Journal of Official Statistics
Vol 4. No. 4. 1988. pp. 285–297

Accessibility by Telephone in Australia: Implications for Telephone Surveys

David Steel and Peter Boal[1]

Abstract: Telephone sampling methods such as random digit dialling (RDD) leave the population who do not own a telephone unrepresented. In this paper, data from the March 1986 Monthly Labour Force Survey and the 1984 Household Expenditure Survey are analysed to examine the characteristics of the population not accessible by telephone and the reasons for not having a telephone. The associated coverage biases are also analysed. Automatic interaction detection (AID) is used to help isolate those variables which are important in affecting accessibility by telephone and the subgroups of the population with relatively low telephone coverage. The effectiveness of post-stratification in reducing the coverage bias is also investigated.

Key words: Telephone surveys; telephone penetration; coverage bias; automatic interaction detection; post-stratification.

1. Introduction

Traditionally, surveys of the Australian population have mainly been conducted using face to face interviewing of a sample of people or households selected using area sampling methods. However, the use of telephones in surveys has increased considerably over the last few years, particularly by non-government survey organisations. Telephones can be used in household surveys in a variety of ways. In a continuing survey, field interviewing can be used the first time a household is included in the survey and telephones can be used for subsequent contact for those households with telephones, with personal interviewing still being used for the remaining households. This approach has been used for many years in the U.S. Current Population Survey and Canada's Labour Force Survey and is being investigated for possible use in the Australian Labour Force Survey.

For one-time surveys this method cannot be used. Moreover to avoid the costs associated with area sampling, telephone sampling either by random digit dialling (RDD) or selecting samples from lists of subscribers can be used. Either of these methods leave those people in households without telephones unrepresented, while using lists of subscribers also means that people with unlisted numbers are unrepresented. Dual frame, mixed mode designs can overcome many biases that arise from confining samples to the population accessible by telephone, but this approach reduces the cost savings yielded by use of

[1] Statistical Services Branch, Australian Bureau of Statistics, Belconnen, A.C.T., 2616, Australia. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the Australian Bureau of Statistics.

telephone surveys. (Groves and Lepkowski (1985) discuss efficient design of dual frame, mixed mode surveys.)

There have been a number of studies of the characteristics of people not accessible by telephone in the U.S.A. (Thornberry and Massey (1978), (1983), (1988), Smith (1987)). These studies show that telephone coverage is lower for households which are low income, rural, in rented accommodation, black, comprise one adult or a lot of adults, whose head is young, never married or divorced, of low education, or male. Jones (1982) uses data from the 1974/75 and 1975/76 Household Expenditure Surveys in Australia and reports data showing telephone ownership lower for households living in flats or rented dwellings and for those households with low income or headed by a young or unemployed person, or manual worker.

The coverage bias that would result from using an RDD survey has been examined by Thornberry and Massey (1978) for certain health characteristics. Whilst the health related characteristics of persons in non-telephone households are very different from those in telephone households, the biases that result from using the telephone population are often not great because of the high level of telephone ownership. Thornberry and Massey also investigate the effectiveness of post-stratification in reducing the bias.

In this paper more up to date data on telephone accessibility in Australia are analysed. In Section 2 the growth in telephone coverage is reported. Reasons for not having a telephone are given in Section 3. In Section 4 the characteristics of the households and persons not accessible by telephone are investigated. In Section 5 the automatic interaction detection (AID) technique is used to investigate those factors affecting telephone connection and those subgroups of the population with relatively low coverage. In Section 6 the coverage biases associated with telephone interviewing are examined and Section 7 investigates the effectiveness of post-stratification in reducing these biases.

2. Growth in Telephone Coverage

In March 1986 all households in private dwellings in the Monthly Labour Force Survey (MLFS) were asked if a telephone was connected, and if no telephone was connected the main reason for not having one.

The MLFS is a multi-stage area sample of households with personal visits by interviewers to all households selected in the survey. The telephone connection survey excluded households in caravan parks and special dwellings such as hotels, motels, hospitals, etc, which account for approximately 4 % of the population. The telephone connection component of the March 1986 MLFS had a response rate of 95.1 % resulting in a sample of approximately 32 000 responding households in private dwellings. Post-stratification by age and sex within each state's metropolitan and non-metropolitan areas is used in the estimation to reduce the potential bias due to non-response and non-contact.

The survey found that 91.3 % of private dwelling households had the telephone connected and these households accounted for 92.8 % of the persons in private dwellings. A similar survey conducted in March 1983 found 85.3 % of households connected covering 87.2 % of persons in such households.

Telephone connection was also collected in the 1975/76 and 1984 Household Expenditure Surveys (HESs). These surveys covered households in private dwellings and caravan parks with samples of 5 900 and 9 600 responding households respectively, but experienced lower levels of response, with 83.5 % of selected households being fully responding in 1984. Although subject to higher levels of non-response the HESs provide more detailed information. The percentage of house-

holds with a telephone connected was estimated at 62 % in 1975/76 and 87.5 % in 1984.

Taken together these four surveys show a steady growth of approximately 2.9 percentage points per year in the proportion of private dwelling households with a telephone. The current high level of telephone accessibility has led to more serious consideration of the use of RDD surveys than was possible with the lower levels of accessibility experienced in the 1970s.

3. Reasons for Not Having a Telephone

The March 1986 survey asked for the main reason a telephone was not connected, and these are given in Table 1. Cost was by far the most common reason given, although 49 % of people gave a reason other than cost.

Table 1. Reasons for non-connection: March 1986

Reason	% of non-connected households
Cost	51
Don't need one	23
Have use elsewhere	6.9
Rented short term accommodation	6.5
Awaiting connection	4.6
Other	7.3
	100.0

4. Characteristics of Non-Accessible Population

The telephone coverage rate for different types of persons and households as estimated from the March 1986 survey are given in Tables 2 and 3. From these tables we see that the rate is relatively low for unemployed persons, and in households in which the head is less than 25, or in which the head has never been married, or in households which have low income, or consists of a single parent family, persons living alone, two or more unrelated persons, or in which the household head is separated or divorced.

Table 2. Telephone coverage for persons: March 1986

Subgroup	Coverage rate %
Employed	94.6
Unemployed	82.2
Not in the labour force	91.5
Males	92.2
Females	93.5

A general rise in the telephone coverage rate with household income is present, as would be expected given that cost was the main reason for not having a telephone.

Table 3. Telephone coverage for households: March 1986

Subgroup	Coverage rate %
1 person HH[1]	83.4
2 person HH	92.0
3 person HH	92.2
4 person HH	95.0
5+ person HH	93.8
income <$A 100 per week	82.8
$A 101< income < $A 150	83.5
$A 151< income < $A 200	87.5
$A 201< income < $A 300	85.9
$A 301< income < $A 400	89.9
$A 401< income < $A 500	93.7
income > $A 500	96.6
some income as pensions or benefits	86.6
no income as pensions or benefits	93.7
Capital city	94.0
Rest of state	86.4
One family with married couple	94.8
One parent family	83.2
Two or more unrelated persons	84.1
Persons living alone	83.4
HOH[2] married	94.7
HOH separated or divorced	85.2
HOH widowed	91.0
HOH never married	77.2
HOH 15–24	72.6
HOH 25–44	92.3
HOH 45–54	93.5
HOH 55–64	93.0
HOH 65–74	92.3
HOH 75+	90.6

[1] HH = household.
[2] HOH = head of household.

More detailed tabulations are given in Australian Bureau of Statistics (1987) which reveal certain subgroups with lower coverage, namely households without a married couple whose head is unemployed (66.9 %), females living alone aged 15–24 (70.4 %), males living alone (73.2 %), especially those aged 15–24 (57.8 %), and households with head aged 15–24 (75.4 %).

These results show that although the overall coverage rate is high, the proportion of households not having a telephone can be large for particular subgroups of the population. We see that the coverage rate is low for young people, people not living in a family, especially those living alone, and for low income, unemployed persons.

Although the HES is not as recent as the March 1986 survey and found a lower telephone coverage rate, it allows us to analyse important variables not collected in the 1986 survey. In particular information is available on housing tenure, type of dwelling occupied, and principal source of income.

Table 4. Telephone coverage for households: 1984 HES

Subgroup	Coverage rate %
1 person HH[1]	81.0
2 person HH	88.9
3 person HH	87.1
4 person HH	90.9
5+ person HH	91.2
1st income decile	82.2
2nd income decile	81.4
3rd income decile	84.0
4th income decile	80.2
5th income decile	84.5
6th income decile	87.6
7th income decile	90.6
8th income decile	92.2
9th income decile	95.5
10th income decile	97.2
Living in a low rise flat	71.1
semi-detached or townhouse	86.7
high rise flat	89.6
separate house	90.8
Dwelling owned	93.4
being bought	94.8
rented from government	68.8
rented privately	69.8
occupied rent free	84.0
Principal source of income – government benefits	81.2
– wages and salary	88.8
– own business	94.0
– superannuation or investments	96.2

[1] HH = household.

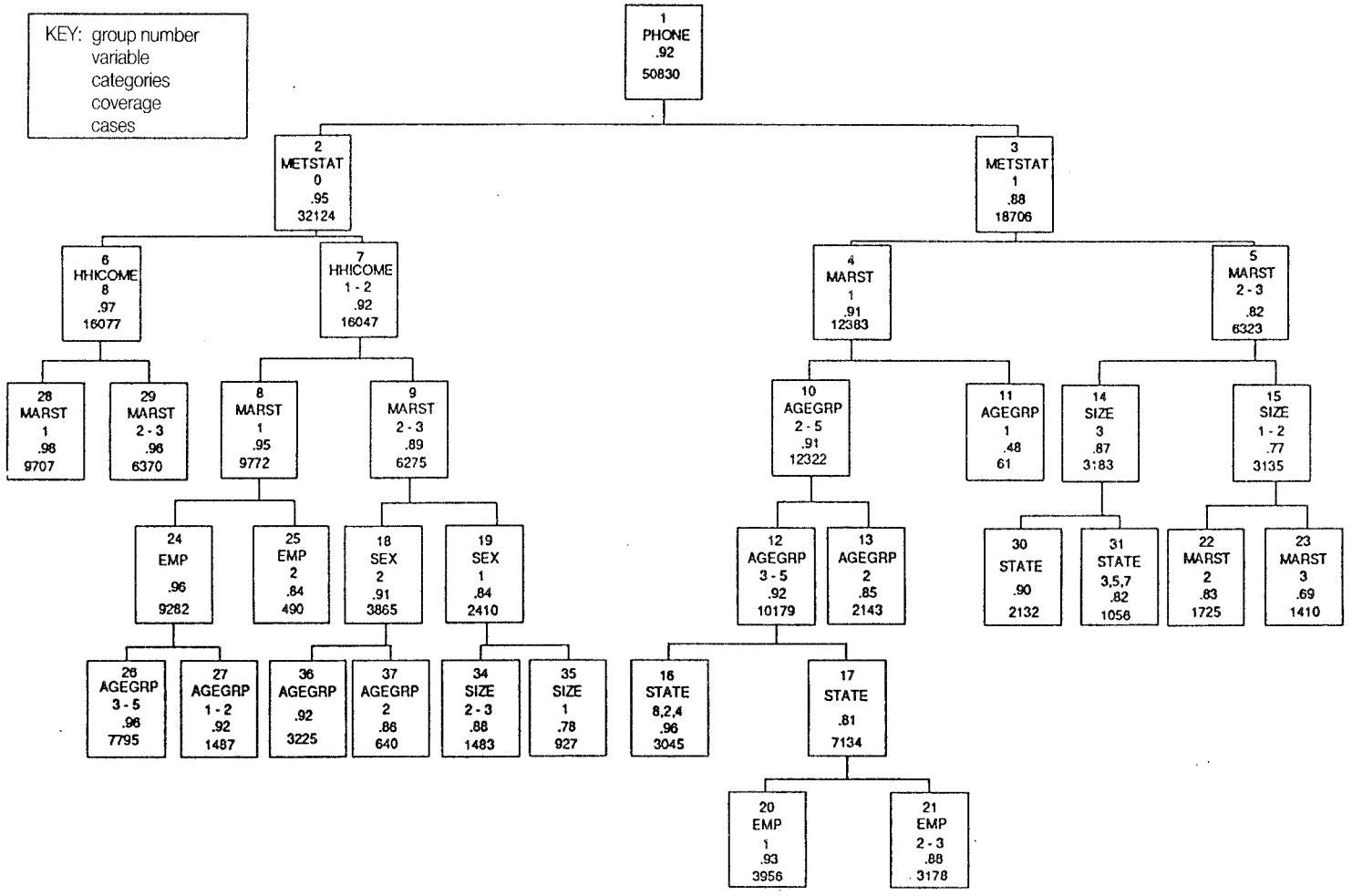
The coverage rates estimated from the 1984 HES are given in Table 4. The rate is relatively low for single person households or households living in a low rise flat or rented accommodation, or whose principal source of income is government benefits. Again, a general rise in coverage rate with household income is present. More detailed information is given in Sparks (1986).

The picture obtained from the March 1986 survey and the 1984 HES is in the main consistent with the results of studies in the USA. These results cast doubt on the suitability of telephone sampling for surveys for which young people, people of low socio-economic status, or in rented accommodation are important.

5. Factors Related to Telephone Accessibility

To gain a better understanding of the variables related to telephone ownership and to help isolate groups in the population with low coverage rate the automatic interaction detection (AID) technique was used. AID is a binary segmentation technique which splits the population under study into two groups so that they are as different as possible with respect to the dependent variable and as internally homogenous as possible. The splitting variable and the split are chosen so that the ratio of the between group sum of squares to the total sum of squares is maximised. The process of segmentation is then continued until the

Fig. 1. March 1986: Telephone coverage AID analysis



groups are too small to split further, the difference in telephone coverage rate obtained by further splitting is not important or until a specified maximum number of unsplit groups is reached. At each step, the next group to be split is chosen on the basis of the sums of squares within the group. More details of the AID procedure can be found in, among others, Fielding and O'Muircheartaigh (1977).

The AID technique was applied to the March 1986 survey data with the dependent variable defined as whether or not a person lived in a household with telephone service. The variables included in the analysis were: age, sex, marital status, state and part of state of residence (capital city/rest of state), employment status, and household income and size. (Category definitions are given in Appendix 2.) Because of problems with computer capacity a random subsample (80 % of the total sample) was used in the analysis.

The analysis was continued until 30 unsplit groups were formed, however, the results down to 20 unsplit groups reveal the main features and are shown in Fig. 1.

The analysis first splits the population into metropolitan areas (i.e., the capital cities of the six states in Australia) and non-metropolitan areas. Within metropolitan areas, household income and marital status appear important. In non-metropolitan areas marital status, age, and household size are important.

The people with average or above average coverage rates in metropolitan areas were mainly those living in households with a weekly income exceeding 500 Australian dollars ($A) or married persons in lower income households. In the non-metropolitan areas, the groups with average or above average telephone coverage were generally those who were married (except for married teenagers).

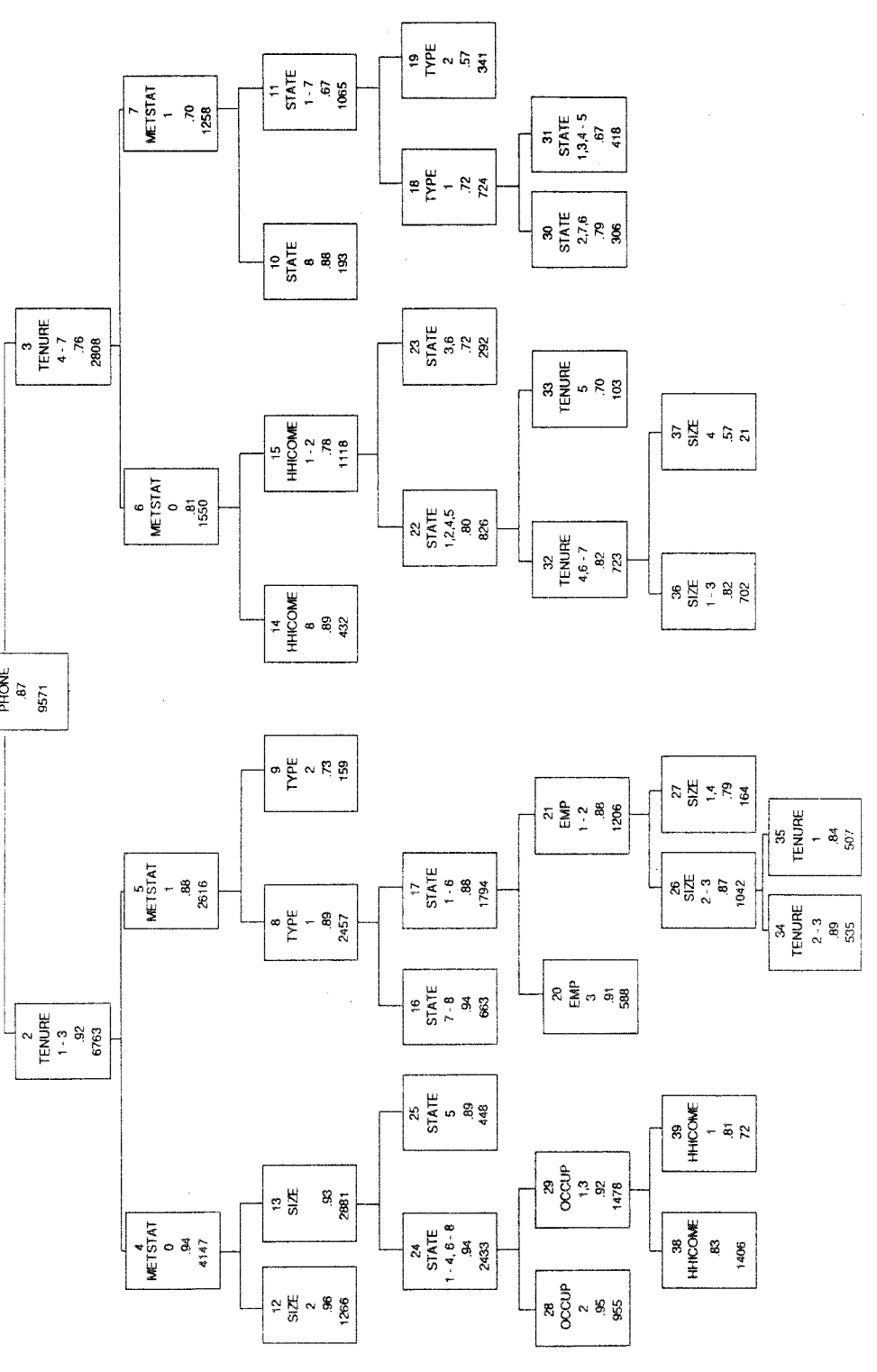
What emerges from this analysis is the importance of income and marital status in metropolitan areas and marital status, household size, and age in non-metropolitan areas in determining telephone coverage.

The AID analysis is useful in identifying groups in the population with relatively low telephone coverage which are not evident from the analysis in Section 4. From Fig. 1 we can see that in metropolitan areas, unmarried males in households with weekly income less than $A 500 have low coverage rates (84 %) as well as married, unemployed persons living in households with weekly incomes of less than $A 500 (84 %). In the non-metropolitan areas, unmarried persons living in households of one or two persons (77 %) have low coverage rates.

Carrying the segmentation on to 30 unsplit groups isolates subgroups in the population with even lower coverage rates. For example, in metropolitan areas unmarried males, living alone and with weekly income less than $A 500 have a coverage of 78 %. In the non-metropolitan areas, never married persons, living in households of one or two persons with a weekly household income of less than $A 500, who are not living in Victoria or the Australian Capital Territory (ACT) have a coverage rate of 55 %. While these subgroups may be small, the analysis shows there are particular groups of people who would be greatly under-represented in a telephone based sample. Clearly, telephone sampling is unsuitable for surveys for which young, unmarried or low income people are important.

The AID technique was applied to the 1984 HES household data. The variables included in the AID analysis were household size and income, dwelling type, age group, employment status and occupation of the household head, state and part of state, and housing tenure. (Category definitions are given in Appendix 2.)

Fig. 2. 1984 HES: Telephone coverage AID analysis



The results of the analysis stopping at 20 unsplit groups are shown in Fig. 2. Housing tenure defines the first split with the coverage rate being low for households in rented accommodation (76 %). Both tenure groups are then split by the variable "part of state." The coverage rate is average or above average for households in non-rented dwellings and for households in rented dwellings in metropolitan areas with weekly incomes of $A 500 or more.

Again subgroups with significantly lower coverage can be identified such as households in metropolitan areas, in rented accommodation with weekly income less than $A 500 in Queensland and Western Australia (72 %). In non-metropolitan areas, the coverage is low in non-rented flats, semi-detached houses, townhouses (73 %), and in rented flats, semi-detached houses and townhouses (outside the ACT) (57 %). Carrying the segmentation on to 30 unsplit groups reveals that households in non-metropolitan areas, occupying a rented flat, semi-detached house, or townhouse, with weekly income less than $A 500 and headed by a person younger than age 40 have low coverage (44 %).

A similar analysis of the 1975/76 HES reported by Jones (1982) showed housing tenure to be an important variable, but also found that the occupation and age of the head of the household were important. In our analysis, occupation first appears in the analysis in the formation of the 28th and 29th groups and age in the formation of the 40th and 41st groups. Our analysis includes "part of state," but the difference in the results is most probably due to the overall level of telephone coverage increasing from 62 % to 87.5 % between the two surveys.

Because of the inclusion of different variables, in particular the inclusion of "housing tenure" and "dwelling type" in the HES analysis, the results of the AID analyses of the March 1986 and HES data differ in some respects. Both analyses reflect the difference in coverage rates between metropolitan and non-metropolitan areas and the influence of household income and size. We believe the appearance of marital status in the analysis of the March 1986 data is probably due to household income, combined with household size, being related to the housing tenure and dwelling type, which appear as important variables in the analysis of the HES data.

6. Biases in Telephone Surveys Due to Non-Coverage

The coverage rates give us some idea of the types of people and households likely to be under-represented in a telephone survey using RDD. An important issue is the biases that are likely to result because of less than 100 % telephone coverage. In Table 5, the mean of several variables are shown for the population potentially accessible by telephone ($\bar{Y}_T$), the population not accessible by telephone ($\bar{Y}_N$) and the mean for the whole population as estimated from the survey. The relative bias in using the mean of the telephone accessible population is also given.

The relative bias for the unemployment rate is large at –12.2 %. The estimate of unemployment rate currently obtained from the MLFS has a relative standard error of 0.4 %. Even the relative bias of 1.1 % for the labour force participation rate is significant compared with the 0.2 % relative standard error on the estimate obtained from the MLFS.

The importance of the biases depends on the overall reliability required. For cases where, because of budget limitations, a small sample is to be used giving large standard error, biases of the order of 5 % to 10 % may be acceptable. Furthermore the lower cost of RDD surveys compared with area sampling methods may still mean that in terms of overall mean squared error an RDD survey is better.

Table 5. Means of variables for telephone and non-telephone populations: March 1986

Variable	$\bar{Y}_T$ %	$\bar{Y}_N$ %	$\bar{Y}$ %	Relative bias %
Unemployment rate	7.26	22.9	8.27	–12.2
Labour force participation rate	64.4	57.1	63.7	+1.1
Single person HH[1]	17.2	55.2	18.9	–9.0
Married couple in HH	70.0	39.0	66.4	+5.5
Single parent HH	4.8	10.0	5.2	–7.7
HH income < $A 150 per week	12.7	26.8	13.9	–8.9
HH income > $A 300 per week	60.5	37.0	58.4	+3.6
HH income > $A 500 per week	35.6	12.9	33.7	+5.6
HH with some pension or benefit income	27.6	46.2	29.2	–5.5
HH with head > 55 years	29.5	25.3	29.0	+1.7
Health insurance	53.7	22.8	51.3	–4.7
Sex (female=0, male=1) proportion	0.49	0.54	0.49	+0.6

[1] HH = household.

7. Reducing Biases by Post-Stratification

Post-stratification is a technique that can be used to reduce the coverage bias associated with telephone sampling. Thornberry and Massey (1978) report some success in using post-stratification by income, region, and race to reduce bias in estimates of health status characteristics. Jones (1982) also suggests post-stratification as a means of reducing the coverage bias.

For any particular characteristic of interest, post-stratification by some set of variables will eliminate the coverage bias if, conditional on the post-stratification variables, the telephone and non-telephone populations have the same mean. Because most surveys collect a large number of variables we would like to find post-stratification variables which eliminate the bias for a large range of variables. This would be possible if we could determine the full set of variables which affect telephone accessibility.

The AID analyses have shown the influence of a variety of variables. We performed a logistic regression using telephone accessibility as the dependent variable that showed the need to include most variables considered and a large number of interaction terms in order to obtain a satisfactory model explaining telephone accessibility. In practice, the availability of post-strata population estimates limits the number of variables that can be used. Moreover, because of the need to have a sufficient sample size in each post-stratum there is a limit to the number of poststrata that can be formed. Taking these factors into account a range of post-stratifications were tried.

Table 6 shows the effects of using a range of post-stratifications for estimates of unemployment rate, labour force participation rate, and health insurance rate. Because of the strong metropolitan/non-metropolitan effect found in the AID analysis, this split is used within each state in all cases. Post-stratification does not always reduce bias mainly because it is affecting the numerator and denominator used in calculating the rate. Generally we see that post-stratification by age and sex, (the traditional post-stratification variables) has little effect. Using household income does

Table 6. Effect of post-stratification

Post-stratification variables[1]	Unemployment rate	Participation rate	Health insurance rate
Full sample	8.45	63.2	51.3
Telephone sample	7.43	63.8	53.7
Age Sex	7.48	64.0	53.6
Household income	7.61	63.1	53.2
Age Sex HH[2] income	7.81	63.6	53.1
Age Sex HH income, HH size	7.80	63.6	53.0
Age Sex HH size, Marital status	7.53	64.0	53.2

[1] A metropolitan/non-metropolitan post-stratification within each state is used in all cases.
[2] HH = Household.

offer some reduction in bias and further improvement is obtained by adding age and sex. Adding household size to the post-stratification offers only marginal improvement and would not be worth the extra complexity. Bearing in mind the need to keep the number of post-stratification variables to a reasonable number the age-sex-income combination seems the most beneficial.

The results in Table 6 also show that while the biases can be reduced, they cannot be eliminated. In the case of the unemployment rate the relative bias of –12.2 % is reduced to –7.6 % using the age-sex-income option. For the labour force participation rate and the health insurance rate the bias is reduced by less than a half.

8. Discussion

The results in this paper show that the suitability of telephone sampling methods will depend on the variables being collected and the groups in the population important to the survey. Coverage rates are such that surveys in which low income, young people, or people in rented accommodations are important should not use telephone samples. Telephone sampling should also not be used for estimating variables which have significantly different means in these low coverage groups, for example, unemployment. Moreover, the traditional way of trying to reduce the coverage bias, post-stratification, while having some beneficial effect, still leaves some significant biases.

9. References

Australian Bureau of Statistics (1987): Household Telephone Connections Australia, March 1986, Catalogue No. 4110.0.

Fielding, A. and O'Muircheartaigh, C.A. (1977): Binary Segmentation in Survey Analysis with Particular Reference to AID. The Statistician, 26, (1), pp. 17–28.

Groves, R.M. and Lepkowski, J.M. (1985): Dual Frame, Mixed Mode Survey Design. Journal of Official Statistics, 1, (3), pp. 263–286.

Jones, R. (1982): Variation in Household Telephone Access: Implications for Telephone Surveys. Australian Journal of Statistics, 24, (1), pp. 18–32.

Smith, T.W. (1987): Phone Home? An Analysis of Household Telephone Ownership. Paper presented at the International Conference on Telephone Survey Methodology, Charlotte, November, 1987.

Sparks, M. (1986): Access to Telephones in Australian Households: An Overview. Statistical Methods Section, Information Paper No. 3, Australian Bureau of Statistics.

Thornberry, O.T. and Massey, J.T. (1978): Correcting for Undercoverage Bias in Random Digit Dialed National Health Surveys. American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 224–229.

Thornberry, O.T. and Massey, J.T. (1983): Coverage and Response in Random Digit Dialed National Surveys. American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 654–659.

Thornberry, O.T. and Massey, J.T. (1988): Trends in U.S. Telephone Coverage Across Time and Subgroups. In Telephone Survey Methodology, edited by R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, and J. Waksberg, pp. 25–49, Wiley, New York.

Appendix 1

Limitations of Analysis

The MLFS uses different probabilities of selection in different states, ranging from 1 in 60 in Tasmania to 1 in 200 in New South Wales and Victoria. In the HES different probabilities of selection are used in different states and within the metropolitan and non-metropolitan areas of the states. These differences in selection probabilities were ignored in the AID analyses and unweighted data was used. However since "part of state" was included in the analysis the effect of the differing probabilities of selection should be minimal.

The response rate to the telephone connection component of the March 1986 MLFS is reasonably high, but since the proportion of households estimated without a telephone is 8.7 % the 4.9 % loss of sample due to non-response and non-contact may affect the results especially if the sorts of households that are difficult to obtain responses from in the MLFS also have low telephone coverage. For example, if the coverage rate of the non-respondents was only 50 %, the true coverage rate for the whole population would be 89.3 % instead of the 91.3 % estimated. Any effect due to non-response may be significant for particular subgroups. The response rate for the HES is lower than that of the MLFS so these problems may also affect the results from that survey.

Received December 1987
Revised February 1989

Appendix 2

Categories, Codes, and Labels Used in AID Analysis

Variable (Label)	Code	Categories
Part of state (METSTAT)	0	Metropolitan (i.e., State Capital, City)
	1	Non-metropolitan
Age group (AGEGRP)	1	15–19 years of age
	2	20–29
	3	30–39
	4	40–59
	5	60+
Household income (HHINCOME)	1	$A 0 –249 per week
	2	$A 250–499 per week
	8	$A 500+ per week
Household size (SIZE)	1	1 person
	2	2 persons
	3	3–5 persons
	4	6+ persons
Marital status (MARST)	1	Married
	2	Widowed, divorced or separated
	3	Single
Employment status (EMP)	1	Employed
	2	Unemployed
	3	Not in labour force
Dwelling type (TYPE)	1	Separate house
	2	Other (including flat, semi-detached or townhouse)
Housing tenure (TENURE)	1	Owned outright
	2	Buying
	3	Renting, government furnished
	4	Renting, government unfurnished
	5	Renting, non-government furnished
	6	Renting, non-government unfurnished
	7	Occupied rent free
Occupation (OCCUP)	1	Professional, managerial
	2	White collar
	3	Other
Sex (SEX)	1	Male
	2	Female
State (STATE)	1	New South Wales
	2	Victoria
	3	Queensland
	4	South Australia
	5	Western Australia
	6	Tasmania
	7	Northern Territory
	8	Australian Capital Territory