1352.0.55.024

Australian Bureau of Statistics

Research Paper

# Design for the Quarterly Economy Wide Survey

New Issue

# Research Paper

# Design for the Quarterly Economy Wide Survey


Edward Szoldra and Louise Gates

Statistical Services Branch


Methodology Advisory Committee

18 June 1999, Canberra


AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30 AM (CANBERRA TIME) WED 1 NOV 2000

ABS Catalogue no. 1352.0.55.024
ISBN 0 642 48128 8



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 Edward Szoldra, Statistical Services Branch on Sydney (02) 9268 4214 or email <edward.szoldra@abs.gov.au>.

# CONTENTS

PART I: EXCLUSION OF SMALL BUSINESSES FROM THE QEWS

1. INTRODUCTION . . . . . . . . . 1
2. SMALL BUSINESSES . . . . . . . . . 2
3. RESULTS . . . . . . . . . 6

APPENDIXES . . . . . . . . . 9

PART II: DIRECT MOVEMENT ESTIMATION IN THE QEWS

1. INTRODUCTION . . . . . . . . . 15
2. BACKGROUND . . . . . . . . . 16
3. THEORETICAL EVALUATION . . . . . . . . . 21
4. EMPIRICAL EVALUATION . . . . . . . . . 28
5. EXCLUSION OF SMALL BUSINESSES . . . . . . . . . 35
6. SUMMARY . . . . . . . . . 38
7. POINTS FOR DISCUSSION . . . . . . . . . 39

APPENDIX . . . . . . . . . 40

# PART I: EXCLUSION OF SMALL BUSINESSES FROM THE QUARTERLY ECONOMY WIDE SURVEY

Edward Szoldra and Louise Gates
Statistical Services Branch

## 1. INTRODUCTION

As a result of the small business deregulation taskforce, the ABS is under an obligation to reduce the compliance cost for small businesses. It has also been suggested that the removal of small businesses from ABS surveys despite the statistical bias would still significantly improve estimates in mean square error terms as there is a high non-sampling error for these businesses. One method of completely removing the cost for small businesses is to remove them from the data collection in ABS surveys, but to estimate for them in some other way. As part of the development of the new Quarterly Economy Wide Survey (QEWS), it was decided to investigate the impact of the removal of small businesses.

This paper discusses the feasibility of this approach.

## 2. SMALL BUSINESSES

Businesses are defined as small based on employment. There are 685,000 businesses with less than 5 employees in scope of the QEWS (77% of in scope businesses) and 862,000 businesses with less than 20 employees (97%). The QEWS is a new survey, consisting of the amalgamation of the surveys: Survey of New Capital Expenditure (CAPEX), Survey of Stocks and Sales (STX and Sales), Survey of Employment and Earnings (SEE) and the Survey of Company Profits. The Survey of Company Profits is not considered further as its scope is already businesses with employment greater than 19.

The following table shows the number and proportion of small businesses in the samples of CAPEX, STX and Sales and SEE. The data item used for the SEE is Wages and Salaries. As can be seen in the table, the proportion of small businesses in the sample is considerably smaller than the proportion of small businesses in the population due to the optimal sampling strategy used. The table also shows that small businesses make up a large proportion of the population size and of the variance, however only a small proportion of the total estimate.

**Table 1. Contribution of small businesses to sample size, total estimate and total variance**

| | *CAPEX* | *STX* | *Sales* | *SEE* |
|---|---|---|---|---|
| Sample < 20 | 4,000 (52%) | 4,200 (55%) | 4,200 (55%) | 7,600 (54%) |
| % estimate < 20 | 20% | 21% | 22% | 26% |
| % variance < 20 | 99% | 95% | 97% | 95% |

The aim of the study is to investigate the impact of removing the small businesses in terms of the changes to bias and variance. As can be seen from the above numbers, the bias introduced to level estimates by removing the small businesses would be quite large. While the bias is not that large for movement estimates, it is still of concern, and there is also a need to adjust for the bias introduced to level estimates. As the bias is so large, the focus of the investigation was to investigate various strategies for adjusting for the bias introduced and to look at the impact on the variance.

The data items investigated were Capital Expenditure, Stocks, Sales and Wages and Salaries.

There are several possible sources of data that could be used to make the adjustment for the bias that were considered. Each of these sources had an assumption underlying them. The possible sources of data identified were

- data collected annually as part of the QEWS,

- data collected from an existing annual collection such as the EAS,
- data from the medium sized businesses

The following sections discuss the use of these various options to estimate for the missing small businesses. In all cases, the same sample size was used, i.e. the sample gained from losing the small businesses was reallocated to the larger businesses.

## 2.1 The use of June quarter data from QEWS

In this investigation, the possibility of collecting the data from the small businesses only once per year was trialed. In the investigation, it was assumed that data would be collected for the June quarter each year only and carried forward for the other quarters. It is of course possible that other quarters could be used.

The assumption under this estimator is that the average data from the small businesses does not change much over the course of a year. In particular, it is assumed that the average data from the small businesses for the September, December and March quarters are not significantly different from data for the June quarter. There are two reasons why there could be differences between the quarters. One is the fact that the data might be seasonal and the second is that there could be some significant changes in the data over time.

The preliminary part of the investigation was to test this assumption. If the test had shown that the assumption had not held, the investigation would not have been continued. Using a t-test, it was tested whether there were any differences for all four variables and the two choices of cutoffs, 0–4 and 0–19 at the stratum level, that is at the state by broad industry level. From this it was discovered that there were no significant differences between the quarters for any of the variables or cutoffs across two years of data. The one exception was Capital Expenditure. For this reason, Capital Expenditure was not considered further.

Once the assumption had been tested, the next step was to compare the estimator with the standard estimator by means of bias and variance. The formula for the level estimator using the annual data from QEWS is given for the non-June quarters by

$$\hat{Y}_q = \sum_{h \in 20+} N_{h,q} \overline{x}_{h,q} + \sum_{h \in 0-19} N_{h,q} \overline{x}_{h,June}$$

where $h$ = stratum (state × broad industry) and it is assumed that the businesses with employment less than 19 are only surveyed once per year.

It can be seen, however that this formula does take into account changes in population size which are quite common in the smaller sized businesses.

In the results below, this estimator is referred to as Annual 5+ QEWS or Annual 20+ QEWS depending on whether the 0–4 units or the 0–19 units were excluded.

## 2.2 The use of annual data from EAS

This investigation is similar to the one above, but instead of using data collected once a year from the QEWS survey, the investigation is into using a source of annual data already available, the Economic Activity Survey (EAS). The advantages of this are decreased cost and decreased respondent load as no data is collected from the small businesses, and a more even workload as the survey area would not experience an increase in work from the quarter in which small businesses are in sample.

As the data collected in EAS is for a full year, there needs to be some assumption about how that data is distributed between the four quarters. Based on the information from the above analysis, it was decided originally to assume that each quarter was equal and therefore divide the annual data by four. Therefore there are two assumptions in the use of this data. One is that the data collected from businesses in the EAS is actually similar to the data collected in the quarterly survey. The other is that the estimate for each quarter is about the same. One problem with this method is that due to the length of time required to process the EAS, the data from the relevant year is not available in time. Therefore, initially the previous year would need to be used. This could of course be updated in the future with the correct year.

To test these assumptions, similar hypothesis tests were conducted. This time it was discovered that there were significant differences between the averages from EAS and the averages from the other surveys for most variables and cutoffs. There was no pattern in which ones were significant and which ones weren't. There didn't however appear to be any significant difference between the quarters. The second part of the investigation was still continued with, in order to determine whether there were some variables for which this estimator would work well.

The formula for this estimator is given by

$$\hat{Y}_q = \sum_{h \in 20+} N_{h,q} \bar{x}_{h,q} + \sum_{h \in 0-19} N_{h,q} \frac{\bar{x}_{h,ann}}{4}$$

In the results below, this estimator is referred to as 5+ Annual (EAS) or 20+ Annual (EAS) depending on whether the 0–4 units or the 0–19 units were excluded. The 5+ Annual (EAS) estimator, however, was not calculated due to a lack of time.

## 2.3 The use of data from the medium sized businesses

One possible use of this data was as a direct substitute for the small businesses. However, as the data from the small businesses was significantly different from the medium sized businesses, this was not considered further. If there had been a significant difference between the quarters for small businesses, another possible use was to take the estimates from either of the above two annual sources and move these forward each quarter using the percentage change in the medium sized businesses. This estimator was intended to pick up any seasonal patterns or any movements in the data and assumed that the movement in the small businesses was the same as the movement in the medium sized businesses. The following table shows the average quarterly movement for both the 0–4 and the 5–19 businesses at the Australia level for all 4 variables.

**Table 2. Average quarterly movement for businesses sized 0–4 and 5–19 at the Australia level for all variables**

| | *CAPEX ($m)* | *STX ($m)* | *Sales ($m)* | *SEE ($m)* |
|---|---|---|---|---|
| Movement 0–4 | 3,100 | –2,590 | 2,700 | 2,060 |
| Movement 5–19 | –4,600 | –3,970 | –1,740 | 3,590 |

From this it can be seen that the pattern of movement in the medium-sized businesses was significantly different from the movement in the small businesses and therefore was not considered further.

## 2.4 The use of seasonal factors

As mentioned above, the difficulty with only using annual data for the small businesses is the loss in detail about quarterly movements for these units, particularly for a seasonal data item. As already stated, there did not appear to be any significant difference between the quarters suggesting that there was no seasonality. The idea of moving the annual sources forward using some kind of seasonal factor was considered, but rejected because of the difficulty of calculating a seasonal factor for the small businesses because they are so variable and also because of the difficulty in keeping this up to date when data from these units is no longer calculated.

# 3. RESULTS

The main assessment of the estimators was to compare the estimates, variances and relative mean square errors under all the different scenarios. The aim of this was to see whether the gain in variance was outweighed by the bias introduced or vice versa. These were calculated for at least eight quarters for both level and movement estimates in order to see how the methods worked over time.

The bias in a new estimator is calculated by taking the difference between the new estimator and the estimator including all units. There has been some discussion about whether or not this is the correct thing to do. This is because it is believed that the data obtained from small businesses has a large non-sampling error associated with it and therefore the estimator including all businesses is inherently biased anyway. This non-sampling error has not been included in the calculations. In order to quantify such a bias, it would be necessary to conduct an extensive empirical study of businesses.

It should be noted that this bias is still present in the estimator where data is used from the June quarter from QEWS, but is likely to be less in the estimator where annual data is used from EAS. This is because small businesses would generally have reconciled their annual accounts for taxation purposes, but would not do so for quarterly accounts.

## 3.1 Level estimates

The following table shows the average RMSEs, RSEs and Relative biases for each of the level estimators of total Australian Stocks across all quarters used. The results for the other variables are in Appendix A.

**Table 3. Summary measures for level estimates for total Australian Stocks (across 8 quarters)**

| | *True* | *Annual 5+ QEWS* | *5+* | *Annual 20+ QEWS* | *20+* | *Annual 20+ EAS* |
|---|---|---|---|---|---|---|
| Mean RMSE% | 2.39 | 1.61 | 12.18 | 1.98 | 29.90 | 2.04 |
| Mean RSE% | 2.39 | 1.49 | 0.89 | 1.88 | 0.37 | 1.50 |
| Mean Relative Bias % | 0.00 | –0.39 | 11.93 | –0.09 | 22.93 | 0.80 |
| Range of Relative Bias % | 0.00 | –1.17, 0.12 | 8.85, 12.08 | –1.36, 0.66 | 20.8, 25.47 | –0.71, 2.63 |

The range of relative bias in the table is the range given across estimates at industry division level.

The two estimators labelled 5+ and 20+ are the estimators calculated by just excluding the businesses with 0–4 and 0–19 employment respectively. The formula for the 5+ estimator is given by

$$\hat{Y}_q = \sum_{h \in 5+} N_{h,q} \bar{x}_{h,q}$$

The estimator for 20+ is similar.

As mentioned earlier, these estimators do not perform very well in that the relative bias for the 5+ estimators are over 10% and over 20% for the 20+ estimators. These large biases also led to large RMSEs.

The two QEWS annual estimators appear to perform the best, with the Annual 5+ QEWS estimator producing lower RMSEs than the unbiased estimator for both Wages and Salaries and Stocks. This is to be expected, considering that the estimator using EAS failed the hypothesis tests.

Appendix B contains graphs for the level estimates and RMSEs for the different techniques for all variables over time. The estimates omitting the small businesses and not adjusting for them have not been included because the biases were unacceptably large.

From these graphs it can be seen that the level estimates under the different approaches are quite similar to the unbiased estimate. It can also be seen that the 5+ Annual QEWS estimator gives a lower RMSE than the unbiased estimate for all quarters. For Stocks, the 20+ Annual QEWS also performs better, but for Wages and Salaries, this is more erratic, due mainly to the increasing bias. The 20+ Annual EAS estimator performs well for the second 4 quarters, i.e. a different EAS survey. This indicates that perhaps this approach is not very robust because different years of EAS can produce vastly different results. This may be partly due to the fact that no unit level edits in EAS are based on changes in data over time.

## 3.2 Movement estimates

Table 4 contains the average RMSEs, RSEs and Relative biases for each of the estimators of movement in total Australian Stocks at a point in time across all quarters used. The results for the other variables are in Appendix C.

**Table 4. Summary measures for estimators of movement in total Australian Stocks (across 8 quarters)**

| | *True* | *Annual 5+ QEWS* | *5+* | *Annual 20+ QEWS* | *20+* | *Annual 20+ EAS* |
|---|---|---|---|---|---|---|
| Mean RMSE% | 0.51 | 0.63 | 0.62 | 0.72 | 2.21 | 1.34 |
| Mean RSE% | 0.51 | 0.25 | 0.25 | 0.22 | 0.19 | 0.91 |
| Mean Relative Bias % | 0.00 | –0.13 | –0.07 | 0.04 | 0.09 | –0.47 |
| Range of Relative Bias % | 0.00 | –1.06, 1.17 | –1.92, 1.02 | –1.41, 3.71 | –3.1, 3.16 | –2.47, 0.80 |

From this table it can be seen that there are no estimators which have a lower RMSE than the true estimate. However the RMSEs produced by the estimators excluding the 0–4 businesses are closest to the RMSE for the true estimate for Stocks. It can be seen that even the two estimators which make no adjustment for the excluded units have quite low RMSEs. It is not essential that the same estimator be used for the level and movement estimates, however if they are not used, then the difference between two level estimates will not equal the movement estimate. However, if the data is collected, it seems logical to use it in both level and movement estimates.

Appendix D contains graphs of the movement estimates and RMSEs for the different estimators for the variables over time.

These graphs support the information found in the table. For Stocks, the 5+ estimator has a lower RMSE for most quarters, however there are a few quarters where the estimator is wildly different, leading to a large bias and therefore a large RMSE. For Wages and Salaries, the 5+ estimator performs better on a few quarters, however on other quarters it is quite different.

# APPENDIX A

**Table A.1 Summary measures for level estimates for total Australian Wages and Salaries (across 8 quarters)**

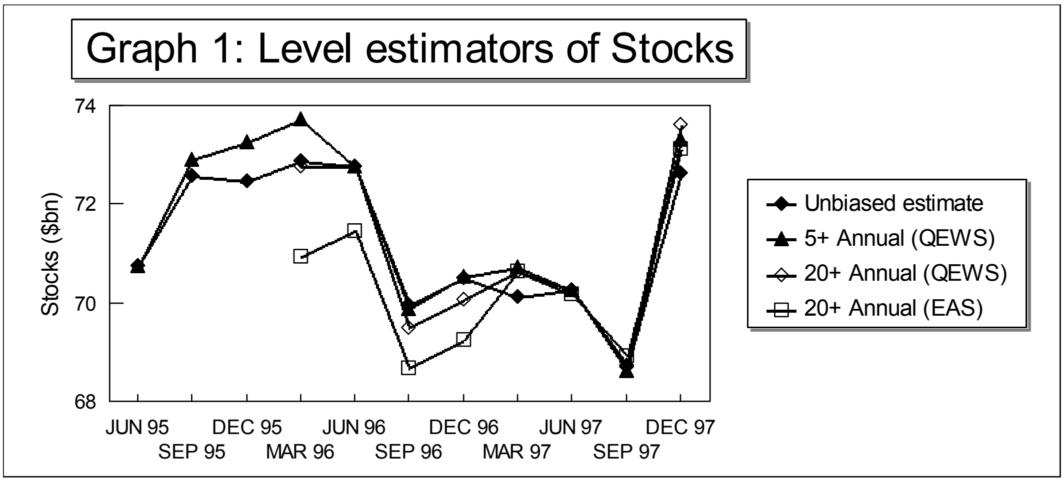
| | *True* | *Annual 5+* | *5+* | *Annual 20+* | *20+* | *20+ EAS 1* |
|---|---|---|---|---|---|---|
| Mean RMSE% | 2.20 | 1.08 | 16.98 | 2.64 | 53.80 | 3.25 |
| Mean RSE% | 2.20 | 0.96 | 0.89 | 2.14 | 0.45 | 1.06 |
| Mean Relative Bias % | 0.00 | –0.20 | 14.20 | –1.15 | 3.50 | 1.31 |
| Range of Relative Bias % | 0.00 | –0.85, 0.72 | 12.40, 15.67 | –3.74, 0.22 | 34.20, 36.46 | –3.24, 6.98 |

**Table A.2 Summary Measures for level estimates for total Australian Sales (across 8 quarters)**

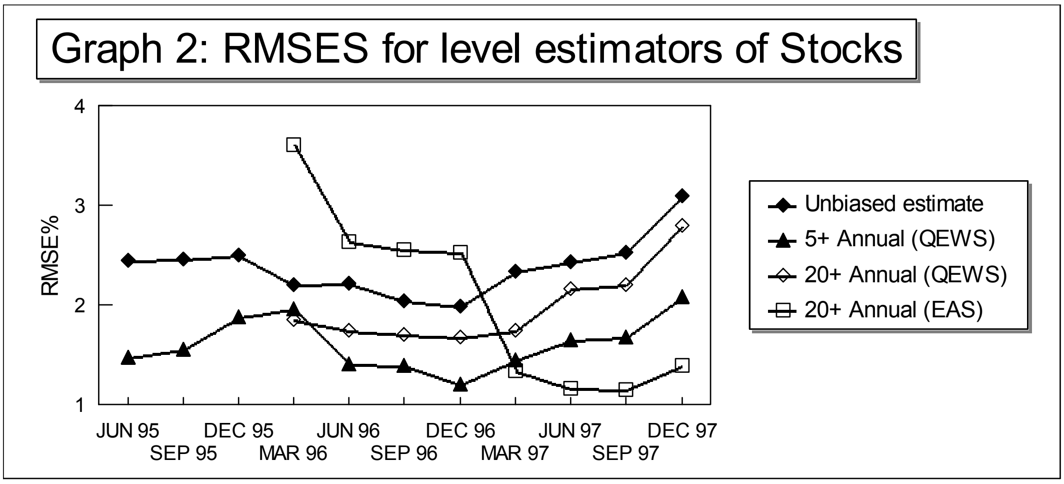
| | *True* | *Annual 5+* | *5+* | *Annual 20+* | *20+* | *20+ EAS 1* |
|---|---|---|---|---|---|---|
| Mean RMSE% | 2.90 | 3.28 | 10.72 | 3.05 | 30.22 | 2.57 |
| Mean RSE% | 2.90 | 3.22 | 1.00 | 2.83 | 0.43 | 1.77 |
| Mean Relative Bias % | 0.00 | 0.03 | 9.63 | 0.34 | 23.12 | –1.29 |
| Range of Relative Bias % | 0.00 | –0.83, 1.26 | 8.44, 11.28 | –1.47, 2.03 | 20.96, 26.23 | –3.56, 0.79 |

# APPENDIX B

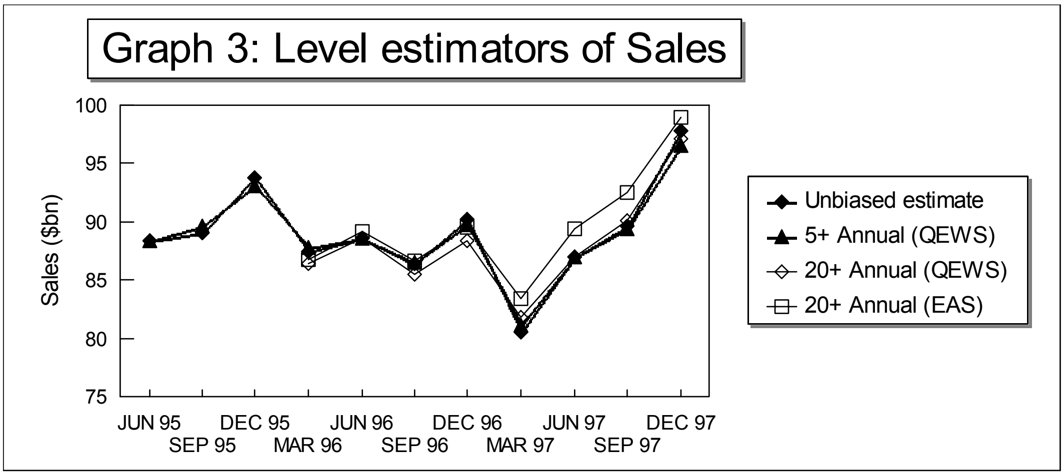
**Graph B.1**



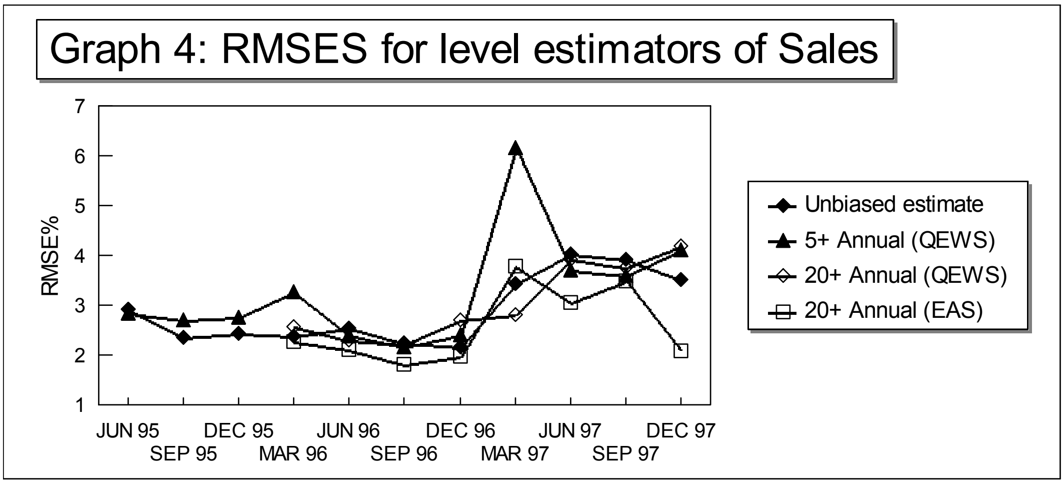

**Graph B.2**



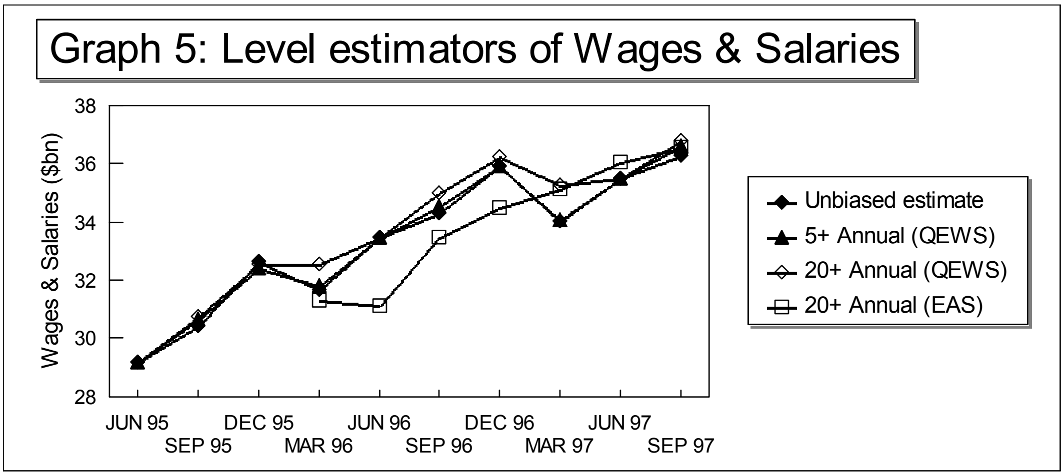

**Graph B.3**



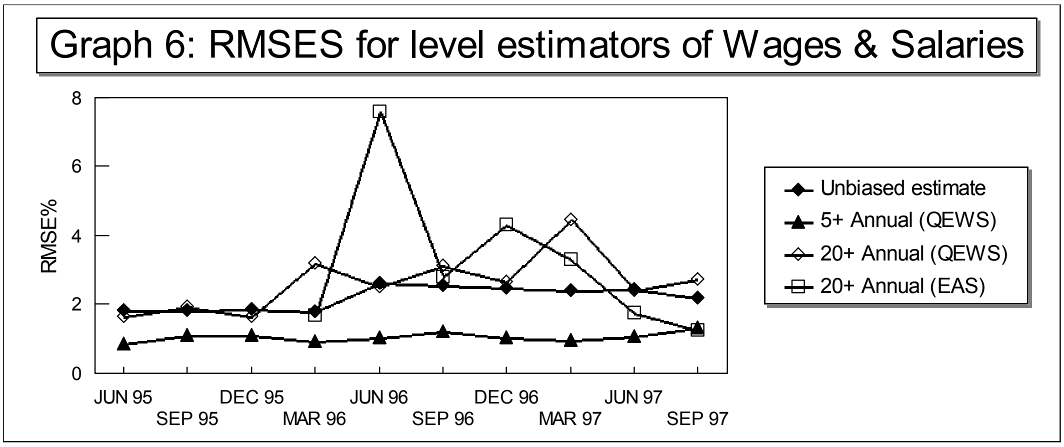
**Graph B.4**




**Graph B.5**




**Graph B.6**



# APPENDIX C

**Table C.1 Summary measures for estimators of movement in total Australian Wages and Salaries (across 8 quarters)**

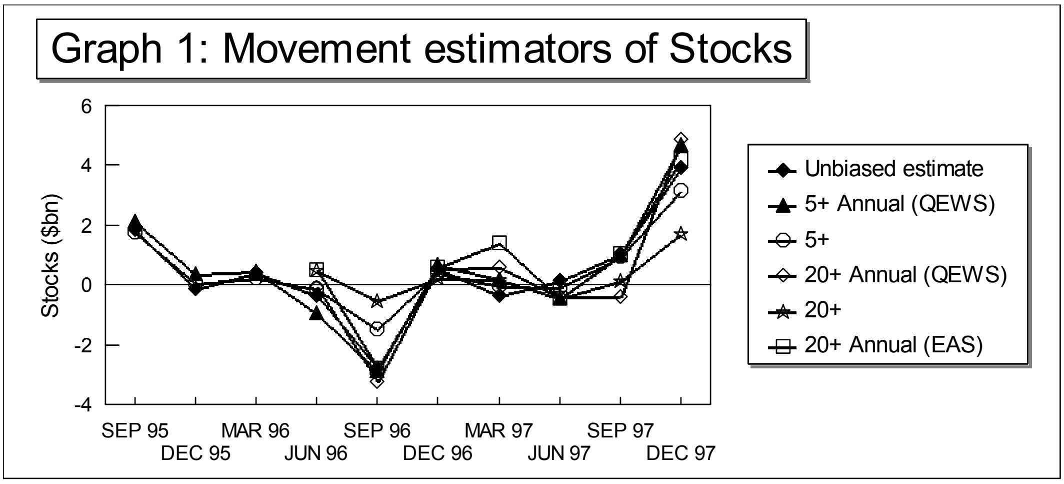
| | *True* | *Annual 5+* | *5+* | *Annual 20+* | *20+* | *20+ EAS 1* |
|---|---|---|---|---|---|---|
| Mean RMSE% | 0.45 | 1.06 | 1.06 | 2.80 | 2.13 | 2.81 |
| Mean RSE% | 0.45 | 0.32 | 0.38 | 0.29 | 0.39 | 0.43 |
| Mean Relative Bias % | N/A | –0.28 | 0.41 | –0.35 | 0.79 | –1.56 |
| Range of Relative Bias % | N/A | –1.70, 1.22 | –1.17, 2.13 | –3.97, 1.22 | –5.40, 6.22 | –8.46, 1.93 |

**Table C.2 Summary measures for estimators of movement in total Australian Sales (across 8 quarters)**

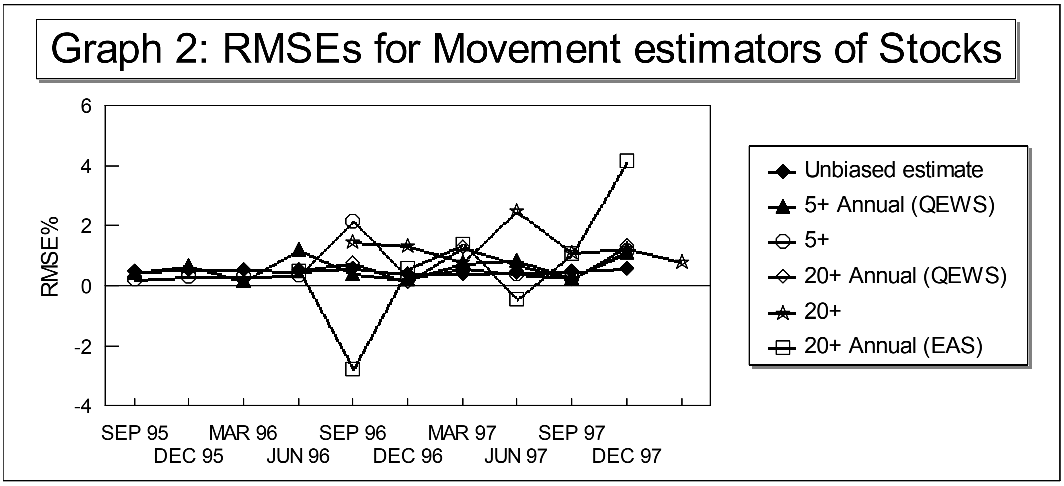
| | *True* | *Annual 5+* | *5+* | *Annual 20+* | *20+* | *20+ EAS 1* |
|---|---|---|---|---|---|---|
| Mean RMSE% | 0.75 | 0.98 | 1.33 | 1.44 | 2.90 | 1.79 |
| Mean RSE% | 0.75 | 0.37 | 0.33 | 0.29 | 0.30 | 0.87 |
| Mean Relative Bias % | N/A | 0.10 | 0.22 | –0.19 | 0.12 | –0.29 |
| Range of Relative Bias % | N/A | –1.60, 1.35 | –1.87, 2.83 | –3.34, 1.47 | –2.79, 5.68 | –3.97, 1.94 |

# APPENDIX D

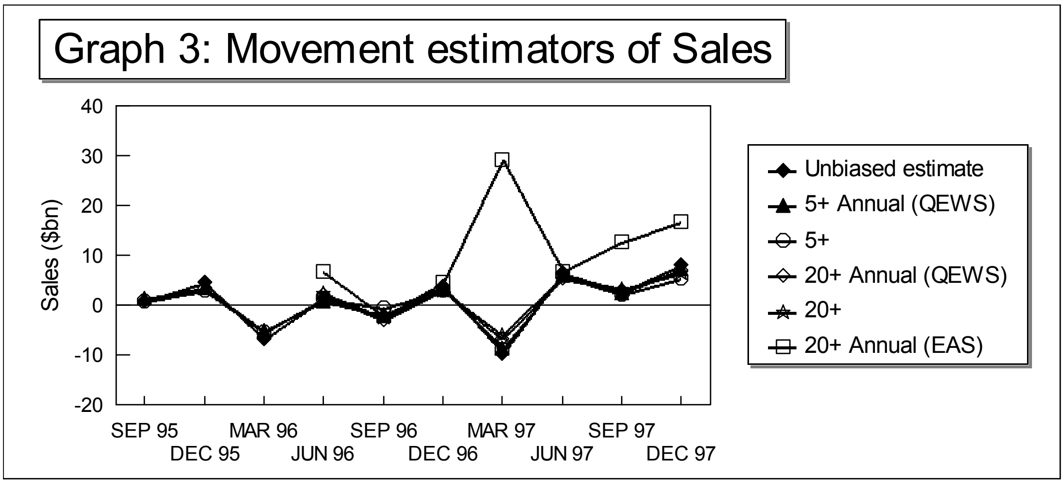
**Graph D.1**



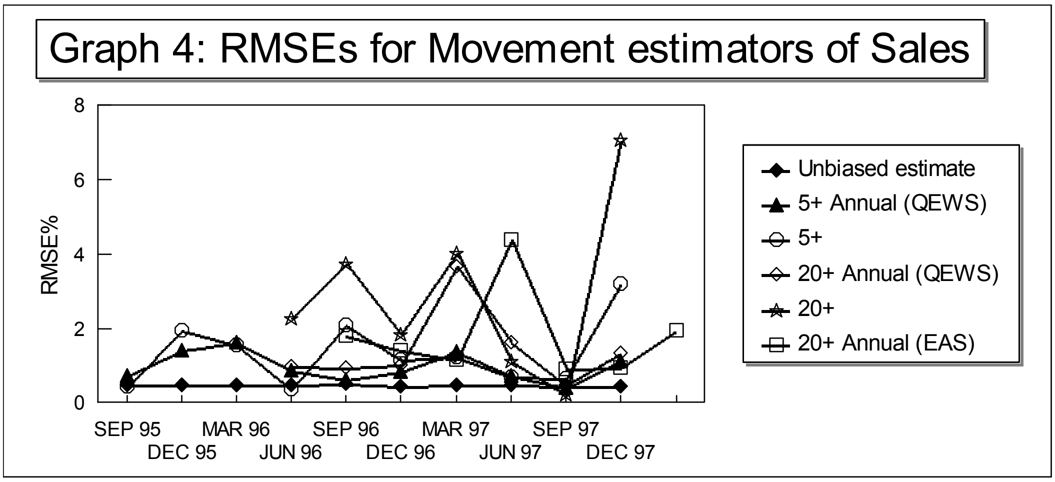

**Graph D.2**



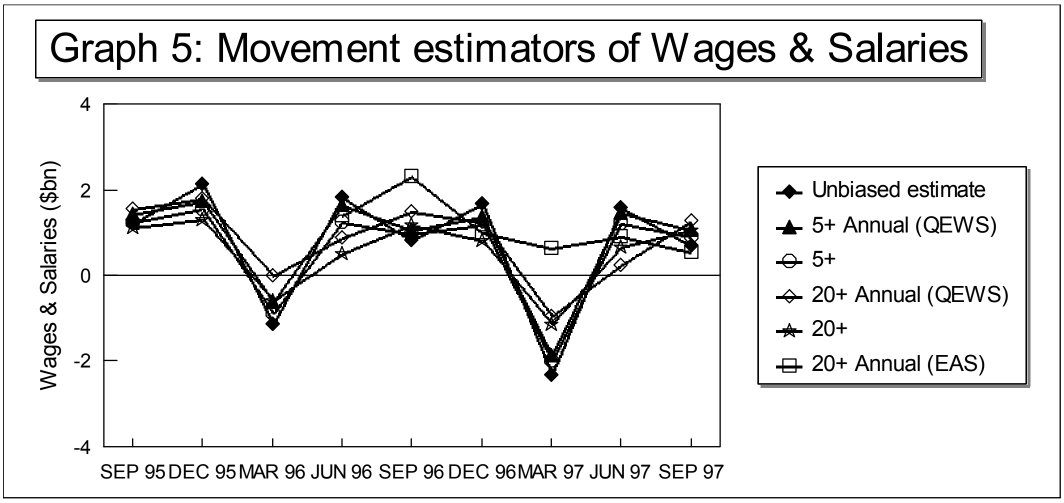

**Graph D.3**



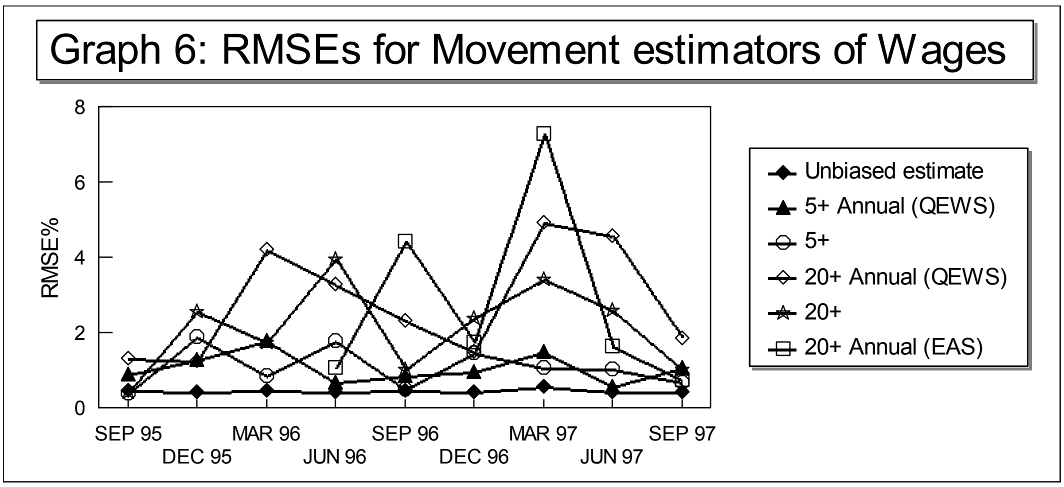
**Graph D.4**




**Graph D.5**




**Graph D.6**



# PART II: DIRECT MOVEMENT ESTIMATION IN THE QUARTERLY ECONOMY WIDE SURVEY

Edward Szoldra and Louise Gates
Statistical Services Branch

## 1. INTRODUCTION

This document presents the results of an evaluation of an alternative method of estimating quarterly movements in regular surveys of businesses. The method relies on the use of only the common sample between quarters, supplemented by the inclusion of frame births and frame deaths between the two periods.

The method, termed 'direct movement estimator' (DIME), is a form of composite estimator of movement. It differs from composite estimation, however, in that it does not attempt to possess optimal weights. It shares with the composite estimator the feature that the estimate of movement is not the difference in the two successive estimates of full-sample level.

The direct-movement method is proposed as part of a wider 'package' of estimation for quarterly surveys. This paper does not propose to address all the aspects of this 'package'. This paper confines itself to investigating two major aspects of the direct movement estimator. Firstly, it investigates how well the direct movement estimator performs as an estimator of quarterly movement. Secondly, it also looks at an estimator of level which is derived from the direct-movement estimator. This new estimator of level, termed the 'direct movement estimator of level (DIMEL)', is simply built by adding successive DIMEs to an estimate of level calculated at some base period. How this base-period estimate of level is calculated is not addressed in this paper. The DIMEL has the intuitively attractive feature of having the DIME being the difference of successive estimates of level.

# 2. BACKGROUND

## 2.1 Current methods

The current methods of estimating movements in ABS business surveys involve estimating levels at two successive time periods, and then subtracting these levels to estimate the movement. This is a convenient method, and also lines up with what intuition would suggest. If the population (represented by the frame) at each of the two time periods is stable (that is, the same units exist at the two time periods), and the sample is also stable, then this method is also an efficient method of estimating movements.

When populations or samples (or both) change, then the situation arises that the composition of the sample units between the two time periods may differ to a degree. Sample rotation (done to help minimise respondent fatigue) brings new units into sample; these units will not have completed a form at the previous time period. Therefore, their unit-level movement is unknown. However, for common units between the time periods, individual movements are known.

This knowledge of which units are common can be used to drastically improve the standard error of the estimate of movement between the two time periods. Units in sampled strata tend to have very similar stocks between successive quarters; in fact, the correlation is about 95%. The DIME uses this correlation to good effect in reducing the variability of the estimate of movement.

In addition, several questions arise:

(i) Are the units dropped similar to the other units in survey, or do they have stocking behaviours which differ to the common units?

(ii) Is the cost to standard error of level estimates outweighed by benefits to standard error of movement estimators?

(iii) What is the impact on smoothed estimators of level (such as trended series) of dropping the non-common units?

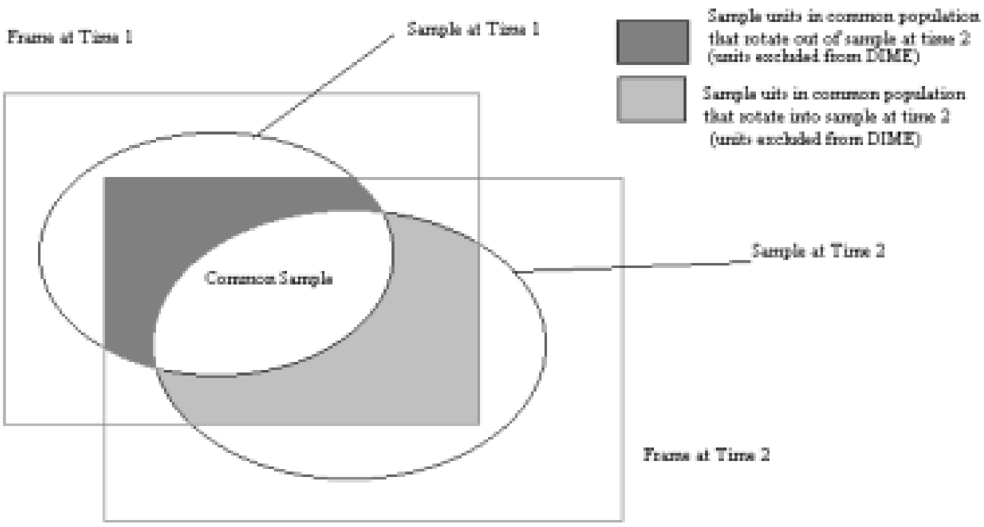
(iv) Is the added complexity in estimation justified?

## 2.2 Direct movements – definition

The DIME (Direct Movement Estimator) is a type of composite estimator. That is, the sample selection information can be used to adjust the weights of the units at the two time points, in order to reduce the variance of the resulting movement estimator. There is a substantial cost, though; the estimator of movement is no longer the difference in levels. This might lead to user confusion and scepticism that ABS movement estimates are in some way counter-intuitive.

A pictorial representation of the direct-movement situation is given in Diagram 1. The units in the shaded portions correspond to the units that the DIME excludes. These units are the sampled units that are rotated out of sample each quarter in order to give the respondents a reprieve from the task of reporting. Each quarter, the ABS aims to rotate out one-twelth of the sample. However, units that rotate out of sample due to being a frame birth or a frame death are included in the DIME.

**Diagram 1: Components of the Population and Sample excluded by the direct movement estimator**




The direct-movement estimator (DIME) can be written as:

$$m''_{21} = x'_{21c} + x'_{2b} - x'_{12c} - x'_{1d}$$

where

$m''_{21}$ = direct-movement estimator between time 2 and 1

$x'_{2b}$ = estimate of level of births on frame at time 2

$= \sum_{i \in b} N_b / n_b \, x_{2i}$ ($N_b$ = births on frame; $n_b$ = births in sample)

$x'_{21c}$ = estimate of level of common units at time 2 (that are common to time 1)

$= \sum_{i \in c} N_c / n_c \, x_{2i}$ ($N_c$ = common on frame between times 1 and 2; $n_c$ = common in sample between times 1 and 2)

$x'_{12c}$ = estimate of level of common units at time 1 (that are common to time 2)

$= \sum_{i \in c} N_c / n_c \, x_{1i}$

$x'_{1d}$ = estimate of level of deaths on frame at time 1

$= \sum_{i \in d} N_d / n_d \, x_{1i}$ ($N_d$ = deaths on frame; $n_d$ = deaths in sample)

Section A.1 in the Appendix gives details regarding the calculation of the DIME.

The DIMEL is defined as follows.

$$x''_k = x'_1 + m'_{21} + m'_{32} + \ldots + m'_{k,k-1}$$

The initial level, $x'_1$ is sourced from an ABS annual series. This paper does not address this aspect of the DIMEL. This paper assumes that $x'_1$ is an estimate of level at time 1 from the same survey that the DIME's are calculated for, but uses the full-sample (i.e. it does not omit any units).

The DIMEL is also defined in Section A.1 of the Appendix. It is presented in its purest form. In practice, as each new DIME is available and added, some revision will be necessary due to late respondents at a previous time period now becoming available for use in estimation. Thus, if the DIMEL is allowed to extend to up to 8 quarters, say, it may be necessary to re-calculate it if more complete data arrives in respect of quarter three in the chain. This should not present a problem; current estimates of movement are revised as new data becomes available.

## 2.3 Surveys used in the study

The three surveys considered to date have been:

- Stocks and Sales Survey,
- Survey of Private New Capital Expenditure,
- Survey of Company Profits.

Each survey estimates for a totally different type of variable.

The Stocks variable is generally very well behaved between quarters; companies generally cannot vary stock levels too much. That is, the between-quarter correlation of the stocks variable is high. This makes Stocks an ideal DIME candidate.

The Capital Expenditure variable is much more volatile. A sampled (and therefore small to medium) company will generally not have much (or any) capital expenditure in any one quarter. In fact, about 80% of small to medium companies generally have no capital expenditure at all in a given quarter. However, when capital expenditure does arise, it is (by its nature) usually quite large, and sometimes, extremely large. This makes between-quarter correlation very small and, most likely, lead to inefficient DIMEs.

The Profits survey has a small sampled sector (only companies with benchmark employment greater than 20 are selected, and of these, only those with employment greater than 30 are used in publications). Company profits tend to be reasonably stable at the quarterly level. Also, profits does not relate particularly well to the stratifying measure, employment, and hence a large within-stratum variability occurs. The between-quarter change in profit is generally not large; these factors should combine to suggest profits as a reasonable candidate for the direct movement estimator.

## 2.4 Theoretical versus empirical study

The purpose of having a theoretical and an empirical study was to be able to make both general and specific comments about the impact of the DIME and the DIMEL. The theoretical study below attempts to demonstrate how these estimators behave with respect to time and also with respect to the most vital data structure; the between-quarter correlation structures.

The empirical study compares the current (full-sample) and proposed (DIMEL) estimators. It is difficult to compare these numbers without the theoretical study; the two estimators produce different results. This in itself is not interesting. However, they are re-assuring in that they do not show any clearly unreasonable behaviour.

## 2.5 Revisions

Current full-sample estimators of level in ABS quarterly business surveys have the property of having a stable standard error. That is, each quarter, a new sample is drawn (which largely overlaps the previous quarter's sample) and this sample used to produce estimators of level. These estimators of level have the same standard error each quarter. The DIMEL does not have this property. The DIMEL is a linear function of the DIMEs and the base-time level estimate (see Section A.1 of the Appendix). Thus, as the DIMEL chain increases, the estimator of level (though still unbiased) assumes an increasing standard error. If this is left to go unchecked, the DIMEL will eventually possess an unacceptably large standard error. This will require the series to be revised.

The method that is proposed to revise the DIMEL quarterly series is to regularly benchmark it to another more accurate series. The ABS also publishes an annual series (known as the Economic Activity Survey), which collects annual data that is similar to the quarterly series in some respects. This annual series possesses a larger sample size than the ABS quarterly series do, and hence this EAS series will be a good candidate for a 'benchmarking series'.

The ABS generally considers the annual series to be more accurate than its quarterly series in other significant ways. Annual accounts are maintained for taxation purposes within all companies, and hence there is a greater chance of obtaining more accurate data from annual sources. In addition, the ABS subjects the annual series to a process known as the Input Output Table reconciliations. This process is beyond the scope of this paper, but essentially it is a process which allows economic series to be adjusted by confronting them against each other (and supplementing with expert knowledge).

Section A.2 of the Appendix shows that the expected size of the revision at time $k$ would be at least 80% (which is the quantity $\sqrt{2/\pi}$ in Section A.2 of the Appendix) of the size of the standard error at that time. That is, the size of the revision is a simple function of the standard error if the DIMEL at time $k$. This assumes that the estimator is unbiased.

Thus, if the DIMEL were to be compared with the full-sample estimator of level at time $k$, it is expected to be considerably different. In fact, the probability that it is within 10% of the full-sample estimator after a couple of years is virtually nil in the surveys studied.

# 3. THEORETICAL EVALUATION

## 3.1 Overview

The study investigated the following issues:

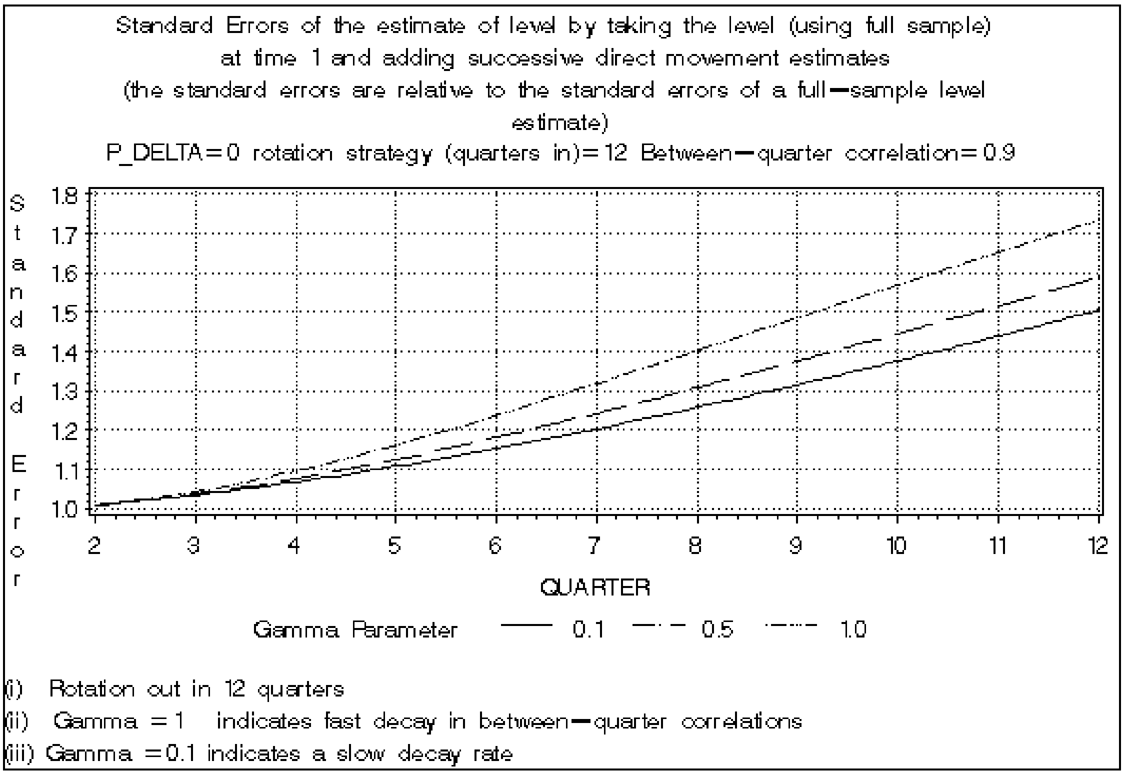
- How quickly does the DIMEL degrade in standard error over time for stable populations ?
- What is the impact of the frame births and deaths on standard error, over and above the variability already present in the matched-sample units
- Is there a condition under which the DIME is less efficient than the full-sample estimator of movement, and can a test be derived to see if the surveys tested are suitable for the DIME ?
- Is it possible to estimate the expected size of revision of the DIMEL-chain ?

### *3.1.1 How quickly does the standard error of the DIMEL increase over time?*

The study used the theory in the Appendix to estimate the standard error of the DIMEL over time. Essentially, the DIMEL will degrade with time more quickly if the correlation between quarters in the target variable is small, which is to be expected. As the correlation increases, the DIMEL is much tighter about its expected value after any given time.

Firstly, assume that the DIMEL does not include any frame births or deaths. That is, all units are common in both the population and the sample. The rate of decay of the DIMEL is shown as Graphs 3.1.1(a) and 3.1.1 (b).

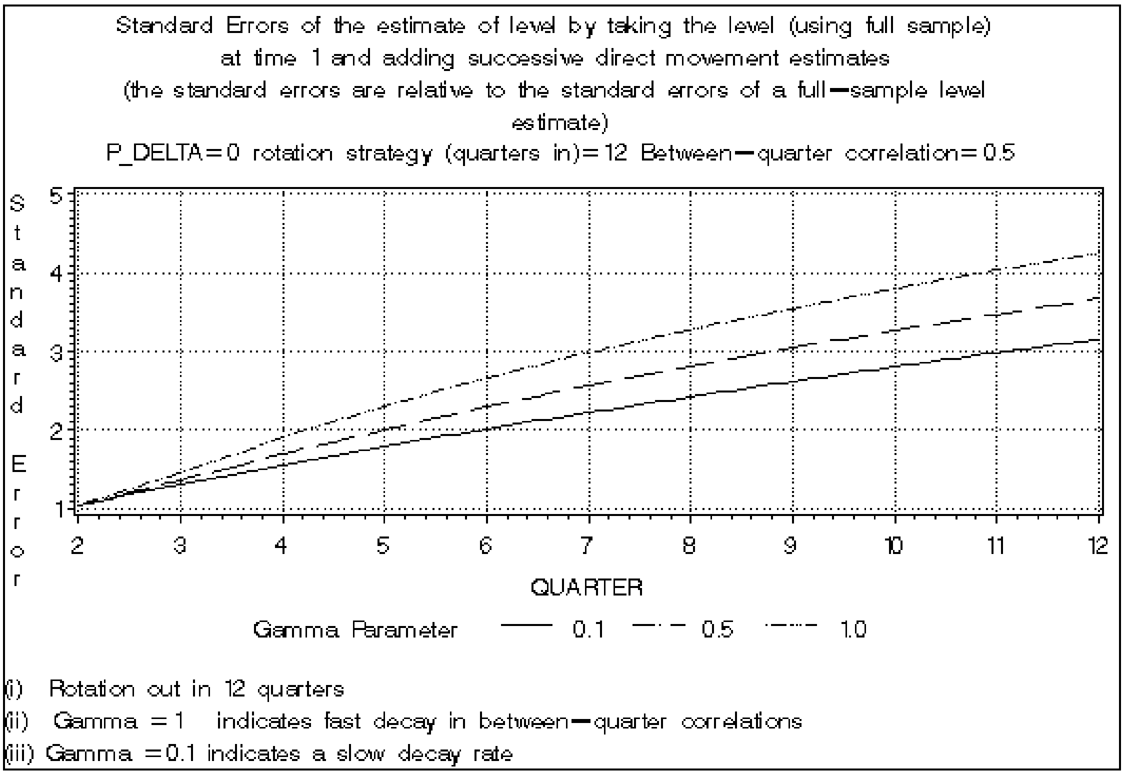
**Graph 3.1.1(a) Rate of decay of DIMEL**
**(increase in standard error over time with a between-quarter population correlation of 0.9)**




FOOTNOTE: P_DELTA is the proportion of the Frame that is occupied by births or deaths.

This Graph demonstrates that after a year (4 quarters), the DIMEL has about 10% more standard error than the usual full-sample estimator. This is a best-case scenario; it assumes a between-quarter correlation of 0.9, and no frame changes. A more pessimistic scenario (correlation of 0.5) gives a more alarming picture.

**Graph 3.1.1(b) Rate of decay of DIMEL**
**(increase in standard error over time with a between-quarter population correlation of 0.5)**




FOOTNOTE: P_DELTA is the proportion of the Frame that is occupied by births or deaths.

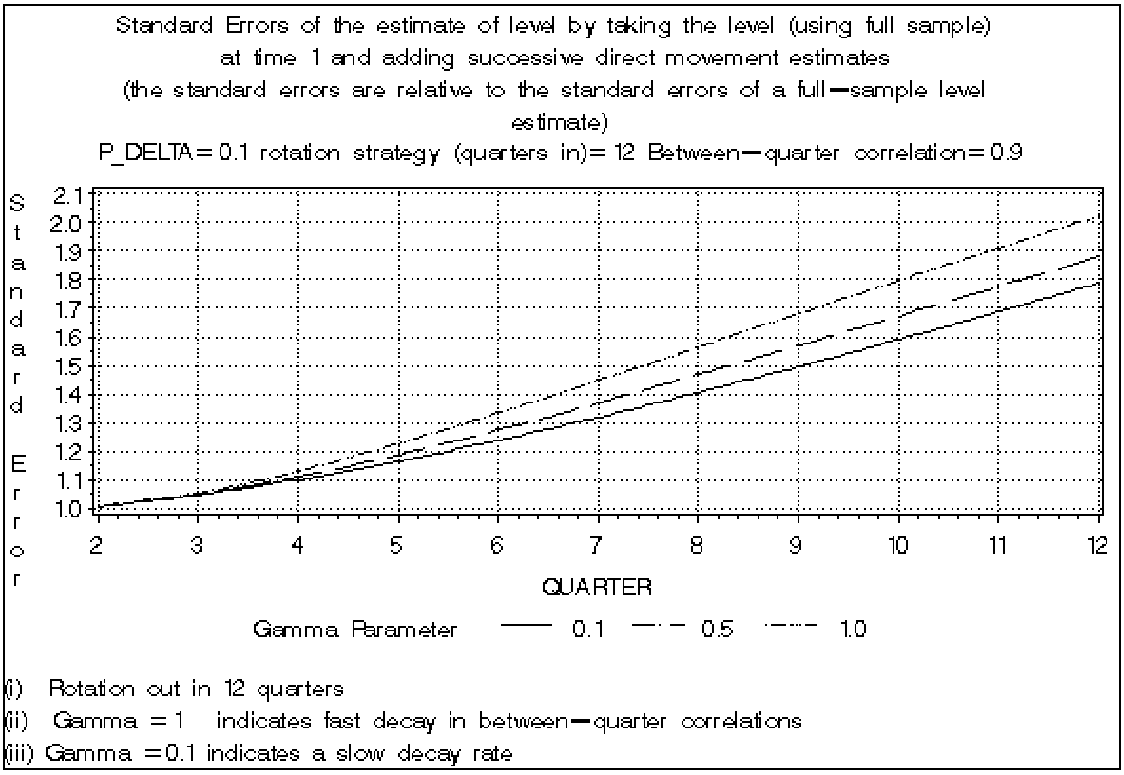
The standard errors are about 70% higher; the longer-term prospects preclude the use of the DIMEL beyond 4 quarters. Thus, any strata with this sort of correlation structure will perform very poorly under the DIMEL (in terms of the level; the movements will retain their benefits).

### *3.1.2 What is the impact of the frame births and deaths on standard error, over and above the variability already present in the matched-sample units?*

The inclusion of frame changes should be expected to raised DIMEL variability a fair amount. The current rotation strategy used in Business Surveys (to rotate units out in 12 quarters) usually exceeds the rate of frame changes. There have been cases where frame changes have exceeded the sample rotation rate (for example, at re-designs and when there is feedback from economic censuses). However, frame deaths can be expected to account for no more than 1% or 2% of frame size, and frame births perhaps slightly more (2% to 3%).

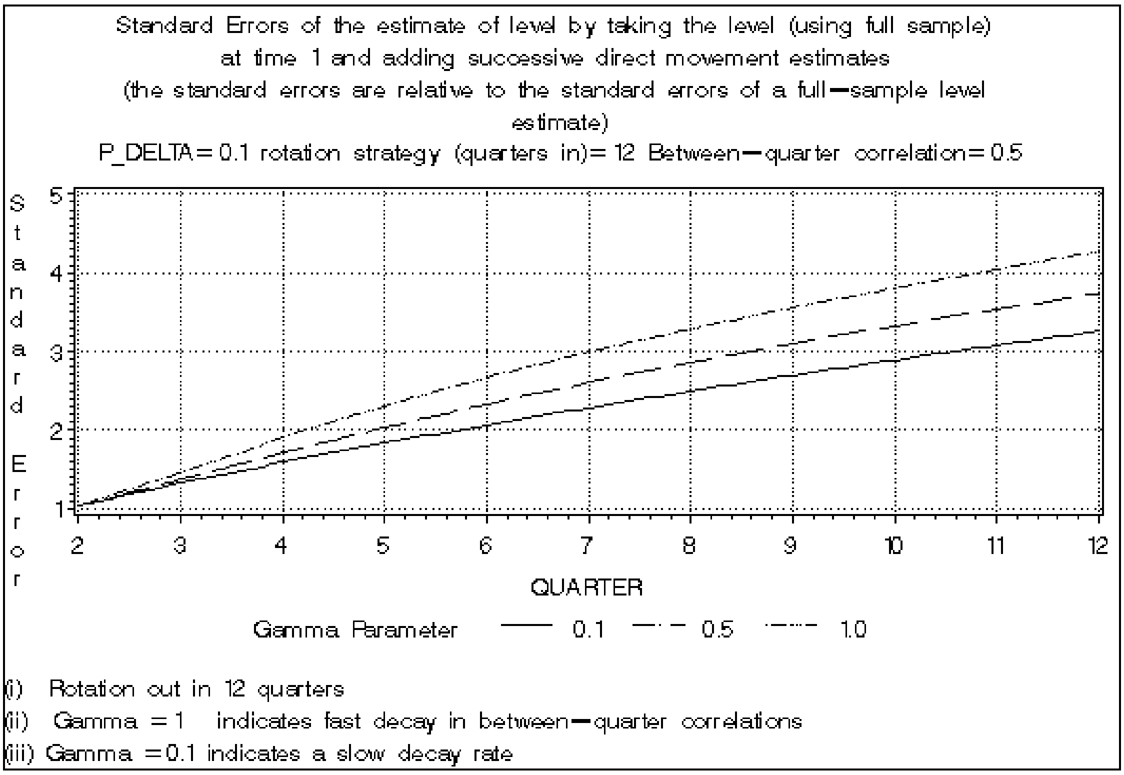
The graphs below demonstrate the impact of including the frame changes.

**Graph 3.1.2(a) Rate of decay of DIMEL including frame births and deaths (increase in standard error over time with a between-quarter population correlation of 0.9)**




FOOTNOTE: P_DELTA is the proportion of the Frame that is occupied by births or deaths.

**Graph 3.1.2(b) Rate of decay of DIMEL including frame births and deaths (increase in standard error over time with a between-quarter population correlation of 0.9)**




FOOTNOTE: P_DELTA is the proportion of the Frame that is occupied by births or deaths.

The impact of including frame births and deaths is quite pronounced in situations where the correlation between quarters is high. In these cases, there is little between-quarter variability, and the contribution of only a small number of non-common units can strongly influence the long-term standard error of the DIMEL. A comparison of Graphs 3.1.1 (b) and 3.1.2 (b) shows that the two graphs are virtually identical. This indicates that when the between-quarter correlations are less than about 0.5, the extra variability of the non-common units is not influential in DIMEL precision.

In addition, Graphs 3.1.1(a) and 3.1.2(a) show a much slower increase in standard error than the Graphs 3.1.1(b) and 3.1.2(b). This demonstrates the importance of having high between-quarter correlation when making use of the DIMEL.

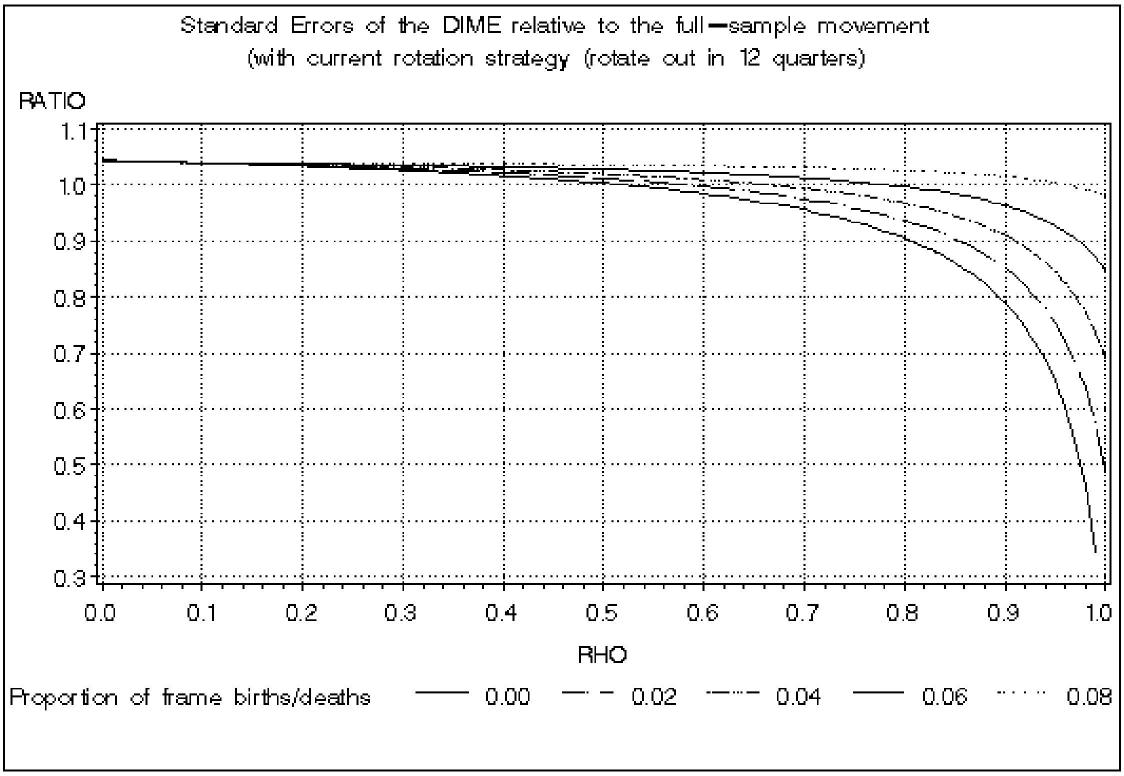
### *3.1.3 Is there a condition under which the DIME is less efficient than the full-sample estimator of movement, and can a test be derived to see if the surveys tested are suitable for the DIME?*

Section A.3 of the Appendix contains details of a test used to discern whether the DIME is of higher or lower standard error than the full-sample estimate. This test has been applied to the Business Surveys to determine if the DIME is appropriate. Essentially, the between-quarter population correlation needs to be of a certain magnitude (relative to sample rotation rates) before the DIME is of benefit.

*Standard Error of the DIME, and long-term correlation behaviour*

Section A.1 of the Appendix contains details of the standard errors of the DIME. Graph 3.1.4 shows the standard error of the DIME relative to the full-sample movement standard error. The graph shows that the DIME performs increasingly better as the correlation increases, and as the proportion of frame occupied by births/deaths increases. To illustrate, the Stocks Survey has between-quarter correlations of about 0.90 to 0.95; this would mean a gain in efficiency of between 5% to 30%. A common sample estimator would benefit considerably more, to the extent shown in graph 3.1.4 by the curve indicated by frame births/deaths being zero.

**Graph 3.1.4 Standard error of the DIME relative to the full-sample movement estimator**




FOOTNOTE: P_DELTA is the proportion of the Frame that is occupied by births or deaths.

### *3.1.4 Is it possible to estimate the expected size of revision of the DIMEL-chain?*

Section A.2 of the Appendix shows that the expected size of the revision can be shown to be a function of the sampling variance of the DIMEL. This is a theoretical evaluation only; the revisions after 4 quarters can be seen in section 4 below. Consideration has to be given to the issue of how to adjust for the increased standard error of the DIMEL (above that of the full-sample estimator); the DIMEL is an unbiased estimator of the underlying level, and (theoretically) any adjustment is simply to account for this increased variability.

In order to make an adjustment for this increased standard error, this paper has assumed the following 2 items:

(i) The average adjustment will be zero (that is, the DIMEL is unbiased and thus will on average be zero)

(ii) The average adjustment, not accounting for the sign of the adjustment (that is, the average absolute adjustment) is the adjustment that is of interest.

The second point is the important one. If the DIMEL needs to be adjusted, the sign of the adjustment is not important. It is the magnitude of the adjustment that is of interest, whether it be positive or negative. The salient result is that the average size of the adjustment in theory, will be about 80% of the standard error of the DIMEL at the time of adjustment. That is, if the DIMEL is adjusted each 4 quarters, then the average size of the adjustment will be about 80% of the standard error of the DIMEL at that point. Reference to the above tables will show the ratio of the DIMEL standard errors to the current full-sample standard errors; for example, for a correlation of 0.9 (as in Stocks), the DIMEL has a standard error about 15% larger than the current standard errors. Therefore, the average revision will be about 80% of 15%, or about 11% of the current standard errors. However, if we revise only every 2 years, the average revision will be about 30% of the current standard errors.

## 4. EMPIRICAL EVALUATION

The empirical evaluation below consists of the production of estimates of DIME and DIMEL estimates for the period March 1997–December 1997. These estimates were produced in order to address two questions:

(i) Are the direct movement estimators producing consistently higher or lower estimates than the full-sample estimates?

(ii) Is there anything of an obvious nature which would make the direct movement estimators unacceptable?

The tables below indicate that the direct movement is not consistently higher or lower than the full-sample estimator. Apart from a few examples, it is also not producing extreme results. The main difference between the two estimators is the exclusion of sample births and deaths that are not frame births and deaths. Their exclusion does not consistently affect the estimates (by industry).

*Notes on the tables*

For Stocks the item used was *total stocks*.

For Capital Expenditure the item used was *total capital expenditure*.

For Profits the item used was *gross operating profit*.

Note that in the following tables, the base quarter is March 1997. Thus in this quarter, the direct movement estimator of level and the full-sample movement are equal.

**STOCKS (sampled sector) (NR=number-raised (full-sample) DM=direct movement)**

| | *Mar 1997* | | *Jun 1997* | | *Sep 1997* | | *Dec 1997* | |
|---|---|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | 1,403 | 1,403 | 1,373 | 1,365 | 1,300 | 1,103 | 1,280 | 1,087 |
| Manufacturing | | | | | | | | |
| Food, beverage and tobacco | 1,206 | 1,206 | 1,274 | 1,251 | 1,238 | 1,401 | 1,339 | 1,483 |
| Textiles, clothing, footwear and leather | 856 | 856 | 816 | 807 | 939 | 919 | 904 | 890 |
| Wood and paper products | 526 | 526 | 607 | 553 | 631 | 526 | 662 | 522 |
| Printing, publishing and recorded media | 497 | 497 | 504 | 503 | 511 | 528 | 556 | 513 |
| Petroleum, coal, chemical and assoc. prods | 1,520 | 1,520 | 1,437 | 1,604 | 1,305 | 1,521 | 1,275 | 1,488 |
| Non-metallic mineral products | 454 | 454 | 485 | 439 | 518 | 434 | 528 | 399 |
| Metal products | 1,594 | 1,594 | 1,517 | 1,521 | 1,280 | 1,122 | 1,272 | 1,043 |
| Machinery and equipment | 2,831 | 2,831 | 2,850 | 2,773 | 2,191 | 2,600 | 2,015 | 2,516 |
| Other manufacturing | 458 | 458 | 531 | 499 | 565 | 462 | 588 | 471 |
| Total manufacturing | 9,941 | 9,941 | 10,020 | 9,951 | 9,179 | 9,512 | 9,138 | 9,325 |
| Wholesale Trade | 14,588 | 14,588 | 15,050 | 15,014 | 16,975 | 16,746 | 16,736 | 18,685 |
| Retail Trade | 10,362 | 10,362 | 11,014 | 10,560 | 10,851 | 10,587 | 10,682 | 10,567 |
| Other services | 414 | 414 | 507 | 407 | 494 | 335 | 548 | 346 |
| Industry Total | 36,708 | 36,708 | 37,964 | 37,297 | 38,799 | 38,283 | 38,384 | 40,010 |

**CAPITAL EXPENDITURE (NR=number-raised (full-sample) DM=direct movement)**

| | *Mar 1997* | | *Jun 1997* | | *Sep 1997* | | *Dec 1997* | |
|---|---|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | 734 | 734 | 579 | 1,312 | 572 | 977 | 680 | 1,004 |
| Manufacturing | | | | | | | | |
| Food, beverage and tobacco | 89 | 89 | 136 | 133 | 133 | 110 | 106 | 78 |
| Textiles, clothing, footwear and leather | 31 | 31 | 44 | 35 | 35 | 26 | 56 | 41 |
| Wood and paper products | 21 | 21 | 29 | 31 | 27 | 26 | 129 | 100 |
| Printing, publishing and recorded media | 65 | 65 | 139 | 124 | 73 | 77 | 116 | 119 |
| Petroleum, coal, chemical and assoc. prods | 88 | 88 | 107 | 114 | 107 | 131 | 84 | 108 |
| Non-metallic mineral products | 192 | 192 | 185 | 171 | 141 | 133 | 161 | 159 |
| Metal products | 59 | 59 | 87 | 83 | 66 | 56 | 143 | 89 |
| Machinery and equipment | 175 | 175 | 232 | 148 | 261 | 164 | 377 | 298 |
| Other manufacturing | 32 | 32 | 55 | 33 | 54 | 25 | 70 | 33 |
| Total manufacturing | 751 | 751 | 1,014 | 872 | 895 | 746 | 1,242 | 1,026 |
| Construction | 262 | 262 | 297 | 345 | 244 | 257 | 352 | 373 |
| Wholesale trade | 366 | 366 | 586 | 1,188 | 548 | 943 | 484 | 894 |
| Retail trade | 84 | 84 | 277 | 211 | 243 | 152 | 357 | 275 |
| Transport and storage | 353 | 353 | 368 | 432 | 424 | 487 | 362 | 448 |
| Services to finance and insurance | 121 | 121 | 177 | 200 | 188 | 188 | 109 | 115 |
| Property and business services | 1,060 | 1,060 | 1,148 | 1,353 | 821 | 965 | 1,033 | 1,056 |
| Other services | 650 | 650 | 808 | 604 | 436 | 246 | 599 | 382 |
| Industry Total | 4,381 | 4,381 | 5,254 | 6,517 | 4,371 | 4,961 | 5,218 | 5,573 |

**COMPANY PROFITS (NR=number-raised (full-sample) DM=direct movement)**

| | *Mar 1997* | | *Jun 1997* | | *Sep 1997* | | *Dec 1997* | |
|---|---|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | 686 | 686 | 571 | 567 | 576 | 578 | 430 | 431 |
| Manufacturing | | | | | | | | |
| Food, beverage and tobacco | 116 | 116 | 146 | 142 | 171 | 183 | 213 | 222 |
| Textiles, clothing, footwear and leather | 68 | 68 | 53 | 51 | 91 | 90 | 77 | 77 |
| Wood and paper products | 50 | 50 | 68 | 66 | 84 | 84 | 80 | 89 |
| Printing, publishing and recorded media | 62 | 62 | 79 | 76 | 87 | 86 | 119 | 110 |
| Petroleum, coal, chemical and assoc. prods | 235 | 235 | 235 | 259 | 251 | 246 | 262 | 265 |
| Non-metallic mineral products | 62 | 62 | 73 | 83 | 89 | 103 | 89 | 103 |
| Metal products | 101 | 101 | 104 | 104 | 164 | 171 | 154 | 155 |
| Machinery and equipment | 137 | 137 | 149 | 163 | 175 | 195 | 180 | 197 |
| Other manufacturing | 12 | 12 | 26 | 24 | 47 | 41 | 31 | 25 |
| Total manufacturing | 843 | 843 | 934 | 967 | 1,160 | 1,198 | 1,205 | 1,242 |
| Construction | 65 | 65 | 135 | 95 | 141 | 66 | 146 | 49 |
| Wholesale trade | 264 | 264 | 460 | 466 | 547 | 539 | 519 | 512 |
| Retail trade | 106 | 106 | 157 | 165 | 151 | 174 | 176 | 200 |
| Transport and storage | 279 | 279 | 217 | 224 | 179 | 204 | 224 | 253 |
| Services to finance and insurance | –62 | –62 | –87 | –98 | –42 | –36 | –49 | –64 |
| Property and business services | 135 | 135 | 200 | 168 | 234 | 191 | 110 | 85 |
| Other services | 233 | 233 | 213 | 222 | 363 | 362 | 315 | 315 |
| Industry Total | 2,548 | 2,548 | 2,800 | 2,776 | 3,308 | 3,277 | 3,076 | 3,024 |

**STOCKS – Direct movement (in the sampled sector)**
**(NR=number-raised (full-sample) DM=direct movement)**

| | Mar–Jun 1997 | | Jun–Sep 1997 | | Sep–Dec 1997 | |
|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | –30.0 | –37.8 | –72.9 | –262.9 | –20.8 | –15.4 |
| Manufacturing | | | | | | |
| Food, beverage and tobacco | 67.4 | 44.9 | –35.8 | 150.1 | 100.8 | 81.1 |
| Textiles, clothing, footwear and leather | –39.6 | –48.2 | 122.9 | 111.2 | –35.2 | –28.4 |
| Wood and paper products | 80.8 | 27.6 | 25.0 | –26.8 | 30.2 | –4.0 |
| Printing, publishing and recorded media | 7.1 | 6.9 | 6.8 | 24.1 | 45.7 | –14.5 |
| Petroleum, coal, chemical and assoc. prods | –82.9 | 84.3 | –131.6 | –83.4 | –30.3 | –33.1 |
| Non-metallic mineral products | 31.3 | –14.9 | 33.0 | –4.9 | 9.9 | –34.7 |
| Metal products | –77.1 | –73.4 | –236.6 | –398.6 | –8.6 | –78.8 |
| Machinery and equipment | 18.8 | –58.6 | –658.8 | –172.8 | –176.4 | –84.0 |
| Other manufacturing | 73.4 | 41.5 | 33.5 | –37.6 | 23.0 | 9.1 |
| Total manufacturing | 79.3 | 10.2 | –841.5 | –438.7 | –40.9 | –187.4 |
| Wholesale Trade | 461.8 | 426.0 | 1925.8 | 1731.7 | –239.5 | 1939.1 |
| Retail Trade | 652.0 | 198.0 | –162.6 | 26.6 | –169.6 | –19.5 |
| Other services | 93.3 | –7.1 | –12.9 | –72.0 | 53.7 | 10.5 |
| Industry Total | 1256.4 | 589.3 | 835.9 | 984.7 | –417.1 | 1727.3 |

**CAPITAL EXPENDITURE – Direct movement (in the sampled sector)**
**(NR=number-raised (full-sample) DM=direct movement)**

| | Mar–Jun 1997 | | Jun–Sep 1997 | | Sep–Dec 1997 | |
|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | –154.9 | 578.0 | –7.0 | –335.1 | 107.7 | 27.0 |
| Manufacturing | | | | | | |
| Food, beverage and tobacco | 47.1 | 44.4 | –3.6 | –23.9 | –26.1 | –31.8 |
| Textiles, clothing, footwear and leather | 13.5 | 4.2 | –8.9 | –8.7 | 20.8 | 15.3 |
| Wood and paper products | 8.4 | 10.9 | –2.0 | –5.4 | 101.9 | 74.3 |
| Printing, publishing and recorded media | 73.7 | 59.5 | –65.7 | –47.8 | 43.6 | 42.5 |
| Petroleum, coal, chemical and assoc. prods | 19.3 | 26.0 | –0.3 | 17.5 | –22.9 | –23.4 |
| Non-metallic mineral products | –6.6 | –20.4 | –44.4 | –38.5 | 20.2 | 26.6 |
| Metal products | 27.7 | 23.5 | –20.8 | –27.3 | 76.9 | 33.5 |
| Machinery and equipment | 56.6 | –27.7 | 28.9 | 16.1 | 116.0 | 134.7 |
| Other manufacturing | 23.6 | 1.2 | –1.7 | –8.4 | 16.0 | 8.4 |
| Total manufacturing | 263.2 | 121.7 | –118.5 | –126.4 | 346.4 | 280.1 |
| Construction | 34.8 | 83.0 | –52.4 | –87.6 | 107.7 | 115.4 |
| Wholesale trade | 219.9 | 822.1 | –38.2 | –245.6 | –63.8 | –48.5 |
| Retail trade | 193.8 | 127.2 | –34.3 | –58.6 | 113.8 | 122.5 |
| Transport and storage | 15.2 | 79.0 | 55.8 | 54.5 | –61.8 | –38.4 |
| Services to finance and insurance | 56.8 | 79.3 | 11.0 | –11.8 | –79.6 | –73.0 |
| Property and business services | 87.8 | 292.9 | –326.8 | –388.2 | 212.7 | 91.9 |
| Other services | 158.0 | –45.7 | –371.8 | –358.0 | 162.8 | 136.3 |
| Industry Total | 873.0 | 2136.0 | –883.0 | –1556.0 | 847.0 | 612.0 |

**COMPANY PROFITS – Direct movement (in the sampled sector)**
**(NR=number-raised (full-sample) DM=direct movement)**

| | *Mar–Jun 1997* | | *Jun–Sep 1997* | | *Sep–Dec 1997* | |
|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | –114.7 | –118.8 | 4.8 | 10.7 | –145.8 | –146.3 |
| Manufacturing | | | | | | |
| Food, beverage and tobacco | 30.6 | 25.7 | 24.8 | 41.6 | 41.9 | 39.0 |
| Textiles, clothing, footwear and leather | –14.7 | –16.3 | 38.4 | 38.7 | –14.0 | –13.6 |
| Wood and paper products | 17.9 | 15.7 | 16.4 | 18.8 | –4.6 | 4.2 |
| Printing, publishing and recorded media | 17.0 | 13.8 | 8.2 | 9.7 | 31.2 | 24.8 |
| Petroleum, coal, chemical and assoc. prods | 0.2 | 23.7 | 15.3 | –12.9 | 11.8 | 19.1 |
| Non-metallic mineral products | 10.8 | 20.7 | 15.5 | 20.3 | 0.2 | –0.2 |
| Metal products | 3.6 | 3.2 | 60.1 | 66.8 | –10.3 | –15.5 |
| Machinery and equipment | 12.7 | 26.1 | 26.0 | 31.8 | 4.3 | 2.1 |
| Other manufacturing | 13.5 | 12.0 | 21.0 | 16.6 | –15.5 | –16.0 |
| Total manufacturing | 91.6 | 124.5 | 225.6 | 231.4 | 44.8 | 44.0 |
| Construction | 70.2 | 30.1 | 5.7 | –29.2 | 4.9 | –17.1 |
| Wholesale trade | 195.7 | 201.7 | 87.3 | 73.1 | –28.2 | –26.8 |
| Retail trade | 50.1 | 58.6 | –5.9 | 9.2 | 25.6 | 25.5 |
| Transport and storage | –61.9 | –55.6 | –38.6 | –19.8 | 45.7 | 49.3 |
| Services to finance and insurance | –25.1 | –35.5 | 45.7 | 61.9 | –7.1 | –28.4 |
| Property and business services | 65.3 | 33.3 | 33.2 | 23.1 | –124.0 | –106.6 |
| Other services | –20.3 | –10.9 | 150.4 | 139.8 | –47.8 | –46.6 |
| Industry Total | 252.0 | 228.0 | 508.0 | 501.0 | –232.0 | –253.0 |

**STOCKS – Full estimates (NR=number-raised (full-sample) DM=direct movement)**

| | *Mar 1997* | | *Jun 1997* | | *Sep 1997* | | *Dec 1997* | |
|---|---|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | 4,313 | 4,313 | 4,468 | 4,460 | 4,463 | 4,265 | 4,368 | 4,176 |
| Manufacturing | | | | | | | | |
| Food, beverage and tobacco | 5,599 | 5,599 | 5,826 | 5,804 | 5,674 | 5,837 | 5,515 | 5,658 |
| Textiles, clothing, footwear and leather | 1,728 | 1,728 | 1,673 | 1,665 | 1,780 | 1,759 | 1,739 | 1,726 |
| Wood and paper products | 1,747 | 1,747 | 1,852 | 1,798 | 1,929 | 1,824 | 1,884 | 1,745 |
| Printing, publishing and recorded media | 822 | 822 | 808 | 808 | 830 | 847 | 863 | 820 |
| Petroleum, coal, chemical and assoc. prods | 5,555 | 5,555 | 5,158 | 5,325 | 5,520 | 5,735 | 5,694 | 5,906 |
| Non-metallic mineral products | 1,333 | 1,333 | 1,337 | 1,291 | 1,353 | 1,268 | 1,312 | 1,183 |
| Metal products | 4,649 | 4,649 | 4,284 | 4,288 | 4,119 | 3,961 | 4,070 | 3,842 |
| Machinery and equipment | 6,250 | 6,250 | 6,032 | 5,955 | 5,559 | 5,968 | 5,716 | 6,217 |
| Other manufacturing | 527 | 527 | 600 | 568 | 638 | 535 | 663 | 546 |
| Total manufacturing | 28,212 | 28,212 | 27,571 | 27,502 | 27,401 | 27,735 | 27,456 | 27,643 |
| Wholesale Trade | 21,660 | 21,660 | 21,474 | 21,439 | 22,402 | 22,172 | 23,097 | 25,045 |
| Retail Trade | 16,959 | 16,959 | 17,254 | 16,800 | 17,203 | 16,938 | 17,270 | 17,156 |
| Other services | 617 | 617 | 708 | 607 | 686 | 526 | 757 | 554 |
| Industry Total | 71,760 | 71,760 | 71,475 | 70,808 | 72,154 | 71,636 | 72,948 | 74,574 |

**CAPITAL EXPENDITURE – Full estimates (NR=number-raised (full-sample) DM=direct movement)**

| | *Mar 1997* | | *Jun 1997* | | *Sep 1997* | | *Dec 1997* | |
|---|---|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | 2,047 | 2,047 | 2,228 | 2,961 | 2,253 | 2,658 | 2,868 | 3,192 |
| Manufacturing | | | | | | | | |
| Food, beverage and tobacco | 501 | 501 | 610 | 608 | 558 | 535 | 547 | 518 |
| Textiles, clothing, footwear and leather | 45 | 45 | 75 | 66 | 55 | 46 | 95 | 80 |
| Wood and paper products | 190 | 190 | 236 | 238 | 162 | 161 | 242 | 214 |
| Printing, publishing and recorded media | 108 | 108 | 178 | 164 | 127 | 131 | 170 | 173 |
| Petroleum, coal, chemical and assoc. prods | 305 | 305 | 347 | 354 | 360 | 384 | 479 | 503 |
| Non-metallic mineral products | 329 | 329 | 291 | 277 | 265 | 257 | 265 | 263 |
| Metal products | 344 | 344 | 459 | 455 | 362 | 351 | 446 | 392 |
| Machinery and equipment | 447 | 447 | 541 | 457 | 550 | 453 | 769 | 690 |
| Other manufacturing | 37 | 37 | 61 | 39 | 63 | 34 | 76 | 40 |
| Total manufacturing | 2,306 | 2,306 | 2,799 | 2,657 | 2,502 | 2,352 | 3,090 | 2,874 |
| Construction | 374 | 374 | 431 | 479 | 377 | 390 | 556 | 577 |
| Wholesale trade | 531 | 531 | 790 | 1,392 | 723 | 1,118 | 809 | 1,219 |
| Retail trade | 401 | 401 | 687 | 620 | 655 | 565 | 907 | 825 |
| Transport and storage | 656 | 656 | 834 | 898 | 646 | 709 | 703 | 789 |
| Services to finance and insurance | 560 | 560 | 694 | 716 | 788 | 788 | 804 | 810 |
| Property and business services | 1,439 | 1,439 | 1,575 | 1,780 | 1,406 | 1,550 | 1,561 | 1,584 |
| Other services | 1,646 | 1,646 | 1,857 | 1,653 | 1,202 | 1,012 | 1,483 | 1,266 |
| Industry Total | 9,960 | 9,960 | 11,894 | 13,157 | 11,552 | 11,141 | 12,781 | 13,137 |

**COMPANY PROFITS – Full estimates (NR=number-raised (full-sample) DM=direct movement)**

| | *Mar 1997* | | *Jun 1997* | | *Sep 1997* | | *Dec 1997* | |
|---|---|---|---|---|---|---|---|---|
| *INDUSTRY (Estimates are in $millions)* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* | *NR* | *DM* |
| Mining | 2,500 | 2,500 | 2,234 | 2,230 | 2,714 | 2,716 | 2,463 | 2,465 |
| Manufacturing | | | | | | | | |
| Food, beverage and tobacco | 838 | 838 | 792 | 787 | 1,024 | 1,036 | 1,313 | 1,322 |
| Textiles, clothing, footwear and leather | 126 | 126 | 110 | 108 | 159 | 158 | 136 | 135 |
| Wood and paper products | 345 | 345 | 394 | 392 | 373 | 373 | 367 | 376 |
| Printing, publishing and recorded media | 344 | 344 | 459 | 456 | 426 | 424 | 499 | 491 |
| Petroleum, coal, chemical and assoc. prods | 816 | 816 | 924 | 948 | 934 | 930 | 1,053 | 1,055 |
| Non-metallic mineral products | 257 | 257 | 337 | 347 | 359 | 374 | 329 | 343 |
| Metal products | 776 | 776 | 403 | 402 | 916 | 922 | 874 | 875 |
| Machinery and equipment | 584 | 584 | 824 | 837 | 733 | 752 | 632 | 649 |
| Other manufacturing | 20 | 20 | 37 | 36 | 60 | 54 | 44 | 38 |
| Total manufacturing | 4,107 | 4,107 | 4,279 | 4,312 | 4,985 | 5,024 | 5,248 | 5,286 |
| Construction | 230 | 230 | 408 | 368 | 306 | 231 | 299 | 202 |
| Wholesale trade | 823 | 823 | 1,048 | 1,054 | 1,280 | 1,272 | 1,030 | 1,023 |
| Retail trade | 526 | 526 | 651 | 660 | 632 | 655 | 1,252 | 1,276 |
| Transport and storage | 784 | 784 | 435 | 441 | 744 | 769 | 758 | 786 |
| Services to finance and insurance | –1 | –1 | 25 | 14 | 54 | 60 | –141 | –156 |
| Property and business services | 200 | 200 | 343 | 311 | 475 | 432 | 307 | 282 |
| Other services | 1,034 | 1,034 | 974 | 984 | 1,386 | 1,385 | 1,381 | 1,381 |
| Industry Total | 10,202 | 10,202 | 10,398 | 10,374 | 12,576 | 12,544 | 12,597 | 12,544 |

# 5. EXCLUSION OF SMALL BUSINESSES

A recent initiative in the QEWS strategy was to test the effectiveness of excluding small businesses from the coverage of the Survey (though not from the scope). Part 1 of this paper presents a study into the impact of excluding the small businesses from some ABS quarterly surveys. To test the effects on the Stocks and Sales Survey of excluding small businesses and using the direct movement together, tables 5.1 and 5.2 have been constructed.

Table 5.1 shows the proportion of population and sample that is contained in possible small business populations. This report looks only at two definitions of small business; companies which employ between 0–4 persons ('micro' businesses), and secondly, companies which employ between 0–19 persons ('small businesses', as given by Register employment at the time of frame extraction). Table 5.1 shows that about two-thirds of Stocks in-scope companies are micro businesses, and about 95% of companies are micro or small companies. The design of Stocks and Sales leads to sampling disproportionately from the larger companies.

**Table 5.1 Number of units by size in the Stocks sample and frame**

| | *Frame* | | *Sample* | |
|---|---|---|---|---|
| *Size* | *Count* | *%* | *Count* | *%* |
| 0–4 | 208,096 | 65.10 | 3,023 | 40.2 |
| 5–19 | 93,416 | 29.30 | 1,181 | 15.7 |
| 20–49 | 11,758 | 3.70 | 817 | 10.9 |
| 50–99 | 3,230 | 1.00 | 740 | 9.8 |
| 100–249 | 1,673 | 0.52 | 695 | 9.2 |
| 250–499 | 557 | 0.17 | 557 | 7.4 |
| 500–1000 | 305 | 0.09 | 305 | 4.1 |
| 1000+ | 202 | 0.06 | 202 | 2.7 |
| TOTAL | 319,237 | | 7,520 | |

The amount of rotation in a ‘standard’ quarter of 8.5% rotation sees the sample rotations as specified in Table 5.2. The amount of rotation that occurs in the 0–19 strata corresponds to about two-thirds of all rotation. Thus, most of the scope for direct movement estimation to be effective will come from the small business strata.

**Table 5.2 Number of units rotating in and out of sample in Stocks**

| Employment Size | Population in Sep 98 | Population in Dec 98 | Sample in Dec 98 | Continuing units Count | Continuing units % | Rotating in Count | Rotating in % |
|---|---|---|---|---|---|---|---|
| Total | 333,113 | 338,838 | 7,509 | 6,968 | 92.8 | 541 | 7.2 |
| 0–19 | 314,186 | 319,707 | 4,200 | 3,841 | 91.5 | 359 | 8.5 |
| 20–49 | 12,701 | 12,853 | 817 | 751 | 91.9 | 66 | 8.1 |
| 50–99 | 3,386 | 3,416 | 732 | 683 | 92.9 | 52 | 7.1 |
| 100–249 | 1,756 | 1,780 | 676 | 638 | 94.5 | 37 | 5.5 |
| 250–499 | 599 | 601 | 599 | 578 | 96.2 | 23 | 3.8 |
| 500+ | 485 | 481 | 485 | 477 | 99.2 | 4 | 0.8 |

Table 5.3 gives the contribution to standard errors for both the usual and direct estimators of movement. The table shows that the benefits from using the direct estimator of movement is mostly in the micro and small business strata. The Direct Movement estimator has a similar RSE to the full-sample estimator when units with 0–19 employment are excluded. There are still gains to be made when only the units 0–4 are excluded; the RSE's for the direct movement estimator is considerably less than the full-sample estimator.

**Table 5.3 RSEs of the Direct movement Estimator when excluding micro and small businesses**

| | Usual full-sample movement estimator: non-September Quarter | Usual full-sample movement estimator: September Quarter | Direct estimator |
|---|---|---|---|
| Stocks RSE(%) | | | |
| Full-Sample | 0.86 | 0.70 | 0.40 |
| Exclude 0–4 | 0.58 | 0.52 | 0.35 |
| Exclude 0–19 | 0.32 | 0.30 | 0.28 |

| | Usual full-sample movement estimator | Direct estimator |
|---|---|---|
| Sales RSE(%) | | |
| Full-Sample | | |
| Manufacturing | 1.11 | 0.62 |
| Wholesale | 2.15 | 0.84 |
| Exclude 0–4 | | |
| Manufacturing | 0.88 | 0.55 |
| Wholesale | 1.70 | 0.72 |
| Exclude 0–19 | | |
| Manufacturing | 0.65 | 0.49 |
| Wholesale | 0.90 | 0.55 |

(The table gives a separate column for the September quarter estimates as these are affected significantly by information that is fedback to the frame from the Manufacturing Survey.  This feedback causes large sample rotation and leads to inefficiency in estimates of movement, as a new sample with a high percentage of new sample units results).

The tables above demonstrate that the direct movement estimator can be profitably used when excluding companies with an employment of 0–4.  However, when excluding companies with employments of 0–19, the gains in using the direct movement estimator become marginal for the stocks variable, though the sales variable does seem to have some gains still.

This paper has not examined the bias that occurs when excluding the small businesses from the direct movement estimator.  Part 1 of this report looked at the effect of excluding small businesses using the usual full-sample estimator of movement.  It is believed that the mean-square error results there would apply to the direct movement estimator as well.

# 6. SUMMARY

This report has detailed two possible enhancements that were considered for inclusion in the ABS Quarterly Economy Wide Survey (QEWS). Both these initiatives were assessed for their ability to increase the precision of estimates.

The first of these, the exclusion of small businesses, does appear to have some potential. It indicates that the information being collected from small businesses may not be enhancing the estimates of change. This report does not speculate as to why this may be occuring.

Secondly, the report has shown that the direct movement estimator may have some value in reducing the standard error of the estimates of movement. It comes at the cost of increased complexity of estimation. In addition, the direct movement estimator derives most of its utility from being applied to the smaller businesses (those with employment less than 20 employees). The exclusion of these companies would appear to invalidate the use of the direct movement estimator. The added complexity and re-basing necessary to make the direct movement estimator acceptable to users, does not appear to be warranted unless the smaller businesses are included.

## 7. POINTS FOR DISCUSSION

- Are other bias reduction strategies that, whilst also meeting the goal of reducing provider load, could lead to better Root Mean-Square Errors
- What are the additional gains of a minimum variance composite estimator over DIME; and whether the gain warrants the additional complexity in computation (noting that the formula for the MV composite estimator is given in the Appendix).
- Are there any gains to be made in using other forms of auxiliary information in helping to measure movement?

# APPENDIX

## A. FORMULATION OF THE DIRECT MOVEMENT ESTIMATORS

### Notation used in the Appendix

Define the following:

$x_i$ = variable of interest (e.g. stock levels for companies)

$x''_k$ = direct movement estimator of level (DIMEL)

$x'_k$ = full-sample estimate of level at time $k$

$x'_{k,j}$ = estimator of level at time $k$ using units common in sample to time $j$

$\sigma_k^2$ = sampling variance of full-sample estimate of level at time $k$

$r$ = rotation rate in the sample (i.e. the proportion of the time 1 sample that leaves the sample between time 1 and time 2)

$\rho$ = population correlation of $x_i$ between two successive quarters

$p$ = proportion of frame that is occupied by frame deaths or frame births

## A.1 Standard error of the DIMEL

The DIMEL for the $k^{\text{th}}$ quarter after a benchmarking is given as:

$$\begin{aligned} x_k'' &= x_1' + (x_{2,1}' - x_{1,2}') + (x_{3,2}' - x_{2,3}') + \ldots + (x_{k,k-1}' - x_{k-1,k}') \\ &= x_1' + m_{2,1}' + m_{3,2}' + \ldots + m_{k,k-1}' \\ &= \mathbf{1}'\boldsymbol{x} \end{aligned} \tag{1}$$

where

$\mathbf{1}'$ = k-vector of ones

$\boldsymbol{x}$ = $\left(x_k', (x_{2,1}' - x_{1,2}'), (x_{3,2}' - x_{2,3}'), \ldots, (x_{k,k-1}' - x_{k-1,k}')\right)$

Assume for the moment that the frame is constant between the two time periods (i.e. there are no frame births or deaths).

The variance of this estimator of level is given by

$$\begin{aligned} Var\left(x_k'\right) &= Var\left(\mathbf{1}'\boldsymbol{x}\right) \\ &= \mathbf{1}'\boldsymbol{\Sigma}\mathbf{1} \end{aligned} \tag{2}$$

where

$$\boldsymbol{\Sigma} = Cov\left(x_{k,j}' - x_{j,k}', x_{l,m}' - x_{m,l}'\right)$$

The determination of the elements of $\boldsymbol{\Sigma}$ can be facilitated by the use of some assumptions regarding the sampling correlation between estimators k-quarters apart. The correlation structure assumed is:

$$\rho^k = \left(\rho, \rho^{1+\gamma}, \ldots, \rho^{1+k\gamma}\right) \tag{3}$$

This correlation structure assumes a type of geometric decay. The rate of decay is determined by the parameter $\gamma$. This two-parameter model is simplistic when dealing with real data. The series that occur in quarterly Economic Surveys do not have such a predictable sampling correlation structure. Typically, there will be a more complex structure, still dominated by the amount of overlapping sample and also by the population correlation that exists between data k-quarters apart. The amount of overlapping sample can be assumed reasonably constant; however, the between-quarter population correlation structure is a less tractable parameter to explain.

Given this sampling correlation structure, the correlation-matrix $\boldsymbol{\Sigma}$ becomes:

$$\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{\Sigma}^* \tag{4}$$

where

$$\Sigma^* = \begin{bmatrix} 1 & & & \\ (\rho-1) & 2(1-\rho)/(1-r) & & \\ \rho(1-\rho)(r-1) & (1-\rho)^2(2r-1) & 2(1-\rho)/(1-r) & \\ \rho[2](1-\rho)(2r-1) & \rho(1-\rho)^2(3r-1) & (1-\rho)^2(2r-1) & 2(1-\rho)/(1-r) \\ \rho[3](1-\rho)(3r-1) & \rho[2](1-\rho)^2(4r-1) & \rho(1-\rho)^2(3r-1) & (1-\rho)^2(2r-1) \\ \rho[4](1-\rho)(4r-1) & \rho[3](1-\rho)^2(5r-1) & \rho[2](1-\rho)^2(4r-1) & \rho(1-\rho)^2(3r-1) \end{bmatrix}$$

where $\rho[k] = \rho^{1+(k-1)\gamma}$ and $\gamma = 1, 2, \ldots$ , with $r$ = rotation rate of the sample.

*Inclusion of Frame births and deaths*

The inclusion of frame births and frame deaths necessitates a modification to $\mathbf{\Sigma}$*. The necessary changes are detailed below.

The impact of frame changes is to significantly raise the variance of the direct movement estimator. In fact, in trails of the direct movement estimator, when the frame changes occupied 10% of the frame, the variance increased by at least 30%.

$$Var\left(x'_{k,j} - x'_{j,k}\right) = 2\left[(1-\rho)(1-p)/(1-r) + p\right]\sigma^2$$

$$Cov\left(x'_1, x'_{k,j} - x'_{j,k}\right) = \rho^d\left(1-\rho(1-p)\right)(dr-1)\sigma^2$$

where $d = k-2$

$$Cov\left(x'_{k,j} - x'_{j,k}, x'_{m,n} - x'_{n,m}\right) = (1-p)\rho^{1+(m-k-2)\gamma}\left(2\rho^\gamma - \rho^{2\gamma} - 1\right)\left((m-k+1)r-1\right)$$

where $\gamma = 1, 2, \ldots$

## A.2 Revisions

The direct movement estimator assumes that an estimate of level will be derived by adding successive estimates of quarterly movement to a base-year estimate of level. This paper has shown that the process will produce an estimate of level (DIMEL) with a rapidly decreasing quality (as measured by variance). However, the estimate of level is unbiased.

In order to correct the DIMEL, it is necessary re-base it to a series producing superior quality estimates of annual level. The amount of noise in the DIMEL series can be easily quantified (as seen above) and the amount of revision to the series is directly related to the amount of noise.

Assume that the error in the DIMEL at time $k$ is $\varepsilon_k$. Also, assume that the error, $\varepsilon_k$, is a normally distributed variable. The expected revision will then be

$$\text{Expected Revision} = \mathrm{E}(|\varepsilon_k|) \text{ where } \varepsilon_k \cap N(0, \sigma_k^2)$$

Then the expected revision is:

$$\begin{aligned} E\left(\left|\varepsilon_k\right|\right) &= \int_{-\infty}^{+\infty} \left|\varepsilon_k\right| f(\varepsilon_k)\, d\varepsilon_k \\ &= 2\int_0^{\infty} \varepsilon_k\, f(\varepsilon_k)\, d\varepsilon_k \\ &= \frac{2}{\sigma_k \sqrt{2\pi}} \int_0^{\infty} \varepsilon_k \exp\left(\frac{-\varepsilon_k^2}{2\sigma^2}\right) d\varepsilon_k \\ &= \sigma_k \sqrt{2/\pi} \end{aligned}$$

This revision does not take into account the amount of error contributed by the annual benchmarking series used. However, it demonstrates that the revision is at least proportional to the standard error of the DIMEL series at time $k$. This indicates that regular re-basing will be required.

## A.3 Standard error of Direct movement estimator

Assume that the population estimate of movement is given as the difference of the level estimators at time 2 and time 1. This is a 'natural' definition of movement. That is,

$$M_{21} = X_2 - X_1$$

The usual full-sample estimator of movement of this quantity is

$$m'_{21} = x'_2 - x'_1$$

where $$x'_2 = \sum_{i=1}^{n} \frac{N}{n} x_{2i}$$

and $$x'_1 = \sum_{i=1}^{n} \frac{N}{n} x_{1i}$$

This estimator assumes a constant sample size between the two time periods.

The summation includes all units in sample.

This estimator of movement suffers from having included in it all non-common units. This inflates the variance of the movement estimator. It is easy to intuitively see this. A continuing unit in a survey such as a stocks survey generally tends to have a highly correlated stocking level between quarters. However, if a unit leaves the sample at after time 1, and another unit comes in to replace this unit at time 2, then this unit will generally have a stocking value significantly different to the other unit. This is simply to say that the between-quarter correlation for a continuing unit is significantly higher than the correlation that exists between two units randomly selected at one time point.

The direct-movement estimator (DIME) can be written as:

$$m''_{21} = x'_{21c} + x'_{2b} - x'_{12c} - x'_{1d}$$

where

$x'_{21c}$ = estimate of total (from the common sample) of the common population using inverse selection probabilities as weights at time 2,

$x'_{2b}$ = estimate of total of the frame births,

$x'_{1d}$ = estimate of total of the units at time 1 that will be dead at time 2.

The estimator above is a type of composite estimator, with the exception being that it does not attempt to use optimal weights. In fact, the weights that have been used in this paper were the post-stratified weights of the categories frame births, deaths and continuing. An optimal composite estimator of movement ($m_{21}^{*}$) would have attempted to determine coefficients $a$, $b$, $c$ and $d$ which minimised the variance of the quantity

$$m_{21}^{*} = Na\,\bar{x}_1 + Nb\,\bar{x}_{1c} + Nc\,\bar{x}_2 + Nd\,\bar{x}_{2c}$$

where

$\bar{x}_{1c}$ = mean of the units at time 1 that are common to time 2

$\bar{x}_1$ = mean of the units at time 1 that will be deaths at time 2

$\bar{x}_{2c}$ = mean of the units at time 2 that were common to time 1

$\bar{x}_2$ = mean of the units at time 2 that are frame births

where the coefficients are constrained by the relationships

$$\begin{aligned} a + b &= -1 \\ c + d &= 1 \end{aligned}$$

which follows from the observation that the expected value of $m_{21}^{*}$ is

$$E\left(m_{21}^{*}\right) = (c+d)X_2 + (a+b)X_1$$

and to make this an estimator of movement, the constraints in the coefficients must hold.

The optimal solution would then go on to show that the optimal coefficients (assuming the same population variance between quarters)

$$a_{opt} = {}^{-r(1-\rho)}\!\big/_{(1-\rho r)}$$

and

$$b_{opt} = {}^{(1-r)}\!\big/_{(1-\rho r)}$$

with

$$c_{opt} = -a_{opt}$$

and

$$d_{opt} = -b_{opt}$$

The composite estimator thus would assign fixed weights to the units.

It is interesting to note that the optimal coefficients will produce an estimator of the form:

$$m_{21}^{*} = \frac{Nr(1-\rho)}{(1-\rho r)}\left[\overline{x}_2 - \overline{x}_1\right] + \frac{N(1-r)}{(1-\rho r)}\left[\overline{x}_{2c} - \overline{x}_{1c}\right]$$

In the case of the Stocks Survey, the correlation, $\rho$, typically assumes a value of about 0.90. The rotation rate, $r$, is usually about 10%. Substituting these, gives values of

$$m_{21}^{*} = \frac{N}{91}\left[\overline{x}_2 - \overline{x}_1\right] + \frac{90N}{91}\left[\overline{x}_{2c} - \overline{x}_{1c}\right]$$

This estimator then weights the continuing units (in $\overline{x}_{2c}$ and $\overline{x}_{1c}$) very highly (almost completely retaining their selection weights), but strongly down-weights the non-continuing units.

The optimal composite estimator is more difficult to implement than the direct movement estimator. It requires an optimal set of weights to be calculated at stratum level at regular intervals. Current ABS generalised estimation programs would be able to handle the direct movement estimator with a small amount of modification. The optimal composite estimator would need further study to see if this is the case. In addition, it is uncertain how non-response could be best handled in the optimal composite estimator.

The direct movement estimator performs in a very similar manner. It downweights the non-continuing units (that are not frame births or deaths) totally. The continuing units essentially maintain their selection weights. However, the direct movement estimator uses the additional information from the frame in the form of the population counts of the continuing and non-continuing units, to produce post-stratified weights.

The variance of the direct movement estimator is easily derived, and is given as

$$Var\left(m_{21}''\right) = 2\sigma^2\left[\frac{(1-\rho)(1-p)}{(1-r)} + p\right]$$

The variance of the standard movement estimator is given as

$$Var\left(m_{21}'\right) = 2\sigma^2\left[1-\rho(1-r)\right]$$

Taking square-roots and dividing, this gives that the ratio of the standard error of the direct movement estimator $(m_{21}'')$to the standard error of the standard movement estimator $(m_{21}')$ is:

$$\frac{SE\left(m_{21}''\right)}{SE\left(m_{21}'\right)} = \sqrt{\frac{(1-\rho)(1-p)+p(1-r)}{(1-r)\left(1-\rho(1-r)\right)}}$$

## 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 |

2000001522974
ISBN 0 642 48128 8

**RRP $11.00**