will be easier to interpret if they are based on the unweighted average value of the variable of interest.

QUESTION 2: Is appropriate to weight the average g-weight by the value of the variable of interest or weight the average value of the variable of the interest by the g-weight to ensure the product of the four factors sum to the survey estimates?

3. Decomposition of Movement Estimates for ABS Repeated Surveys

3.1 Decomposition into Four Primary Factors

Using a standard decomposition methodology the movement estimator (2.5) can be decomposed into the following terms:

$$\begin{aligned}\hat{Y}_c^{(m)} =& \sum_h [N_h^{(2)} - N_h^{(1)}]\hat{p}_{hc}^{(1)}\bar{g}_{hc}^{(1)}\bar{y}_{hc}^{(1)} + \sum_h [\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}]N_h^{(1)}\bar{g}_{hc}^{(1)}\bar{y}_{hc}^{(1)} \\ &+ \sum_h [\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}]N_h^{(1)}\hat{p}_{hc}^{(1)}\bar{y}_{hc}^{(1)} + \sum_h [\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]N_h^{(1)}\hat{p}_{hc}^{(1)}\bar{g}_{hc}^{(1)} \end{aligned} \quad (3.1)$$

$$\begin{aligned}&+ \sum_h [N_h^{(2)} - N_h^{(1)}][\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}]\bar{g}_{hc}^{(1)}\bar{y}_{hc}^{(1)} + \sum_h [N_h^{(2)} - N_h^{(1)}][\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}]\hat{p}_{hc}^{(1)}\bar{y}_{hc}^{(1)} \\ &+ \sum_h [N_h^{(2)} - N_h^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]\hat{p}_{hc}^{(1)}\bar{g}_{hc}^{(1)} + \sum_h [\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}][\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}]N_h^{(1)}\bar{y}_{hc}^{(1)} \\ &+ \sum_h [\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]N_h^{(1)}\bar{g}_{hc}^{(1)} + \sum_h [\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]\hat{p}_{hc}^{(1)}\bar{y}_{hc}^{(1)} \end{aligned} \quad (3.2)$$

$$\begin{aligned}&+ \sum_h [N_h^{(2)} - N_h^{(1)}][\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}][\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}]\bar{y}_{hc}^{(1)} \\ &+ \sum_h [N_h^{(2)} - N_h^{(1)}][\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]\bar{g}_{hc}^{(1)} \\ &+ \sum_h [N_h^{(2)} - N_h^{(1)}][\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]\hat{p}_{hc}^{(1)} \\ &+ \sum_h [\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}][\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]N_h^{(1)} \end{aligned} \quad (3.3)$$

$$+ \sum_h [N_h^{(2)} - N_h^{(1)}][\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}][\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}][\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}] \quad (3.4)$$

The first four terms (3.1) are the main effects for the four factors, while the other terms (3.2), (3.3) and (3.4) are the two, three and four way interaction terms between the four factors. These interaction terms are difficult to interpret and are usually relative small compared to the main effects (Holt and Skinner, 1989). Although this decomposition methodology has the primary factors standardised at the first time period, it would have also been acceptable to have the primary factors standardised at the second time period.

In order to avoid including the interaction terms, Das Gupta (1991) developed an alternative decomposition methodology, along the lines suggested by Kitagawa (1955). Using the alternative decomposition methodology the movement estimator (2.5) can be decomposed into the following four primary effects:

$$\hat{Y}_c^{(m)} = \sum_h [N_h^{(2)} - N_h^{(1)}]\varphi(\hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad (3.5)$$

$$+\sum_h [\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}]\varphi(N_h^{(1)}, N_h^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad (3.6)$$

$$+\sum_h [\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad (3.7)$$

$$+\sum_h [\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.8)$$

where $\varphi(a^{(1)}, a^{(2)}, b^{(1)}, b^{(2)}, c^{(1)}, c^{(2)}) = (3a^{(1)}b^{(1)}c^{(1)} + a^{(1)}b^{(1)}c^{(2)} + a^{(1)}b^{(2)}c^{(1)} + a^{(1)}b^{(2)}c^{(2)}$
$+ a^{(2)}b^{(1)}c^{(1)} + a^{(2)}b^{(1)}c^{(2)} + a^{(2)}b^{(2)}c^{(1)} + 3a^{(2)}b^{(2)}c^{(2)})/12$

The first terms (3.5) measures the effect of changes in the number of units on the survey frame, the second term (3.6) measures the effect of changes in the estimated proportion of units in the domain of interest, the third term (3.7) measures the effect of changes in the weighted average g-weight, and the fourth term (3.8) measures the effect of changes in the average value of the variable of interest.

3.1.1 Empirical Evaluation of the Contribution of Interaction Terms

The Retail Business Survey was used to evaluate the relative size of the interaction terms to determine whether is it appropriate to distribute these interaction terms into the main effects. The Retail Business Survey produces monthly estimates of the value of turnover for retail businesses classified by state and industry. The principal objective of the Retail Business Survey is to show month to month movements of turnover for retail industries.

The contribution of the main effects and interaction terms to the relative movements between the June and July 2005 estimates of retail turnover are presented in Table 1, while a comparison of the main effects and the primary effects to the relative movements between the June and July 2005 estimates of retail turnover are presented in Table 2.

Table 1: Contribution of Main Effects and Interaction Terms to Relative Movements[#]

State	Net Contribution to Relative Movements				Percentage Contribution to Relative Movements		
	Main Effects	2nd Order Interaction Terms	3rd Order Interaction Terms	Relative Movements	Main Effects	2nd Order Interaction Terms	3rd Order Interaction Terms
NSW	0.89	0.13	-0.05	0.97	91.4%	13.4%	-4.8%
Vic	0.54	0.25	-0.08	0.71	76.7%	34.6%	-11.3%
Qld	4.85	-0.11	0.00	4.73	102.4%	-2.4%	-0.0%
SA	3.78	-0.25	0.02	3.55	106.6%	-7.2%	0.5%
WA	1.91	0.03	0.00	1.93	98.6%	1.4%	0.0%
Tas	3.36	-0.05	0.19	3.51	96.0%	-1.5%	5.5%
NT	6.07	-0.46	0.00	5.62	108.1%	-8.1%	0.0%
ACT	-0.16	0.93	-0.22	0.55	-29.0%	169.1%	-40.1%
Aust	1.99	0.08	-0.03	2.03	97.9%	3.8%	-1.7%

[#] All main effect and interaction terms which involve differences in the estimated proportion of units within the domain of interest (i.e. State) between the two time periods will be equal to zero, since the Retail Business Survey is stratified by State.

Table 2: Comparison of Main Effects and Primary Effects

State	Net Contribution to Relative Movements of Main Effects			Net Contribution to Relative Movements of Primary Effects		
	Frame Change Effect	Weight Change Adjustment Effect	Average Value Change Effect	Frame Change Effect	Weight Change Adjustment Effect	Average Value Change Effect
NSW	0.81	0.38	-0.30	0.81	0.44	-0.27
Vic	0.44	-0.30	0.41	0.48	-0.23	0.46
Qld	0.34	2.22	2.28	0.31	2.20	2.23
SA	0.72	2.62	0.45	0.58	2.58	0.38
WA	0.79	1.11	0.01	0.76	1.14	0.03
Tas	0.35	0.33	2.68	0.37	0.41	2.72
NT	0.93	0.30	4.85	0.71	0.29	4.62
ACT	-0.01	3.18	-3.33	0.13	3.47	3.05
Aust	0.59	0.87	0.53	0.58	0.91	0.54

In most cases the percentage contribution to the relative movements of the main effects was between 90 and 110 percent. In those cases where the interaction terms contributed more than 10 percent of the relative movements, the relative movements were small (i.e. less than one percent) as were the net contribution to the relative movements. Furthermore, the net contribution to the primary effects are similar to the net contribution to the main effects.

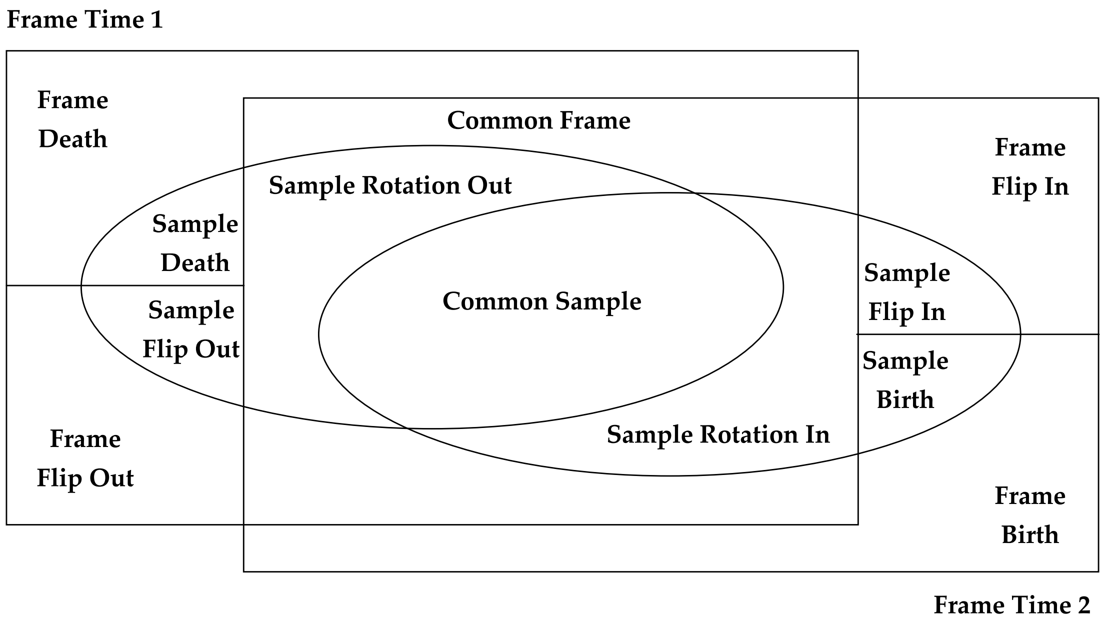
QUESTION 3: Is it appropriate to distribute the interaction terms into the main effects?

3.2 Decomposition into Secondary Factors

The movement estimates are influenced by more factors than just the four primary factors described in Section 3.1. For example, stratum level changes in the number of units on the survey frame can occur as a births and deaths in the population as well as units changing stratum. Furthermore, changes in the average value of the variable of interest can occur as a result of changes in the values of the common sample units as well as differences between the values of the uncommon units.

A pictorial illustration of the stratum level changes in the survey frame and survey sample between the two time periods is presented in Figure 1.

Figure 1: Pictorial Illustration of Stratum Level Changes in the Survey Frame and Survey Sample



The units on the survey frames within the stratum in the two time periods are represented by the left and right large rectangles respectively. The units on the common frame within the stratum are represented by the intersection of the two large rectangles, while the units not on the common frame within the stratum can be split into frame births, frame deaths, frame flips in and frame flips out. Those units which move into and out of the scope of a survey will be considered frame births and frame deaths.

The units in the survey samples within the stratum in the two time periods are represented by the left and right ovals respectively. The units in the common survey sample within the stratum are represented by the intersection of the two ovals, while the units not in the common survey samples within the stratum can be split into sample rotations in, sample rotations out, sample flips in, sample flips out, sample births and sample deaths.

The purpose of the decomposition into secondary factors is to split the effect of changes in the number of units on the survey frame into frame births, frame deaths, frame flips in and frame flips out effects, and the effect of changes in the average value of the variable of interest into common sample, sample rotations in, sample rotations out, sample flips in, sample flips out, sample births and sample deaths effects.

3.2.1 Frame Change Decomposition

The difference in the number of units on the survey frames between the two time periods, can be written as:

- differences in the number of frame births and frame deaths (i.e. frame growth); plus
- differences in the number of frame flips in and frame flips out (i.e. frame stratum flips).

$$[N_h^{(2)} - N_h^{(1)}] = [N_h^{(2b)} - N_h^{(1d)}] + [N_h^{(2fi)} - N_h^{(1fo)}]$$

Therefore, the frame change effect(3.5) can be decomposed into a frame growth effect (3.9) and a frame stratum flip effect (3.10):

$$\hat{Y}_c^{(mf)} = \sum_h [N_h^{(2)} - N_h^{(1)}]\varphi(\hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)})$$

$$= \sum_h [N_h^{(2b)} - N_h^{(1d)}]\varphi(\hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad (3.9)$$

$$+ \sum_h [N_h^{(2fi)} - N_h^{(1fo)}]\varphi(\hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad (3.10)$$

3.2.2 Average Value Change Decomposition

The difference in the average value of units in the survey samples between the two time periods can be written as a function of:

- differences in the average value of common sample units between the two time periods;
- differences in the average value of sample rotations in and common sample units at the second time period;
- differences in the average value of sample flips in and common sample units at the second time period;
- differences in the average value of sample births and common sample units at the second time period;
- differences in the average value of sample rotations out and common sample units at the first time period;
- differences in the average value of sample flips out and common sample units at the first time period; and

- differences in the average value of sample deaths and common sample units at the first time period.

$$\left[\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}\right] = \left[\bar{y}_{hc}^{(2c)} - \bar{y}_{hc}^{(1c)}\right]$$

$$+ \frac{n_{hc}^{(2ri)}}{n_{hc}^{(2c)}}\left[\bar{y}_{hc}^{(2ri)} - \bar{y}_{hc}^{(2c)}\right] + \frac{n_{hc}^{(2fi)}}{n_{hc}^{(2c)}}\left[\bar{y}_{hc}^{(2fi)} - \bar{y}_{hc}^{(2c)}\right] + \frac{n_{hc}^{(2b)}}{n_{hc}^{(2c)}}\left[\bar{y}_{hc}^{(2b)} - \bar{y}_{hc}^{(2c)}\right]$$

$$+ \frac{n_{hc}^{(1ro)}}{n_{hc}^{(1c)}}\left[\bar{y}_{hc}^{(1c)} - \bar{y}_{hc}^{(1ro)}\right] + \frac{n_{hc}^{(1fo)}}{n_{hc}^{(1c)}}\left[\bar{y}_{hc}^{(1c)} - \bar{y}_{hc}^{(1fo)}\right] + \frac{n_{hc}^{(1d)}}{n_{hc}^{(1c)}}\left[\bar{y}_{hc}^{(1c)} - \bar{y}_{hc}^{(1d)}\right]$$

Therefore, the average value change effect (3.8) can be decomposed into a common sample unit effect (3.11), a sample rotations in effect (3.12), a sample stratum flip in effect (3.13), a sample birth effect (3.14), a sample rotations out effect (3.15), a sample stratum flip out effect (3.16) and a sample death effect (3.17):

$$\hat{Y}_c^{(mu)} = \sum_h \left[\bar{y}_{hc}^{(2)} - \bar{y}_{hc}^{(1)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)})$$

$$= \sum_h \left[\bar{y}_{hc}^{(2c)} - \bar{y}_{hc}^{(1c)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.11)$$

$$+ \sum_h \frac{n_{hc}^{(2ri)}}{n_{hc}^{(2c)}}\left[\bar{y}_{hc}^{(2ri)} - \bar{y}_{hc}^{(2c)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.12)$$

$$+ \sum_h \frac{n_{hc}^{(2fi)}}{n_{hc}^{(2c)}}\left[\bar{y}_{hc}^{(2fi)} - \bar{y}_{hc}^{(2c)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.13)$$

$$+ \sum_h \frac{n_{hc}^{(2b)}}{n_{hc}^{(2c)}}\left[\bar{y}_{hc}^{(2b)} - \bar{y}_{hc}^{(2c)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.14)$$

$$+ \sum_h \frac{n_{hc}^{(1ro)}}{n_{hc}^{(1c)}}\left[\bar{y}_{hc}^{(1c)} - \bar{y}_{hc}^{(1ro)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.15)$$

$$+ \sum_h \frac{n_{hc}^{(1fo)}}{n_{hc}^{(1c)}}\left[\bar{y}_{hc}^{(1c)} - \bar{y}_{hc}^{(1fo)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.16)$$

$$+ \sum_h \frac{n_{hc}^{(1d)}}{n_{hc}^{(1c)}}\left[\bar{y}_{hc}^{(1c)} - \bar{y}_{hc}^{(1d)}\right]\varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \quad (3.17)$$

A description of the decomposition components (3.5 - 3.17) is presented in Appendix 1.

Under this average value change decomposition, the common sample unit effect (3.11) could be considered to be inconsistent with the other average value effects (3.12 - 3.17). The common sample unit effect compares the differences in the average value of common sample units between the two time periods, while the other average value effects compare differences in the average value of uncommon sample units and common sample units at a single time period.

There are two main reason for adopting this average value change decomposition: (i) to ensure that the average value effects sum to the average value change effect (3.8); and (ii) it would appears to be logical to compare the average value of uncommon sample units with the average value of common sample units at a the same time period, because if the uncommon units have the same values as the common units then the decomposition effects would be equal to zero.

In order for the other average value effects to be consistent with the common sample unit effect, then the other average value effects would need to compare differences in the average value of uncommon sample units at one time period and common sample units at the other time period. However, the average value effects would not sum to the average value change effect.

QUESTION 4: Is it necessary for the common sample unit effect to be consistent with the other average value effects?

In order to better interpret changes in the movement estimates, the primary and secondary factors can be collapsed into the following six broad components:

$$\hat{Y}_c^{(m)} = \hat{Y}_c^{(muc)} + \hat{Y}_c^{(mbd)} + \hat{Y}_c^{(msf)} + \hat{Y}_c^{(mr)} + \hat{Y}_c^{(md)} + \hat{Y}_c^{(mw)} \quad (3.18)$$

where $\hat{Y}_c^{(muc)}$ is the common sample unit effect (3.11), $\hat{Y}_c^{(mbd)} = \hat{Y}_c^{(mfg)} + \hat{Y}_c^{(mub)} + \hat{Y}_c^{(mud)}$ is the net birth-death effect (3.9, 3.14, 3.17), $\hat{Y}_c^{(msf)} = \hat{Y}_c^{(mfg)} + \hat{Y}_c^{(mufi)} + \hat{Y}_c^{(mufo)}$ is the net stratum flip effect (3.10, 3.13, 3.16) , $\hat{Y}_c^{(mr)} = \hat{Y}_c^{(muri)} + \hat{Y}_c^{(muro)}$ is the net rotations effect (3.12, 3.15) , $\hat{Y}_c^{(md)}$ is the domain change effect (3.6), and $\hat{Y}_c^{(mw)}$ is the weight adjustment change effect (3.7).

3.3 Decomposition of Movement Estimates of Rates

The movement estimates of a rate between the two time periods can be written as a function of differences between the estimates of the numerator between the two time periods and differences between the estimates of the denominator between the two time periods. Therefore, the movement estimates of a rate can be decomposed into a numerator effect and a denominator effect:

$$\hat{Y}_c^{(m)} = \left[\frac{\hat{A}_c^{(2)}}{\hat{B}_c^{(2)}} - \frac{\hat{A}_c^{(1)}}{\hat{B}_c^{(1)}} \right]$$

$$= \left[\left(\hat{A}_c^{(2)} - \hat{A}_c^{(1)}\right) \frac{\left(\hat{B}_c^{(1)} + \hat{B}_c^{(2)}\right)}{2\hat{B}_c^{(1)}\hat{B}_c^{(2)}} \right] \quad (3.19)$$

$$+ \left[-\left(\hat{B}_c^{(2)} - \hat{B}_c^{(1)}\right) \frac{\left(\hat{A}_c^{(1)} + \hat{A}_c^{(2)}\right)}{2\hat{B}_c^{(1)}\hat{B}_c^{(2)}} \right] \quad (3.20)$$

The numerator effect and denominator effect can be further decomposed by decomposing the movement estimates for the numerator $\left(\widehat{A}_c^{(2)} - \widehat{A}_c^{(1)}\right)$ and the denominator $\left(\widehat{B}_c^{(2)} - \widehat{B}_c^{(1)}\right)$ into the various decomposition components (3.5 - 3.17).

4. Decomposition "Replicate Variance" Methodology

4.1 The Bootstrap Variance Estimator

The bootstrap procedure for a stratified random sample selected without replacement (Shao and Tu 1995) is to: (i) Select simple random samples s_{rh} without replacement (SRSWOR) of $m_h = [n_h/2]$ units from the original sample of n_h units, independently within each stratum h. Let δ_{rhi} equal 1 if $i \in s_{rh}$ and zero otherwise. Calculate the bootstrap sampling weights:

$$w_{ri} = w_i \left(1 - \sqrt{\frac{(1-f_h)m_h}{(n_h - m_h)}} + \sqrt{\frac{(1-f_h)m_h}{(n_h - m_h)}} \frac{n_h}{m_h} \delta_{rhi} \right) \tag{4.1}$$

where $f_h = \dfrac{n_h}{N_h}$ is the sampling fraction within stratum h. Calculate the bootstrap estimator of Y_c:

$$\widehat{Y}_{rc} = \sum_{i \in s} w_{ri} g_{ri} y_i \tag{4.2}$$

(ii) Independently replicate step (i) a large number of times, R, and calculate the bootstrap estimates, $\widehat{Y}_{1c}, \widehat{Y}_{2c}, \ldots, \widehat{Y}_{Rc}$.

(iii) The bootstrap variance estimator is given by the Monte Carlo approximation:

$$Var(\widehat{Y}_c) = \frac{\sum_r \left(\widehat{Y}_{rc} - \overline{\widehat{Y}}_c \right)^2}{R-1} \tag{4.3}$$

where $\overline{\widehat{Y}}_c = \dfrac{\sum_r \widehat{Y}_{rc}}{R}$.

4.2 The Bootstrap Variance Estimator for Movement Estimates

The bootstrap procedure for estimating the variance of movement estimates for stratified random samples selected without replacement is to: (i) Divide the units selected in the survey into the three categories: $s_h^{(c)} = s_h^{(1)} \cap s_h^{(2)}$, $s_h^{(1\bar{c})} = s_h^{(1)} \cap \overline{s_h^{(2)}}$ and $s_h^{(2\bar{c})} = \overline{s_h^{(1)}} \cap s_h^{(2)}$. (ii) Within each of the three categories select simple random samples $s_{rh}^{(c)}$, $s_{rh}^{(1\bar{c})}$ and $s_{rh}^{(2\bar{c})}$ without replacement (SRSWOR) of $m_h^{(c)} = [n_h^{(c)}/2]$, $m_h^{(1\bar{c})} = [n_h^{(1\bar{c})}/2]$ and $m_h^{(2\bar{c})} = [n_h^{(2\bar{c})}/2]$ units from the original samples of $n_h^{(c)}$, $n_h^{(1\bar{c})}$ and $n_h^{(2\bar{c})}$ units, independently within each stratum h. Let $\delta_{rhi}^{(c)}$, $\delta_{rhi}^{(1\bar{c})}$ and $\delta_{rhi}^{(2\bar{c})}$ equal 1 if $i \in s_{rh}^{(c)}$, $i \in s_{rh}^{(1\bar{c})}$ and $i \in s_{rh}^{(2\bar{c})}$, and zero otherwise. Calculate the bootstrap sampling weights:

$$w_{ri}^{(t)} = w_i^{(t)} \left(1 - \sqrt{\frac{(1-f_h^{(c)})m_h^{(c)}}{(n_h^{(c)} - m_h^{(c)})}} + \sqrt{\frac{(1-f_h^{(c)})m_h^{(c)}}{(n_h^{(c)} - m_h^{(c)})}} \frac{n_h^{(c)}}{m_h^{(c)}} \delta_{rhi}^{(c)} \right) \quad \text{, if } i \in s_h^{(c)}$$

$$w_{ri}^{(t)} = w_i^{(t)}\left(1-\sqrt{\frac{(n_h^{(t)}(1-f_h^{(t)})-n_h^{(c)}(1-f_h^{(c)}))m_h^{(t\bar{c})}}{n_h^{(t\bar{c})}(n_h^{(t\bar{c})}-m_h^{(t\bar{c})})}}\right) + w_i^{(t)}\left(\sqrt{\frac{(n_h^{(t)}(1-f_h^{(t)})-n_h^{(c)}(1-f_h^{(c)}))m_h^{(t\bar{c})}}{n_h^{(t\bar{c})}(n_h^{(t\bar{c})}-m_h^{(t\bar{c})})}}\frac{n_h^{(t\bar{c})}}{m_h^{(t\bar{c})}}\delta_{rhi}^{(t\bar{c})}\right), \text{ if } i\in s_h^{(t\bar{c})} \tag{4.4}$$

where $f_h^{(c)} = \dfrac{n_h^{(1)}n_h^{(2)}N_h^{(c)}}{n_h^{(c)}N_h^{(1)}N_h^{(1)}}$. Calculate the bootstrap movement estimator of $Y_c^{(m)}$:

$$\hat{Y}_{rc}^{(m)} = \sum_{i\in s^{(2)}} w_{ri}^{(2)}g_{ri}^{(2)}y_i - \sum_{i\in s^{(1)}} w_{ri}^{(1)}g_{ri}^{(1)}y_i \tag{4.5}$$

(iii) Independently replicate step (ii) a large number of times, R, and calculate the bootstrap estimates, $\hat{Y}_{1c}^{(m)}, \hat{Y}_{2c}^{(m)}, \ldots, \hat{Y}_{Rc}^{(m)}$.

(iv) The bootstrap variance estimator is given by the Monte Carlo approximation:

$$Var(\hat{Y}_c) = \frac{\sum_r\left(\hat{Y}_{rc}^{(m)} - \overline{\hat{Y}_c^{(m)}}\right)^2}{R-1} \tag{4.6}$$

where $\overline{\hat{Y}_c^{(m)}} = \dfrac{\sum_r \hat{Y}_{rc}^{(m)}}{R}$.

4.3 The Bootstrap Variance Estimator for Decomposition Components

The bootstrap movement estimator (4.5) can be written as the difference between a product of four factors:

$$\hat{Y}_{rc}^{(m)} = \sum_h N_h^{(2)}\hat{p}_{rhc}^{(2)}\bar{g}_{rhc}^{(2)}\bar{y}_{rhc}^{(2)} - \sum_h N_h^{(1)}\hat{p}_{rhc}^{(1)}\bar{g}_{rhc}^{(1)}\bar{y}_{rhc}^{(1)} \tag{4.7}$$

where $\hat{p}_{rhc}^{(t)}$, $\bar{g}_{rhc}^{(t)}$ and $\bar{y}_{rhc}^{(t)}$ represent the bootstrap estimates of $\hat{p}_{hc}^{(t)}$, $\bar{g}_{hc}^{(t)}$ and $\bar{y}_{hc}^{(t)}$:

$$\hat{p}_{rhc}^{(t)} = \frac{\sum_{i\in s_{hc}^{(t)}} w_{ri}^{(t)}}{\sum_{i\in s_h^{(t)}} w_{ri}^{(t)}}$$

$$\bar{g}_{rhc}^{(t)} = \frac{\sum_{i\in s_{hc}^{(t)}} w_{ri}^{(t)}g_{ri}^{(t)}y_i^{(t)}}{\sum_{i\in s_{hc}^{(t)}} w_{ri}^{(t)}y_i^{(t)}}$$

$$\bar{y}_{rhc}^{(t)} = \frac{\sum_{i\in s_{hc}^{(t)}} w_{ri}^{(t)}y_i^{(t)}}{\sum_{i\in s_{hc}^{(t)}} w_{ri}^{(t)}}$$

Furthermore, the decomposition components (3.5 - 3.17) can be written as a function of $N_h^{(t)}$, $\hat{p}_{hc}^{(t)}$, $\bar{g}_{hc}^{(t)}$ and $\bar{y}_{hc}^{(t)}$:

$$\hat{Y}_c^{(*)} = \psi^{(*)}(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \qquad (4.8)$$

Therefore, the bootstrap procedure for estimating the variance of decomposition components is to: (i) Calculate the bootstrap estimator of $\hat{Y}_{1c}^{(*)}$, $\hat{Y}_{2c}^{(*)}$, ..., $\hat{Y}_{Rc}^{(*)}$, where:

$$\hat{Y}_{rc}^{(*)} = \psi^{(*)}(N_h^{(1)}, N_h^{(2)}, \hat{p}_{rhc}^{(1)}, \hat{p}_{rhc}^{(2)}, \bar{g}_{rhc}^{(1)}, \bar{g}_{rhc}^{(2)}, \bar{y}_{rhc}^{(1)}, \bar{y}_{rhc}^{(2)}) \qquad (4.9)$$

(ii) The bootstrap variance estimator of the decomposition components is given by the Monte Carlo approximation:

$$Var\left(\hat{Y}_c^{(*)}\right) = \frac{\sum_r \left(\hat{Y}_{rc}^{(*)} - \overline{\hat{Y}_c^{(*)}}\right)^2}{R-1} \qquad (4.10)$$

where $\overline{\hat{Y}_c^{(*)}} = \frac{\sum_r \hat{Y}_{rc}^{(*)}}{R}$.

QUESTION 5: Is it appropriate to use the bootstrap replicate weights to estimate the variance of the decomposition components?

5. Simulation Study

A Monte Carlo simulation study was undertaken to measure the bias and variability of the decomposition components and the bootstrap variance estimator of the decomposition components. In the Monte carlo study, an artificial population U of 12,000 units was split into three categories: $U^{(c)} = U^{(1)} \cap U^{(2)}$ of 9,000 units, $U^{(1\bar{c})} = U^{(1)} \cap \overline{U^{(2)}}$ of 1,000 units and $U^{(2\bar{c})} = \overline{U^{(1)}} \cap U^{(2)}$ of 2,000 units. The characteristics of the population was generated using the following models:

$$y_i^{(1)} = |x_i + \psi_i^{(1)}|$$
$$y_i^{(2)} = 1.1 \times |x_i + \psi_i^{(2)}|$$

where $x_i = abs(\zeta_i)$, $\zeta_i = N(0, 22.5)$ if $i \in U^{(1\bar{c})}$, $\zeta_i = N(0, 25)$ if $i \in U^{(c)}$, $\zeta_i = N(0, 27.5)$ if $i \in U^{(2\bar{c})}$, $\psi_i^{(1)} = N(0, \sqrt{x_i})$ and $\psi_i^{(2)} = N(0, \sqrt{x_i})$.

All units in the population were randomly assigned to one of two stratum groups, and then within these two stratum groups units were assigned to one of five strata, using stratum boundaries 0-9, 10-19, 20-29, 30-49, 50+ based on $z_i^{(1)} = int|x_i + \xi_i^{(1)}|$ at the first time period and $z_i^{(2)} = int|x_i + \xi_i^{(2)}|$ at the second time period, where $\xi_i^{(1)} = N(0, \sqrt{x_i})$ and $\xi_i^{(2)} = N(0, \sqrt{x_i})$.
All units in the population were randomly assigned to one of two domains of interest. The simulation population characteristics of the variable of interest are presented in Table 3.

Table 3: Simulation Population Characteristics of the Variable of Interest

Domain	Time 1 Population	Time 2 Population	Movement Population	Time 1 Common Population#	Time 2 Common Population#	Movement Common Population#
1	97,847	122,950	25,103	55,344	61,099	5,756
2	100,478	125,135	24,657	56,879	62,745	5,866
Domain	**Time 1 Death Population**	**Time 2 Birth Population**	**Movement Birth Death Population**	**Time 1 Flip In Population**	**Time 2 Flip Out Population**	**Movement Stratum Flip Population**
1	8,487	24,484	15,997	34,106	37,367	3,351
2	9,263	24,689	15,426	34,336	37,701	3,365

#The common population is defined as those units in the same stratum at both time periods.

A total of 10,000 independent simulated stratified samples were selected from the population. In the first time period, stratified simple random samples without replacement of 200 units, with equal allocation of $n_h^{(1)} = 20$, were selected from the population at the first time period, $U^{(1)}$. In the second time period, a stratified synchronised random samples of 220 units, with equal allocation of $n_h^{(2)} = 22$ and $n_h^{(c)} = 14$, were selected from the population at the second time period, $U^{(2)}$.

The estimator used this study was the generalised regression estimator (2.2), with $\underset{\sim}{X}$ equal to the population totals for the two auxiliary variables $\underset{\sim}{x}_i = \{x_{1i}, x_{2i}\}$ at the stratum group level, and $c_i\sigma^2 = 1$.

The relative biases (RB) and the relative root mean squared errors (RRMSE) of the decomposition components were estimated by:

$$RB\left(\hat{Y}_c^{(*)}\right) = \frac{1}{Y_c^{(m)}}\left[\frac{1}{10,000}\sum_{k=1}^{10,000}\left(\hat{Y}_{kc}^{(*)} - Y_c^{(*)}\right)\right]$$

$$RRMSE\left(\hat{Y}_c^{(*)}\right) = \frac{1}{Y_c^{(m)}}\sqrt{\frac{1}{10,000}\sum_{k=1}^{10,000}\left(\hat{Y}_{kc}^{(*)} - Y_c^{(*)}\right)^2}$$

where $\hat{Y}_{kc}^{(*)}$ are the decomposition components (3.5 - 3.17) under the kth simulation sample and $Y_c^{(*)}$ are the decomposition components (3.5 - 3.17) under a complete census of the population at the two time periods. The estimated RB and RRMSE of the decomposition components for the variable of interest are presented in Table 4.

Table 4: Estimated Relative Biases and Relative Root Mean Squared Errors of the Decomposition Components

Domain	Movement $\hat{Y}_c$	Common Sample Unit Effect $\hat{Y}_c^{(muc)}$	Net Birth Death Effect $\hat{Y}_c^{(mbd)}$	Net Stratum Flip Effect $\hat{Y}_c^{(msf)}$	Net Rotations Effect $\hat{Y}_c^{(mr)}$	Domain Change Effect $\hat{Y}_c^{(md)}$	Weight Adjust Change Effect $\hat{Y}_c^{(mw)}$
	Population Values						
1	25,103	10,870	13,531	235	0	468	0
2	24,657	10,894	14,260	-11	0	-485	0
	Average Simulation Estimates						
1	24,897	10,777	13,794	-492	19	338	461
2	25,011	10,884	14,568	-621	2	-325	505
	Estimated Relative Biases (%)						
1	-0.8	-0.4	1.0	-2.9	0.1	-0.5	1.8
2	1.4	0.0	1.2	-2.5	0.0	0.6	2.0
	Estimated Relative Root Mean Squared Errors (%)						
1	29.3	6.0	9.4	8.6	5.9	26.6	7.5
2	29.9	6.1	9.7	8.6	6.0	27.0	8.1

The results of the simulation study found that the estimated relative biases of the decomposition components are quite small. These results indicate that the expected value of the common sample unit effect (3.11) is approximately equal to the common population unit effect, but it is not equal to the difference in the population values of the common population units:

$$E(\hat{Y}^{(muc)}) \approx Y^{(muc)} \neq Y^{(2c)} - Y^{(1c)}$$

Furthermore, the expected value of the net birth-death effect (3.10 plus 3.13 plus 3.16) and the net stratum flip effect (3.9 plus 3.14 plus 3.17) are approximately equal to the true population net birth-death effect and net stratum flip effect, but they are not equal to the difference in the population values of their respective population units:

$$E(\hat{Y}^{(mfg)} + \hat{Y}^{(mub)} + \hat{Y}^{(mud)}) \approx Y^{(mfg)} + Y^{(mub)} + Y^{(mud)} \neq Y^{(2\bar{c})} - Y^{(1\bar{c})}$$

$$E(\hat{Y}^{(mff)} + \hat{Y}^{(mufi)} + \hat{Y}^{(mufo)}) \approx Y^{(mff)} + Y^{(mfi)} + Y^{(mufo)} \neq Y^{(2fi)} - Y^{(2fo)}$$

QUESTION 6: Is it necessary for the expected value of the common sample unit effect, the net birth-death effect, and the net stratum flip effect to be equal to the differences in the population values of their respective population units?

The RB and RRMSE for the bootstrap standard error estimator of the decomposition components were estimated by:

$$RB\left(SE\left(\hat{Y}_c^{(*)}\right)\right) = \frac{1}{\widehat{SE(Y_c^{(*)})}}\left[\frac{1}{10,000}\sum_{k=1}^{10,000}\left(SE\left(\hat{Y}_{kc}^{(*)}\right) - \widehat{SE(Y_c^{(*)})}\right)\right]$$

$$RRMSE\left(SE\left(\hat{Y}_c^{(*)}\right)\right) = \frac{1}{\widehat{SE(Y_c^{(*)})}}\sqrt{\frac{1}{10,000}\sum_{k=1}^{10,000}\left(SE\left(\hat{Y}_{kc}^{(*)}\right) - \widehat{SE(Y_c^{(*)})}\right)^2}$$

where $SE\left(\hat{Y}_{kc}^{(*)}\right)$ are the bootstrap standard errors of the decomposition components under the k[th] simulation sample and $\widehat{SE(Y_c^{(*)})} = \sqrt{\frac{1}{10,000}\sum_{k=1}^{10,000}\left(\hat{Y}_{kc}^{(*)} - Y_c^{(*)}\right)^2}$ is the estimated true standard error of the decomposition components. The estimated RB and RRMSE for the bootstrap standard error estimator of the decomposition components for the variable of interest are presented in Table 5.

Table 5: Estimated Relative Biases and Relative Root Mean Squared Errors for the Bootstrap Standard Error of the Decomposition Components

Domain	Movement $\hat{Y}_c$	Common Sample Unit Effect $\hat{Y}_c^{(muc)}$	Net Birth Death Effect $\hat{Y}_c^{(mbd)}$	Net Stratum Flip Effect $\hat{Y}_c^{(msf)}$	Net Rotations Effect $\hat{Y}_c^{(mr)}$	Domain Change Effect $\hat{Y}_c^{(md)}$	Weight Adjust Change Effect $\hat{Y}_c^{(mw)}$
	Estimated Population Values						
1	7,351	1,507	2,337	2,023	1,489	6,665	1,827
2	7,358	1,516	2,375	2,023	1,483	6,649	1,933
	Average Simulation Estimates						
1	7,488	1,516	2,466	2,249	1,632	6,797	2,259
2	7,498	1,531	2,494	2,225	1,620	6,784	2,357
	Estimated Relative Biases (%)						
1	1.9	0.6	5.6	11.2	9.6	2.0	23.6
2	1.9	1.0	5.0	10.0	9.2	2.0	21.9
	Estimated Relative Root Mean Squared Errors (%)						
1	8.8	11.4	16.6	23.6	23.2	8.6	30.0
2	8.8	12.1	16.9	24.1	22.9	8.7	28.2

The results of the simulation study found that the estimated relative biases for the bootstrap variance estimator of the decomposition components are reasonably small, with the exception of the weight change adjustment effect.

QUESTION 7: Is the size of the estimated relative biases for the bootstrap variance estimator of the weight change adjustment effect problematic?

6. Using the Decomposition Methodology to Identify Irregularities

The decomposition methodology has dual purposes. Firstly, it can be used as a tool to better understand the key factors that are driving the movements in the estimates. It has the potential to identify problems which might influence the movements in the estimates, such as unexpected levels of birthing or deathing of units in the population of interest, unusual combinations of units rotating into and out of the sample, and errors in the auxiliary benchmark values or auxiliary benchmark totals.

Secondly, it can be used as a macro editing tool to assist in the output editing of sample surveys. It has the potential to identify suspicious movements in the aggregate level estimates (using the decomposition components presented in Section 3.1 and 3.2), and then drill down to the micro level data to identify units with the largest impact on the movement estimates (using the unit level contributions to the decomposition components presented in Section 6.1). The suspicious movements in the aggregate level estimates can be investigated by looking at the size of decomposition components relative to the standard errors on the decomposition components as well as looking at the size of the decomposition components compared to the time series of the decomposition components.

6.1 Unit Contributions to Decomposition Components

The unit contributions to the movement estimates will have contributions from each of the four factors. The contribution of unit i to the frame change effect (3.5) is give by:

$$\hat{Y}_{ci}^{(mf)} = \varphi(\hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in U_h^{(fi)}, U_h^{(b)}$$

$$\hat{Y}_{ci}^{(mf)} = -\varphi(\hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in U_h^{(fo)}, U_h^{(d)}$$

The contribution to unit i to the domain change effect (3.6) is given by:

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}}{2n_{hc}^{(1)}} + \frac{\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}}{2n_{hc}^{(2)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in s_h^{(c)}$$

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}}{2n_{hc}^{(2)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in s_h^{(ri)}, s_h^{(fi)}, s_h^{(b)}$$

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\hat{p}_{hc}^{(2)} - \hat{p}_{hc}^{(1)}}{2n_{hc}^{(1)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in s_h^{(ro)}, s_h^{(fo)}, s_h^{(d)}$$

The contribution of unit i to the weight adjustment change effect (3.7) is given by:

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}}{2n_{hc}^{(1)}} + \frac{\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}}{2n_{hc}^{(2)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in s_h^{(c)}$$

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}}{2n_{hc}^{(2)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \quad \text{, if } i \in s_h^{(ri)}, s_h^{(fi)}, s_h^{(b)}$$

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\bar{g}_{hc}^{(2)} - \bar{g}_{hc}^{(1)}}{2n_{hc}^{(1)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{y}_{hc}^{(1)}, \bar{y}_{hc}^{(2)}) \qquad \text{, if } i \in s_h^{(ro)}, s_h^{(fo)}, s_h^{(d)}$$

The contribution of unit i to the average value change effect (3.8) is given by:

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{y_{hi}^{(2)}}{n_{hc}^{(c)}} - \frac{y_{hi}^{(1)}}{n_{hc}^{(c)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \qquad \text{, if } i \in s_h^{(c)}$$

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{y_{hi}^{(2)}}{n_{hc}^{(2)}} - \frac{\bar{y}_{hc}^{(2)}}{n_{hc}^{(2)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \qquad \text{, if } i \in s_h^{(ri)}, s_h^{(fi)}, s_h^{(b)}$$

$$\hat{Y}_{ci}^{(mu)} = \left[\frac{\bar{y}_{hc}^{(1)}}{n_{hc}^{(1)}} - \frac{y_{hi}^{(1)}}{n_{hc}^{(1)}} \right] \varphi(N_h^{(1)}, N_h^{(2)}, \hat{p}_{hc}^{(1)}, \hat{p}_{hc}^{(2)}, \bar{g}_{hc}^{(1)}, \bar{g}_{hc}^{(2)}) \qquad \text{, if} i \in s_h^{(ro)}, s_h^{(fo)}, s_h^{(d)}$$

6.2 Ability of the Decomposition Methodology to Identify Irregularities

(i) Example 1: Unusual Rotation Combinations

A single simulation sample from the simulation study was used to examine the impact of unusual combinations of units rotating into and out of the sample. Various combinations of the smallest and largest units were rotated into and out of the sample at the second time period. The estimates and standard errors on the decomposition components of using these unusual rotation combinations are presented in Table 6.

Table 6: Impact on Decomposition Components of Unusual Rotation Combinations

Rotation Combinations	Time 1 Estimate $\hat{Y}_c^{(1)}$	Time 2 Estimate $\hat{Y}_c^{(2)}$	Movement Estimate $\hat{Y}_c^{(m)}$	Common Sample Effect $\hat{Y}_c^{(muc)}$	Net Birth Death Effect $\hat{Y}_c^{(mbd)}$	Net Stratum Flip Effect $\hat{Y}_c^{(msf)}$	Net Rotations Effect $\hat{Y}_c^{(mr)}$	Domain Change Effect $\hat{Y}_c^{(md)}$	Weight Adjust Change Effect $\hat{Y}_c^{(mw)}$
No adjustments	103,302	129,965	26,664* (7,197)	12,004* (1,467)	14,533* (2,580)	-206 (2,679)	474 (1,524)	-468 (6,539)	307 (2,213)
Smallest units rotated out	103,302	133,915	30,614* (7,176)	12,745* (1,573)	13,598* (2,531)	-133 (2,529)	4,383* (1,847)	2,358 (6,991)	-2,337 (2,167)
Largest units rotated out	103,302	125,404	22,102* (8,227)	11,596* (1,526)	14,770* (2,599)	-470 (2,368)	-6,616* (2,323)	-246 (6,446)	3,068 (2,302)
Smallest units rotated in	103,302	122,518	19,216* (6,965)	11,955* (1,473)	14,109* (2,636)	-81 (2,685)	-7,259* (2,902)	-2,739 (6,295)	3,230 (2,189)
Largest units rotated in	103,302	125,503	22,202* (8,188)	11,407* (1,424)	14,089* (2,490)	-104 (2,594)	7,598* (3,287)	-5,787 (6,748)	-5,002 (2,648)

* Significantly different from zero at the 95% confidence level.

The unusual rotation combinations had substantial impacts on the net rotations effect. Since the net rotations effect were significantly different from zero at the 95% confidence level, further investigations would be undertaken in order to identify the underlying reasons behind the large net rotations effects. The unusual rotation combinations also had an large impact on the weight adjustment change effect. The larger weight adjustment change effects occurred because units with the very small and very large auxiliary benchmark values were rotated into and out of the sample at the second time period, since the reported values are highly correlated with benchmark variables.

(ii) Example 2: Invalid Benchmark Totals

A single simulation sample from the simulation study was used to examine the impact of errors in the auxiliary benchmark totals. The auxiliary benchmark totals at the second time period were modified by constant factors ranging between 0.95 to 1.05. The estimates and standard errors on the decomposition components of using these invalid auxiliary benchmark totals are presented in Table 7.

Table 7: Impact on Decomposition Components of Invalid Benchmark Totals

Benchmark Total Adjust Factor	Time 1 Estimate $\hat{Y}_c^{(1)}$	Time 2 Estimate $\hat{Y}_c^{(2)}$	Movement Estimate $\hat{Y}_c^{(m)}$	Common Sample Effect $\hat{Y}_c^{(muc)}$	Net Birth Death Effect $\hat{Y}_c^{(mbd)}$	Net Stratum Flip Effect $\hat{Y}_c^{(msf)}$	Net Rotations Effect $\hat{Y}_c^{(mr)}$	Domain Change Effect $\hat{Y}_c^{(md)}$	Weight Adjust Change Effect $\hat{Y}_c^{(mw)}$
1.00	100,141	125,212	25,072* (8,400)	11,805* (1,729)	15,780* (2,516)	-120 (2,271)	-317 (1,555)	-1,811 (7,696)	-267 (2,088)
0.95	100,141	92,474	-7,666 (6,910)	10,269* (1,541)	13,598* (2,198)	-120 (2,047)	-258 (1,349)	-1,722 (6,619)	-29,434* (2,839)
0.99	100,141	117,980	17,840* (8,000)	11,466* (1,687)	15,298* (2,445)	-120 (2,222)	-304 (1,509)	-1,791 (7,457)	-6,710* (2,049)
1.01	100,141	132,815	32,674* (8,853)	12,162* (1,774)	16,287* (2,590)	-120 (2,325)	-331 (1,603)	-1,831 (7,947)	6,507* (2,266)
1.05	100,141	167,242	67,102* (11,214)	13,778* (1,978)	18,582* (2,932)	-120 (2,567)	-393 (1,824)	-1,924 (9,089)	37,179* (4,039)

* Significantly different from zero at the 95% confidence level.

Errors in the auxiliary benchmark totals as little as one percent had substantial impacts on the weight adjustment change effect. Since the weight adjustment change effect were significantly different from zero at the 95% confidence level, further investigations would be undertaken in order to identify the underlying reasons behind the large weight adjustment change effect.

(ii) Example 3: Extreme Reported Values

A single simulation sample from the simulation study was used to examine the impact of extreme reported values. A total of four units; one common sample unit (#01011), one sample rotation in (#02609), one sample flip in (#01500) and one sample birth (#01080); were modified by adding 100 to their reported values at the second time period. The estimates and standard errors on the decomposition components of using these four units with extreme reported values are presented in Table 8.

Table 8: Impact on Decomposition Components of Extreme Reported Values

Extreme Values	Time 1 Estimate $\hat{Y}_c^{(1)}$	Time 2 Estimate $\hat{Y}_c^{(2)}$	Movement Estimate $\hat{Y}_c^{(m)}$	Common Sample Effect $\hat{Y}_c^{(muc)}$	Net Birth Death Effect $\hat{Y}_c^{(mbd)}$	Net Stratum Flip Effect $\hat{Y}_c^{(msf)}$	Net Rotations Effect $\hat{Y}_c^{(mr)}$	Domain Change Effect $\hat{Y}_c^{(md)}$	Weight Adjust Change Effect $\hat{Y}_c^{(mw)}$
No extreme values	96,651	123,561	26,910* (6,939)	10,570* (1,516)	17,905* (2,675)	-906 (2,618)	-475 (1,114)	280 (6,163)	-463 (2,318)
Extreme values	96,651	146,833	50,182* (14,536)	20,379* (9,976)	17,856* (4,045)	4,482 (6,914)	4,808 (4,850)	1,925 (6,553)	733 (2,593)

* Significantly different from zero at the 95% confidence level.

Although these four units with extreme reported values did not result in the net stratum flip effect or the net rotations effect to be significantly different from zero at the 95% confidence level, further investigations would be undertaken in order to identify the underlying reasons behind the unexpectedly large movement estimate. The ten largest unit contributions to the movement estimates using these four extreme reported values are presented in Tables 9 and 10.

Table 9: Largest Unit Contributions to Movement Estimates Without Extreme Values

Unit	Time 1 Value	Time 2 Value	Decomposition Effect	Frame Change Contribution	Domain Change Contribution	Weight Adjustment Change Contribution	Unit Value Contribution	Total Contribution
A	n.a.	13.6	Flip In	18.3	115.3	-11.8	-1,392.9	-1,271.1
B	28.2	n.a.	Flip Out	-5.4	-38.0	14.3	-1,124.9	-1,154.0
C	n.a.	64.0	Birth	18.3	115.3	-11.8	765.8	887.6
D	27.3	n.a.	Flip Out	-18.3	144.1	-14.7	774.2	885.2
E	53.3	60.3	Common Unit	0.0	259.4	-26.5	596.9	829.8
F	n.a.	24.4	Birth	5.4	-38.0	14.3	803.6	785.2
G	59.2	64.4	Common Unit	0.0	259.4	-26.5	448.4	681.2
H	46.6	n.a.	Rotations Out	-23.1	-89.7	-23.9	-533.9	-670.6
I	38.8	43.8	Common	0.0	259.4	-26.5	430.3	663.2
J	56.2	n.a.	Flip Out	-18.3	144.1	-14.7	-769.4	-658.3

Table 10: Largest Unit Contributions to Movement Estimates With Extreme Values

Unit	Time 1 Value	Time 2 Value	Decomposition Effect	Frame Change Contribution	Domain Change Contribution	Weight Adjustment Change Contribution	Unit Value Contribution	Total Contribution
K	5.5	105.5	Common Unit	0.0	274.1	181.1	9,810.9	10,266.1
L	n.a.	125.5	Flip In	10.2	6.9	9.6	6,566.8	6,593.5
M	n.a.	100.6	Rotations In	6.5	112.1	74.1	4,968.6	5,161.3
N	n.a.	166.2	Birth	32.9	-45.4	-37.3	1,510.9	1,461.2
A	n.a.	13.5	Flip In	18.3	115.3	-11.8	-1,392.9	-1,271.1
B	28.2	n.a.	Flip Out	-5.4	-38.0	14.3	-1,124.9	-1,154.0
C	n.a.	64.0	Birth	18.3	115.3	-11.8	765.8	887.6
D	27.3	n.a.	Flip Out	-18.3	144.1	-14.7	774.2	885.2
E	53.3	60.3	Common Unit	0.0	259.4	-26.5	596.9	829.8
O	n.a.	1.6	Birth	6.5	112.1	74.1	-1,003.3	-810.5

The four units with extreme reported values were the four largest contributors to the movement estimates, and since the contribution of three of these units were much greater than the other units, these units would be given the highest priority in the output editing of the survey.

7. Conclusion

The proposed decomposition of the movement estimates into primary and secondary factors is just one of many possible approaches to the decomposition of the movement estimates. This decomposition methodology is consistent with the generalised regression estimator currently used in ABS business surveys, and the existing bootstrap variance estimation methodology can be used to produce approximate variances for the various decomposition components.

This decomposition methodology can be used as a tool to better understand the key factors that are driving the movements in the estimates and as a macro editing tool to assist in the output editing of sample surveys. It has the ability to identify suspicious movements in the aggregate level estimates and then to drill down to the micro level data to identify units with the largest impact on the movement estimates.

This decomposition methodology has been implemented in a number of ABS Business Surveys, although the variances for the decomposition components have yet to be implemented. Although the decomposition methodology has proven useful to users, the users have had some difficulties with the interpretation of the decomposition components and understanding the significance of the size of the decomposition components. It is envisaged that the availability of variances for the decomposition components, longer time series of the decomposition components, and training on the interpretation of the decomposition components might alleviate this limitation. Another limitation of the decomposition methodology is the lack of a link to the final published seasonally adjusted movement estimates.

8. References

Brewer, K.R.W., Gross W.F., and Lee G.F. (1999). PRN Sampling: The Australian Experience, *ISI Proceedings: Invited Papers, Helsinki August 10-18, 1999*, 155-163.

Das Gupta, P. (1991). Decomposition of the difference between two rates and its consistency when more than two populations are involved. *Mathematical Population Studies,* 3(2), 105-125.

Holt, D. and Skinner, C.J. (1989). Components of change in repeated surveys. *International Statistics Review,* 57, 1-18.

Kitagawa, E.M. (1955). Components of the difference between two rates. *Journal of the American Statistical Association,* 50, 1168-1194.

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

Shao, J. and Tu, D. (1995). *The Jacknife and Bootstrap.* Springer, New York.

Appendix: Descriptions of the Movement Decomposition Components

Decomposition Component	Description	Positive Effect	Negative Effect
Frame Change Effect (3.5)	Measures the impact on the movement estimates of changes in the number of units on the survey frame between the two time periods.		
Frame Growth Effect (3.9)	Measures the impact on the movement estimates of the net difference between the number of units birthed and deathed on the survey frame between the two time periods.	Larger number of units birthed on the survey frame than deathed on the survey frame, or units deathed on the survey frame occur within strata with a lower average value of units, or units birthed on the survey frame occur within strata with a higher average value of units.	Smaller number of units birthed on the survey frame than deathed on the survey frame, or units deathed on the survey frame occur within strata with a higher average value of units, or units birthed on the survey frame occur within strata with a lower average value of units.
Frame Stratum Flip Effect (3.10)	Measures the impact on the movement estimates of the net difference between the number of units on the survey frame that flipped into and out of strata between the two time periods.	Units have flipped out of strata with a lower average value of units and flipped into strata with a higher average value of units.	Units have flipped out of strata with a higher average value of units and flipped into strata with a lower average value of units.
Domain Change Effect (3.6)	Measures the impact on the movement estimates of changes in the estimated proportion of units within the domain of interest between the two time periods.	Higher estimated proportion of units within domain of interest at the second time period than at the first time period.	Lower estimated proportion of units within domain of interest at the second time period than at the first time period.
Weight Adjustment Change Effect (3.7)	Measures the impact on the movement estimates of changes in the weighted average g-weight between the two time periods.	Higher weighted average g-weights at the second time period than at the first time period.	Lower weighted average g-weights at the second time period than at the first time period.
Average Value Change Effect (3.8)	Measures the impact on the movement estimates of changes in the average value of units of the variable of interest between the two time periods.		
Common Sample Unit Effect (3.11)	Measures the impact on the movement estimates of differences in the average value of common sample units between the two time periods.	Higher average value of common sample units at the second time period than at the first time period	Lower average value of common sample units at the second time period than at the first time period

Decomposition Component	Description	Positive Effect	Negative Effect
Sample Rotations In Effect (3.12)	Measures the impact on the movement estimates of differences in the average value of sample rotations in and common sample units at the second time period.	Higher average value of sample rotations in than common sample units at the second time period.	Low average value of sample rotations in than common sample units at the second time period.
Sample Stratum Flip In Effect (3.13)	Measures the impact on the movement estimates of differences in the average value of sample that flipped into strata and common sample units at the second time period.	Higher average value of sample flips in than common sample units at the second time period.	Low average value of sample flips in than common sample units at the second time period.
Sample Birth Effect (3.14)	Measures the impact on the movement estimates of differences in the average value of sample births and common sample units at the second time period.	Higher average value of sample births than common sample units at the second time period.	Low average value of sample births than common sample units at the second time period.
Sample Rotations Out Effect (3.15)	Measures the impact on the movement estimates of differences in the average value of sample rotations out and common sample units at the first time period.	Lower average value of sample rotations out than common sample units at the first time period.	Higher average value of sample rotations out than common sample units at the first time period.
Sample Stratum Flip Out Effect (3.16)	Measures the impact on the movement estimates of differences in the average value of sample that flipped out of strata and common sample units at the first time period.	Lower average value of sample flips out than common sample units at the first time period.	Higher average value of sample flips out than common sample units at the first time period.
Sample Death Effect (3.17)	Measures the impact on the movement estimates of differences in the average value of sample deaths and common sample units at the first time period.	Lower average value of sample deaths than common sample units at the first time period.	Higher average value of sample deaths than common sample units at the first time period.

FOR MORE INFORMATION . . .

INTERNET	**www.abs.gov.au** the ABS web site is the best place for data from our publications and information about the ABS.
LIBRARY	A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our web site for a list of libraries.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our web site, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

PHONE	1300 135 070
EMAIL	client.services@abs.gov.au
FAX	1300 135 211
POST	Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All ABS statistics can be downloaded free of charge from the ABS web site.

WEB ADDRESS www.abs.gov.au

2000001524220
ISBN 0 642 48177 6

RRP $11.00

1352.0.55.073 • RESEARCH PAPER: DECOMPOSITION OF MOVEMENT ESTIMATES AS A DIAGNOSTIC TOOL FOR REPEATED BUSINESS SURVEYS (METHODOLOGY ADVISORY COMMITTEE) • November 2005