

**Testing a new questionnaire for the Australian Bureau of  
Statistics Labour Force Survey**

**Alan D. Herning, Elizabeth J. Finlay**

**Methodology Division**

**Australian Bureau of Statistics**

**November 1998**

# CONTENTS

1 Introduction	2
2 The Labour Force Survey	2
2.1 Scope	2
2.2 Sample	2
2.3 Collection Methodology	3
3 Questionnaire Redesign	3
3.1 Objectives	3
3.2 Constraints	4
3.3 Questionnaire Changes	4
4 Testing The New Questionnaire	6
4.1 Testing Strategies	6
4.2 Analysis Of Testing	7
5 Composite Estimation Of Questionnaire Effect	9
5.1 Strategy For Assessing The Statistical Impact Of The New Questionnaire	9
5.2 A Model For An Additive Questionnaire Effect	10
5.3 Level Estimate	11
5.4 Variance Of A Linear Combination Of RG Estimates	11
6 Comparison Of Assessment Strategies	12
6.1 Estimating The Variances	12
6.2 Discussion Of Results	12
6.3 The Optimal Strategy	13
7 Conclusion	14
8 Issues For Discussion	14

# 1 Introduction

1. In repeated surveys the focus is often on the movement of time series of key characteristics of interest. Changes to the methodology used in these surveys can often affect these key estimates, resulting in trend breaks or large movements. When changes in methodology are contemplated it is desirable to attempt to quantify these effects.

2. This paper describes the testing of a new questionnaire for the Australian Labour Force Survey (LFS). Testing strategies used in the development of the new questionnaire are described, followed by a discussion of the statistical design and estimation of the statistical impact of the new questionnaire on the key series. The statistical analysis and associated testing strategy that have been developed take advantage of the strong correlations between estimates from the same individuals over time. Under the assumption that a respondent's replies do not depend on the questions they were asked in a previous month, the statistical estimator described in this paper is the best linear unbiased estimator of the new questionnaire effect.

3. Section 2 of this paper provides some background to the LFS. Section 3 describes the motivation behind the introduction of the new LFS questionnaire. Section 4 describes testing strategies used in the questionnaire development. Section 5 describes three competing statistical designs which could be used in the estimation of the statistical impact of the new questionnaire. Also described in this Section is a simple additive model for questionnaire effect, and a composite estimator of this effect is derived. The composite estimator is unbiased and has minimum variance under a known autocovariance structure (which in practice is estimated from past data). The three statistical strategies are compared in Section 6.

## 2 The Labour Force Survey

### 2.1 Scope

4. The LFS includes both private dwellings (PDs) and special dwellings (SDs) in its sample. This also includes those dwellings in sparsely settled areas. Both usual residents and visitors in the household are included, however, Australian military personnel are excluded.

### 2.2 Sample

5. The LFS achieves a sample of approximately 35,000 households, which equates to approximately 62,000 completed questionnaires. The dwellings included in the survey are selected by using a multi-stage sampling scheme. Firstly, geographic areas known as Collection Districts (CDs) are selected and, secondly, dwellings are selected within those CDs. The sample of CDs is divided into eight 'rotation groups' (RGs), with the dwellings in one RG being rotated out of the sample each month and being replaced with other sampled dwellings from the same CD. This scheme ensures that selected dwellings remain in the survey for eight months, before being replaced by nearby dwellings.

6. The LFS provides statistics relating both to Australia as a whole and to each State separately; this requires that the sample be spread over each State. The survey has been designed to provide very precise estimates - the estimate of employed persons for Australia has a relative standard error of around 0.2 %.

7. Further details about the design of the LFS sample are provided in Australian Bureau of Statistics (1997a).

## **2.3 Collection Methodology**

8. The 'any responsible adult' (ARA) methodology is used. Under this methodology, the survey questions are asked of the first responsible adult (usual resident aged 18 years or over) contacted by the interviewer. This person answers the questions on behalf of all members of the household in on scope and coverage for the survey. More information about the collection methodology and conceptual framework can be found in Australian Bureau of Statistics (1995).

9. For the first month that a dwelling is included in the survey, an interviewer makes contact with the usual residents of the dwelling and conducts the interview face-to-face. Where possible, the second and following months that the dwelling is in the survey, the interview is conducted over the phone (telephone interviewing, TI). Where dwellings do not have a phone, or do not wish to be interviewed by phone, interviewers continue to conduct monthly interviews face-to-face. Further details about the collection methodology and the effects of TI are given in Australian Bureau of Statistics (1997b).

## **3 Questionnaire Redesign**

### **3.1 Objectives**

10. The LFS has been conducted quarterly from 1960 to 1977, and monthly from February 1978 onwards. The last time a major redesign of the LFS questionnaire was undertaken was prior to the introduction of monthly surveys in 1978. For further information about the questionnaire currently used in the LFS, see Australian Bureau of Statistics (1993).

11. In view of changes and restructuring in the labour market, Australian Bureau of Statistics (ABS) has concluded that a redesign of the survey questionnaire is warranted. While the current questionnaire continues to provide a wide range of relevant labour force information, ABS believes it could be improved to better measure contemporary labour force characteristics. Some aspects of the labour market are now more prominent than when the current questionnaire was designed, while others have declined in importance. For example, part-time work has become increasingly important, and there are more non-standard work arrangements in the labour market.

12. As well, users of the survey data have requested additional monthly labour force data items. Developments in international standards and practices for collecting labour statistics also indicate that some of the current LFS data items need to be reviewed.

13. The main objectives and strategies of the LFS questionnaire redesign are as follows.
- Meet user requirements for additional and more relevant data by:
    - a) adding new data items; and
    - b) extending or altering existing data items.
  - Improve the quality and stability of current data by:
    - a) applying the survey concepts more accurately and consistently; and
    - b) improving the field interview to reduce respondent load and other sources of non-sample error.
  - Better cater for current labour market conditions, statistical standards and collection methodologies by:
    - a) ensuring relevance to current ABS and labour market conditions and using up-to-date terminology; and
    - b) taking current international standards for collection of labour force statistics into account.

## **3.2 Constraints**

14. The redesign has been undertaken with an overriding requirement of ensuring the comparability of time series for key labour force variables produced by the current and new questionnaires. That is, no unquantifiable breaks in series will result for the key data items of employment, unemployment and not in the labour force. The statistical impact of the new questionnaire on these items will be evaluated prior to implementation, to ensure this constraint has been met.

15. While questions relating to key labour force variables will not change, a number of minor definitional changes will be made to these items. In order to rectify inconsistencies with the latest international recommendations on labour statistics, three conceptual changes will be made. These will be implemented by additional questions, or by changes in data derivations. Each of the changes is quantifiable; two are of negligible size and the other is of a size yet to be determined. As these are directly measurable, comparability of key labour force series will be maintained. To alleviate respondent load, one additional change will be made to a key data item. Section 3.3 discusses these changes in further detail.

16. An additional constraint of the redesign is that the new questionnaire must be cost neutral.

## **3.3 Questionnaire Changes**

### **3.3.1 Key Labour Force Variables**

17. The following changes have been made to key labour force variables.
- People who are not working, who are actively seeking work, but who are unavailable to start work during the reference week due to temporary illness, will be classified as not in the labour force instead of unemployed, as is currently the case.
  - Persons who are unpaid helpers in businesses operated by a direct member of the person's family (contributing family workers) who are away from work will no longer be classified as employed. These people will be classified as

unemployed, or not in the labour force, depending on their recent job search activity and availability for work.

- The new questionnaire will more completely identify people waiting to start a job they have already obtained. This will result in some new job starters being classified as unemployed, instead of not in the labour force, as is currently the case.
- The problem of respondent load for elderly people who are permanently not intending to work will be addressed in the redesign. It is proposed to alter the question which checks whether they worked in the reference week to include a new response category. This category will identify people aged 65 years or more who clearly volunteer that they are permanently not intending to work. They will be asked no further questions.

### **3.3.2 Non-Key Labour Force Variables**

18. It is not appropriate in this paper to provide a comprehensive list of the many changes that are being made to non-key items. While it has been possible, within the constraints of the redesign, to make significant changes to questions that collect non-key items, the focus has been to maintain continuity for all series. The basic interview approach used currently has been retained, and most survey questions have been altered where necessary, rather than re-written. Where new data series are being produced, current data series will, in most instances, continue to be available.

19. Perhaps the most significant change to non-key items, in terms of reduction of respondent load, will be applied to the unemployed population. Currently the details of the last full-time job held are collected each month. In the new questionnaire, this information will be collected in relation to the last job held (that lasted two weeks or more) and will be collected quarterly. This will be a sizeable decrease in respondent load, particularly for the long-term unemployed, and is comparable to the way in which job details are collected for employed persons.

20. The following data will be collected in greater detail in the new questionnaire:

- persons who are marginally attached to the labour force;
- the number of hours actually worked in the reference week by multiple job holders; and
- when unemployed persons last worked.

21. The following data will be collected in the new questionnaire that are not collected currently:

- job tenure (quarterly);
- availability of underemployed persons to work more hours in the reference week; and
- the number of hours that are usually worked by employed persons.

## 4 Testing The New Questionnaire

22. A thorough review of the LFS questionnaire took place in order to assess the changes that would need to be made and then tested. This review took a number of forms.

- The users of the data were consulted to identify what data are used extensively and what data are only of minimal interest.
- The concepts used in the questionnaire were assessed against the latest international standards.
- Extensive research was conducted into the changes that other statistical agencies had made to the questionnaires used in their surveys of employment and unemployment.
- Focus groups of interviewers were organised to identify the areas of the questionnaire that cause difficulty during enumeration.
- Formal field queries submitted over many years documenting situations for which interviewers had to seek clarification were examined and analysed.

### 4.1 Testing Strategies

23. One of the primary aims of testing the new questionnaire is to minimise non-sampling error. Part of this minimisation will be achieved by ensuring that the concepts and definitions embodied in the questionnaire are understood and applied consistently by interviewers. An additional, and equally important part, is that the question wording is able to be understood consistently by respondents, resulting in reliable data.

24. The testing process comprises more than fine-tuning the question wording. It also comprises testing the accuracy of the questionnaire as a collection instrument (e.g. the accuracy with which sequence guides are followed by interviewers) and the optimal positions of questions and modules of questions. Moving beyond issues relating directly to the questionnaire, minimising non-sampling error also necessitates testing that the interviewer and office documentation, training materials and office procedures perform in accordance with expectations.

25. Rigorous testing of this kind will produce a questionnaire for which the statistical impact can be effectively evaluated without confounding factors such as inconsistently applied definitions.

26. For the redesign of the LFS questionnaire, the testing strategy comprises two pilot tests and a dress rehearsal. The first pilot test was conducted on approximately 600 dwellings which were contacted once only. This test was used to test a questionnaire which incorporated changes resulting from the review process described previously, as well as contemporising the questionnaire by making it relevant to the more diverse labour market that exists in Australia today (in comparison to twenty years ago).

27. The second pilot test was conducted on approximately 750 dwellings, contacted twice with a 4 week gap between contacts. This test provided the opportunity to fine-tune and finalise the content and wording of the questionnaire. In addition, the second pilot test was used to test the associated processing systems and office procedures. Enumerating the questionnaire over two contacts allowed telephone

interviewing to be introduced to test the questionnaire and procedures in this environment. This was done by conducting face-to-face interviews at the first contact and requesting phone numbers in order for the second contact interviews to be conducted by telephone.

28. The dress rehearsal will be conducted on approximately 500 dwellings which will be contacted twice, again with a four week gap between contacts. This will test the new questionnaire (in face-to-face and telephone interviewing environments), associated training procedures, office procedures and processing systems. The expectation is that these will remain unchanged from the dress rehearsal for the measurement of statistical impact and then for implementation.

## **4.2 Analysis Of Testing**

29. While many aspects of the LFS have been evaluated as part of the redesign, when choosing areas of the survey to analyse in depth, a judgement has been made as to which issues, data items or questions are likely to have the biggest impact on published estimates. These areas are discussed briefly below.

### **4.2.1 Qualitative Analysis**

30. The purpose of qualitative analysis is to identify areas of the questionnaire, documentation, training or office procedures which lead to poor quality data. For example, incomplete documentation may result in some situations being dealt with differently by different interviewers because of lack of explanation and guidance on the issue. Another example is that some words are interpreted slightly differently by respondents in different areas of Australia. This is identified by interviewer feedback.

31. The types of qualitative analysis used in the redesign of the LFS questionnaire explored how interviewers applied concepts and dealt with problems during the test. In depth information was also sought on respondent reaction and apparent understanding of the questions.

32. As part of each test, interviewers were required to complete a debriefing form which focused on specific aspects of the questionnaire and enumeration, as well as giving them an opportunity to make general comments. This form was completed by the end of enumeration and before the interviewers met together at the debriefing session. This ensured that comments made on the debriefing forms were the interviewers' independently held opinions, unswayed by other interviewers' thoughts and contributions.

33. At the end of the enumeration of each test, interviewers were brought together to attend a debriefing session. The debriefing session explored areas of concern more thoroughly than the debriefing form as the facilitator was able to guide the discussion and probe in depth when that was required. However, this form of evaluation is less likely to obtain interviewers' independently held opinions - rather, after discussion of the issues, it will obtain the group consensus.



34. Some examples of changes that were made as a result of this process for the first pilot test are:

- directions were identified that needed to be included in the interviewer documentation in order to obtain country of birth information that is able to be coded to the classification;
- the structure of the front page of the form was altered so that all the questions that relate to visitors are in the one column instead of spread over two columns; and
- the flow of questions for persons who held more than one job in the reference week was altered so the shift between questions relating to their main job and questions relating to their second job occurred less frequently.

35. Another type of qualitative analysis to be conducted will be to examine the responses for each question in the context of the response history for the equivalent question in the current questionnaire. A response history for a number of questions has been compiled from data covering the period 1990 to 1997. This will be used to prepare control charts which will indicate whether the test data falls into the historical ranges for the series. Obviously, this would only detect gross departures in the series (of the kind that may indicate sequencing errors). More subtle shifts in the series would not be detected using this analysis.

#### ***4.2.2 Quantitative Analysis Of Pilot Test Data***

36. One key area of interest was the comparison of new questions with old questions collecting information about the number of hours actually worked during the reference week. There was sufficient power in the second pilot test sample to measure this relatively accurately. The number of hours actually worked in the reference week is currently collected using a number of prompts which encourage people to remember the days they worked last week and whether they worked extra time or had any time off in the reference week. In the new questionnaire, these prompts have been combined into the one question instead of being asked separately.

37. It is important to identify whether the new question is collecting information of the same quality as the current question. The question asked currently (including the prompts) was placed at the end of the second pilot test questionnaire so that people who were at work last week were asked both the new question and the current question.

38. Given the relatively large sample involved (about 750 dwellings), it is anticipated that the data will be initially analysed by using a standard z-statistic. The test will involve calculating, at a unit record level, the difference between the responses given to the two different questions about the hours actually worked. The dataset of differences will then be used to test the hypothesis that the mean difference in responses is zero. The distribution of differences will also be examined and outliers analysed to determine any reasons for large differences (e.g. clerical error). This analysis will be complete by the end of November.

## 5 Composite Estimation Of Questionnaire Effect

### 5.1 Strategy For Assessing The Statistical Impact Of The New Questionnaire

39. To determine what effect, if any, the new questionnaire will have on the key labour force estimates, the ABS is planning to test the questionnaire in the field and compare the estimates derived to those obtained using the existing questionnaire. The general approach for measuring the differences between two methodologies is to conduct a 'parallel run'. However, this option involves setting up a parallel sample and is often too expensive. In the case of the LFS questionnaire redesign, 'live' LFS sample will be split allowing the estimate of difference between the two methodologies to be based on comparisons between two portions of the 'live' sample.

40. Three possible testing strategies are proposed and compared in this paper. The first strategy is for seven eighths of the 'live' LFS sample to continue to be surveyed using the current questionnaire and the remaining eighth to use the new questionnaire for a number of months with a simple level estimator being used to calculate the difference between the two questionnaires. The second strategy is the same as the first except that a composite estimator is used to calculate the difference between the two questionnaires. The third strategy also uses an eighth of the 'live' sample to test the new questionnaire, but a different eighth of the sample is used each month with composite estimation. To give an example of these strategies the diagrams below show the proposed strategies if the statistical impact study was conducted over 4 months.

**Diagram 1: Proposed testing strategies 1 and 2, for constant eighth of "live" sample to use new questionnaire**

Rotation Group

Month	7	8	1	2	3	4	5	6
Oct '99								
Nov								
Dec				Q				
Jan '00				Q				
Feb				Q				
Mar				Q				
Apr								
May								

41. In diagrams 1 and 2, Q denotes that the new questionnaire is being used and the shaded regions indicates when a new rotation group is rotated into the sample. In Diagram 1, a single rotation group is interviewed using the new questionnaire for four successive months, as in strategies 1 and 2. Diagram 2 shows strategy 3, in which a different rotation group uses the new questionnaire each month. Each month, the rotation group which is being interviewed for the fourth time is interviewed using the new questionnaire. Any bias which may be present in the responses as a result of the

rotation pattern are expected to be close to zero by the fourth month in the survey (see Bell (1998)).

**Diagram 2: Proposed testing strategy 3, for rotating eighth of 'live' sample to use new questionnaire**

Rotation Group

Month	7	8	1	2	3	4	5	6
Oct '99								
Nov								
Dec		Q						
Jan '00			Q					
Feb				Q				
Mar					Q			
Apr								
May								

42. These strategies only require a study over a short period, and with a minimum of extra expenditure. Options involving a large or extended parallel sample were rejected. There are advantages in using the 'live' LFS sample for the study, besides the obvious advantage of lower cost. In particular, for the 'live' sample, previous data for the same dwellings based on a different questionnaire is available, which can be used to improve the comparisons.

## 5.2 A Model For An Additive Questionnaire Effect

43. It is computationally easier to assign questionnaires to sampled units according to the rotation group (RG) the unit was sampled in. There are eight RGs and, in any month, each RG contains dwellings that can be characterised by the number of months that the dwellings have been surveyed. Thus RGs can be identified by a "months-in-survey" figure giving the number of months that dwellings in the RG have been included in the survey.

44. Analysis will be done by comparing estimates from RGs using the new questionnaire to those from other RGs using the current questionnaire. The estimates will be made as comparable as possible by a benchmarking process which will ensure that each RG represents the same mix of individuals (by geographic region, age group and sex).

45. Let  $\tilde{y}_t^j$  be the estimated value of a key labour force variable for RG with months-in-survey  $j$  at time  $t$  and let  $Y_t$  be the corresponding population value at time  $t$ . Also let  $G(t)$  be the set containing the months-in-survey values for the RGs that are using the new questionnaire at time  $t$ . We shall assume that the effect of the new questionnaire is constant over the study period.

46. The following model at the RG level is used.

$$\begin{aligned}\tilde{y}_t^j &= Y_t + e_t^j, & j \notin G(t) \text{ (this RG is not using the new questionnaire), and} \\ &= Y_t + Q + e_t^j, & j \in G(t) \text{ (this RG is using the new questionnaire)}\end{aligned}\quad (1)$$

where  $Q$  is the new questionnaire effect and  $e_t^j$  is the sampling error for RG  $j$  at time  $t$  and  $E(e_t^j) = 0$ .

### 5.3 Level Estimate

47. Under the model (1), an estimate of the questionnaire effect can be made by comparing groups using the new questionnaire to groups using the existing questionnaire at a single point in time. This will be referred to as the level estimate. Let  $G(t)$  be as defined above, and let  $\#G(t)$  be the number of RGs in this set. In our case for the impact study of the new questionnaire  $\#G(t) = 1$  as there is only one RG using the new questionnaire at any given time. Then the level estimate is given by

$$\hat{Q}_{LEV,t} = \sum_{j \in G(t)} \tilde{y}_t^j - \frac{1}{8-1} \sum_{j \notin G(t)} \tilde{y}_t^j \quad (2)$$

48. The variance of the level estimator can be calculated since the variance-covariance structure of the RG estimates is known (derived in Bell (1998)).

### 5.4 Variance Of A Linear Combination Of RG Estimates

49. The composite estimator,  $\alpha' \tilde{y}$ , is a linear combination of the RG estimates from a number of months of data. Assume that  $L + 1$  months of data are to be used for the composite estimator and  $\tilde{y}$  and  $\alpha$  are the column vectors with elements  $\{\tilde{y}_{t-l}^j\}$ , as defined above and  $\{\alpha_{t-l}^j\}$  respectively for  $j=1, \dots, 8$  and  $l = 0, \dots, L$  where the elements of  $\alpha$  lie in the interval  $[-1, 1]$ . The general formula for the variance of a linear combination of the RG estimates mentioned previously is given by

$$\text{var}(\alpha' \tilde{y}) = \alpha' \text{var}(\tilde{y}) \alpha \quad (3)$$

50. As mentioned previously, the variance-covariance structure of the RG estimates is known. Consequently,  $\text{var}(\tilde{y})$ , the covariance matrix of the RG estimates can be estimated. This matrix has covariances between different rotation groups as zeros and strong covariances between successive months for the same rotation group. The expected value of a linear combination of the  $\tilde{y}_t^j$  under the model (1) is given by

$$E(\alpha' \tilde{y}) = \sum_{l=0}^L \{ Y_{t-l} (\sum_{j=1}^8 \alpha_{t-l}^j) + \sum_{j \in G(t-l)} Q \alpha_{t-l}^j \} \quad (4)$$

since  $E(e_t^j) = 0$ .

51. The constraints on the choice of  $\alpha$  for the linear combination  $\alpha' \tilde{y}$  to be an unbiased estimator of  $Q$  are

$$\sum_{j=1}^8 \alpha_{t-l}^j = 0 \quad \text{for } l = 0, \dots, L, \text{ and} \quad (5)$$

$$\sum_{l=0}^L \sum_{j \in G(t-l)} \alpha_{t-l}^j = 1 \quad (6)$$

52. Under these constraints,  $E(\alpha/\tilde{y})=Q$ , under the model with constant new questionnaire effect. The optimal composite estimator based on this data will be the choice of  $\alpha$  that minimises the variance (3) under these constraints.

53. Writing the constraints in the form  $C'\alpha = c$  and setting  $V = \text{var}(\tilde{y})$ , the optimal  $\alpha$  is given by  $\alpha^* = V^{-1}CD^{-1}c$ , for  $D^{-1}$  any generalised inverse of  $(C'V^{-1}C)$  (Rao, 1973). So the estimate of variance is

$$\sigma^2 = \text{var}(\hat{Q}) = \alpha^{*'} \text{var}(\tilde{y}) \alpha^* \quad (7)$$

## 6 Comparison Of Assessment Strategies

### 6.1 Estimating The Variances

54. It was mentioned previously that the variance-covariance structure of the RG estimates is known. The variances and sampling errors autocorrelations were estimated using RG estimates from January 1990 to July 1996. Separate RG estimates are obtained for each time in survey  $j$  and lag  $m$  and are based on Bell (1998) and Bell and Carolan (1998).

55. The work described in those papers can be summarised as follows. Using the covariance matrix of the RG estimates to estimate  $\text{var}(\tilde{y})$ , the variance estimate of the level estimator was calculated. Also, using the covariance matrix of RG estimates to estimate  $\text{var}(\tilde{y})$  and the constraints on  $\alpha$ , an optimisation program was written in SAS to find the optimal  $\alpha$ 's which would minimise the variance equation (3) for the composite estimator  $\alpha/\tilde{y}$ . Then using (7), the variance for the composite estimator for the questionnaire effect could be calculated.

### 6.2 Discussion Of Results

56. Table 1 shows the respective standard errors of the optimal composite estimators for each of the proposed testing strategies for varying number of months. Comparing the standard errors of the level and composite estimator shows the gains made in terms of the accuracy of the estimator by using composite estimation to calculate the difference between the two questionnaires. For a test of one month, the standard error for the composite estimator is half that of the level estimator. Further gains can also be made on the composite estimator by using a different eighth of the LFS sample each month. For a test conducted over 5 months, the standard error for the employment rate estimator using a rotating sample is less than half the standard error for when the same portion of the 'live' sample is used.

57. The improvement comes from utilising the additional movement in the estimates each month, i.e. a movement from the current questionnaire to the new questionnaire. If a composite estimator was used with a constant sample (strategy 2), this movement would only happen once (i.e. moving to the new questionnaire and then back to the old one). Utilising the movement that occurs when the portion of the sample using the new questionnaire returns to the existing questionnaire (in the last month of testing) achieves further improvements.

**Table 1: Standard errors (in % points) of estimators for split sample testing strategies**

Employment Rate	Level Estimator with fixed sample	Composite Estimator with fixed sample	Composite Estimator with rotating sample
Months = 1	0.64	0.32	0.32
Months = 2	0.6	0.29	0.22
Months = 3	0.58	0.28	0.17
Months = 4	0.56	0.27	0.15
Months = 5	0.55	0.27	0.13
Unemployment Rate			
Months = 1	0.53	0.35	0.35
Months = 2	0.48	0.3	0.24
Months = 3	0.45	0.29	0.19
Months = 4	0.43	0.28	0.16
Months = 5	0.41	0.28	0.15

### **6.3 The Optimal Strategy**

58. From the results discussed above, it can be seen that using a composite estimator to calculate differences in the two versions of the questionnaire can greatly improve the standard error of the estimate. Also, using a different portion of the sample each month to test the new questionnaire (as in strategy 3) further improves the standard error of the estimator.

59. It can be seen from Table 1 that the improvements in standard errors for all strategies between impact studies of four and five months are small. As it is necessary to conduct the study over a short time period and with a minimum of extra expenditure the benefits in extending the study beyond five months would be minimal. Hence, it was decided that the composite estimator with a rotating sample (Strategy 3) for five months duration would be the strategy to test the statistical impact of the new questionnaire. Under this strategy, difference of 0.32 % points in the employment rate can be detected at the 90 % confidence level with 80 % power.

60. It was assumed for the purpose of this analysis that the questionnaire effect will be constant throughout the duration of the testing of the new questionnaire. If it appears during the impact study of the new questionnaire that the effect is changing over time different constraints can be imposed to better model the effect (see Section 3.3 Bell (1998)).

61. One concern with using live sample to test the new questionnaire is that this may impact published Labour Force series over the period of the test. It is likely that any such effect would be small, being restricted to an eighth of the sample in any month. The nature of the changes to the questionnaire are such that there is not expected to be any gross changes in the survey estimates for the key labour force series. Furthermore,

any gross changes in the estimates could be expected to be picked up during the pilot tests and dress rehearsals, so appropriate action would be taken before the statistical impact study and final implementation of the new questionnaire.

62. Other concerns with the form of test proposed in strategy 3 relate to how well it corresponds to a live survey. It is possible that the effect of the new questionnaire is different for people who have previously reported using the old questionnaire, or for individuals interviewed for the first time using the new questionnaire. The most realistic test for these situations would involve a separate sample interviewed for a number of months using the new questionnaire. However, such a test would need to be very large, or it would only be able to detect large changes (being roughly as efficient as strategy 1 in the table). It is considered preferable to rely on the pilot testing and dress rehearsals to identify any large effects related to these issues.

## **7 Conclusion**

63. A large amount of evaluation, both qualitative and quantitative, is being conducted before a decision is made to introduce a new questionnaire to the Australian LFS. Different components of the evaluation are being conducted at various stages of the development process to ensure that a good questionnaire is developed and to detect whether this questionnaire will have an impact on the key estimates produced from the LFS. It is our view that the testing strategies used will provide ABS management with the information to decide whether to implement the new questionnaire, defer implementation pending further testing and assessment, or continue with the existing questionnaire.

## **8 Issues For Discussion**

- Is the overall testing strategy appropriate?
- Are there any other worthwhile analyses that could be undertaken using the pilot test data?
- Is the strategy for assessing statistical impact of the new questionnaire appropriate?
- Can the statistical design of estimation be improved?

## References

Australian Bureau of Statistics 1993, 'Information Paper: Questionnaires Used in the Labour Force Survey', ABS Cat. no. 6232.0, Canberra.

Australian Bureau of Statistics 1995, 'Information Paper: Measuring Employment and Unemployment', ABS Cat. no. 6279.0, Canberra.

Australian Bureau of Statistics 1997a, 'Information Paper: Labour Force Survey Sample Design', ABS Cat. no. 6269.0, Canberra.

Australian Bureau of Statistics 1997b, 'The Effect of Telephone Interviewing on Labour Force Estimates' in *Labour Force June 1997*, ABS Cat. no. 6203.0, pp1-7, Canberra

Bell, P. A. 1998, 'Using State Space Modelling and Composite Estimation to measure the effects of telephone interviewing on labour force estimates', in *Working papers in Econometrics and Applied Statistics*, ABS Cat. no. 1351.0, no.98/2, Canberra.

Bell, P. A. & Carolan, A. 1998, 'Trend estimation for small areas from a continuing survey with controlled sample overlap', in *Working papers in Econometrics and Applied Statistics*, ABS Cat. no. 1351.0, no.98/1, Canberra.

Rao, C.R. (1973) *Linear statistical inference and its applications*, second edition, John Wiley and Sons, USA