

## CHAPTER 2

# AUSTRALIAN HOUSEHOLD EXPENDITURE SURVEYS

*Readers will uncover how the Australian Bureau of Statistics (ABS) collects various versions of the Household Expenditure Survey (HES) data. They will also learn how these data are presented, their limitations and uses for academic and policy analyses, and how to improve the quality of the HES data for future social policy research.*

### 1. INTRODUCTION

The Australian Household Expenditure Survey (HES) is a multi-purpose socio-economic enquiry of the nation conducted by the Australian Bureau of Statistics (ABS) in the form of successive 'series'. Each series took approximately a year for data collection. The enquiry on consumer expenditure has been a regular of all the HES series, beginning with the first series, which was conducted during July 1974 to June 1975 and confined to capital cities only. Later HESs covered the whole of Australia (except remote and sparsely populated areas).

These HESs were designed to find out how the expenditure patterns of private households vary according to different income levels and characteristics of households. Most of the information was collected from households on a recall basis with a particular reference period (which varied according to the type of expenditure) using interview techniques. In addition, all members of households aged 15 years and over were requested to record all 'payments made' over a two to four week period in a diary provided to each of them. The stratified multi-stage probability (proportional to the households and collector's districts) sampling procedure was followed for selecting the households, who were interviewed evenly throughout the survey year. Any expenditure made by members of the selected households for business purposes were not considered in these surveys. These surveys collected 'household expenditure', which included expenditures on those goods (both durable and non-durable) and services made by the members of the selected households for private consumption. Other components of household expenditure such as income tax, superannuation contributions, life insurance premiums, purchases of and deposits on dwellings and land are classified as 'other payments'.

Estimates for the broad (15 groups) and medium (120 major items) expenditure items were available from the beginning of the HES series by broad weekly

household income and all households for the capital cities only. The estimates for urban and rural sectors were always distinguished and (see point 6 on page vii of the ABS catalogue No. 6516.0) became available from the 1975-76 HES in different broad income groups for different states and territories and for the whole of Australia. More importantly, the ABS began to release 'confidentialised' Unit Record Tapes for public use from the 1984 HES. Detailed tabulations covering various expenditure items can be found from the ABS publications for various versions of the HES. The concepts, definitions, and procedures remained more or less the same over the various versions of the HES, with some noticeable exceptions discussed in the following section.

Many important academic and policy research works are based on the HES data. Among many others, the following are some important and useful works based on these survey data.

- \* Studies on the functional forms of Engel curves by Haque (1988, 1989a, 1989b, 1996), Bewley (1982) and Hoa, Ironmonger and Manning (1983).
- \* Studies on income distribution by Yates (1991), Kakwani (1986), Houghton (1988), Harding (1984, 1995, 1997), Borland and Wilkins (1996), Borland (1998), Page and Simmons (2000) and NATSEM (1998, 2000).
- \* Studies on taxes and social welfare benefits on various classes of the population by Warren (1986, 1991), Dixon, Foster and Gallagher (1985), Agrawal (1989), Saunders (1987, 1992, 1994, 1997, 1998), Castles (1987), Australian Bureau of Statistics [ABS] (1987), and Mitchell, Harding and Gruen (1994).
- \* Studies on the cost of children by McDonald (1989, 1993), Bradbury (1989a) and Edgar (1989).
- \* Studies on patterns of household expenditure by Bradbury (1996), Haque (1984, 2000, 2001), Perkins (1991), Newell (1990), Whiteford (1991), Powles, Hage and Cosgrove (1990), and Newell, Ham and Coady (1987).

Despite many problems, researchers, planners and policy makers are using the HES data increasingly. It is necessary to continually examine the validity of the data and to take all possible steps for improving the usefulness of the data for planning and policy-making purposes at regional and national levels, and also for certain target groups such as aborigines and disadvantaged groups.

In this chapter, we make some observations on the limitations of HES data based on our experience, processing, tabulation and analysis of such data for various investigations. Among other things, we emphasise: (a) the need for field as well as technical studies, which examine the validity of the data being collected; (b) the need for studies on the inter-temporal comparability of the HES data from different versions; and (c) the need for timely release and/or publication of these data in a more meaningful way for some broad groups of consumption items.

## 2. CHANGES IN DATA COLLECTION PROCEDURES FOR VARIOUS VERSIONS

The basic data collection methodology remained approximately the same over the various versions of the HES. However, a few changes were made from time to time in the procedure of collecting HES data. Some important changes are mentioned below.

- (1) The geographical coverage of the 1974-75 HES was confined to the six state capital cities and Canberra only. However, the subsequent surveys covered the entire country (both urban and rural areas) except for remote and sparsely populated areas.
- (2) In the first series of the HES, any usual member of the household (including the head) absent at the time of interview and not returning within 7 days was excluded from the survey, whereas in the later series if the head of the household was absent at the time of interview and was expected to return after 7 days, but within 6 weeks of that date then the person was included as a member of the household. Income and expenditure data were collected for absent heads by a recall method without any diary information relating to the household head.
- (3) The diary-keeping period was 'two weeks' in the first series. During the second and third series the length of the diary-keeping period was two weeks for urban households and four weeks for rural households. But, in the 1988-89 HES the diary keeping period was two weeks for all households (both urban and rural).
- (4) In the first survey the value (retail price) of only free goods and services obtained from an employer was included in the 'income-in-kind' component. However, this concept was extended to cover goods and services provided by an employer at reduced prices and free of charge in the later surveys. In addition, from the second series, households were requested to keep records of all details of any product taken from their back yard and/or vegetable gardens in the supplied diary.
- (5) In the first series of HES any income earned by a member of the household aged less than 15 years was excluded from the family income. However, children's income was included in the family income from the second series.
- (6) The different treatment and incidence of non-monetary income commenced from the third series. Income in kind received from an employer is counted in the 1984 HES as both income and expenditure of the household concerned. However, the items are restricted in the income questionnaire to cars, housing, electricity and telephones, and to any goods obtained during the diary-keeping period. The respondent is asked for their current retail value (and the amount if any that they cost him). Unlike previous series only the last financial year's income from investment and/or self-employment is considered from the third series.

- (7) Negative income (in the case of business loss) was collected to estimate the total household income from the 1988-89 HES. This income was treated as 'zero' in previous surveys.
- (8) In order to estimate household expenditure, the first two series adopted a mixture of the 'payments (i.e. payments made during the reference period for goods and services, whether or not acquired or consumed during that period)' and 'acquisitions (i.e. the cost of those items acquired during the period, whether or not fully consumed or paid for)' approaches. However, only the acquisition approach was followed from the 1984 HES.
- (9) Important changes in categorisation in the HES Commodity Code List (HESCCCL) were made between the first two series and other series: separate identification of current housing costs for selected dwelling (item code 01); allocation of the principal component of mortgage payments to form part of 'other capital housing costs' (item code 16); and collapsing and splitting of previously established expenditure items. The number of expenditure items in different surveys is given in Table 2.1. This table shows that more detailed expenditure items are collected in recent HESs than earlier HESs, but major expenditure items remain approximately the same in various HESs except the 1998-99 HES when 609 detailed consumption items were collapsed into 123 major items; while total number of broad expenditure items remained same in Australia in various HESs.

*Table 2.1. Number of expenditure items in various HES in Australia: 1974-1975 to 1998-1999*

<i>Description</i>	<i>1974-75</i>	<i>1975-76</i>	<i>1984</i>	<i>1988-89</i>	<i>1993-94</i>	<i>1998-99</i>
Detailed items	300	300	440	430	626	609
Major items	102	90	100	100	99	123
Broad items	15	17	17	17	17	17

Source: Australian Household Expenditure Surveys: 1974-1975 to 1998-1999.

- (10) The Australian Standard Classification of Occupations (ASCO) has replaced the Classification and Classified List of Occupations (CCLO) since the 1988-1989 HES.
- (11) A three-month recall period was used for some infrequently purchased items such as refrigerators, washing machines, etc., since the third series of the HES, while a two-week or four-week recall period was used for such items in earlier versions.
- (12) Different sample sizes of households were used in different series of the HES. These are given in Table 2.2 below. This table shows that sample size of households vary from HES to HES.

Table 2.2. Number of households sampled in various HESs in Australia: 1974-75 to 1998-1999

<i>Year of survey</i>	<i>Sample size (number of households)</i>
1974-75	9095
1975-76	5869
1984	9571
1988-89	7405
1993-94	8389
1998-99	6893

Source: Australian Household Expenditure Surveys: 1974-1975 to 1998-1999.

As a result, the estimates of the change in expenditure between different series of HESs are subject to various degrees of sampling and non-sampling errors. Standard errors for some variables can differ significantly from version to version.

- (13) Financial years were chosen as a survey period for the first two series of the HES. However, a calendar year was chosen for the third series, 1984 HES. The data collection period for the fourth series was from July 1988 to July 1989.
- (14) The main differences of the 1993-1994 HES from the previous HESs are as follows.
  - (a) The survey weighting process for the 1993-1994 HES used independent estimates of the number of households in Australia as benchmarks.
  - (b) The number of households contributing to the 1993-1994 HES (approximately 900) increased, replacing some missing items by its imputed values calculated on the basis of information reported for similar households.
  - (c) Estimated income tax was calculated for all households according to the taxation criteria for 1993-94, and using the income and other characteristics of household members, as reported in the survey.
- (15) New techniques were introduced to collect the 1998-99 HES data. The main differences of the 1998-99 HES from that of the previous HESs are given below.
  - (a) A computer assisted interviewing (CAI) technique was used to collect data from households and individuals through a written diary.
  - (b) More household and individual estimates were used as the weighting process for the 1998-99 HES.
  - (c) Extra detail was collected for characteristics such as mobile phones, taxes, childcare, education, gambling, and income-in-kind through improved methods.

- (d) New questions on financial stress and lump sum payments were included, and the definition of a dependent child changed. The new definition of a dependent child includes people 15 years and under, and full-time students aged between 15 and 24 years (instead of 15 to 20 years, as used previously), who have a parent in the household, but no child of their own.
- (e) A Household Expenditure Classification was introduced instead of the HES Commodity Code List (HESCCL) as used previously.

These changes were made over various versions of the HES with a view to improve the quality of data. The idea behind all these changes was to obtain a more comprehensive picture of the economic activities of the households. However, there is a serious danger in interpreting and analysing such data, particularly for intertemporal comparisons among various versions of the HES data. Intertemporal comparability of household income and expenditure data is very important and useful to understand the performance of the economy, and the standard and quality of living of households. The continuing changes in the data collection procedures make it difficult to fulfill such objectives.

### 3. VALIDITY OF THE HES DATA

The HES data are based on recorded information in a diary (during a two to four weeks period depending on the series and locality of the responding households) as well as answers provided by a responding household to questions which were asked about their consumption of various goods and services during a certain period of time (using a recall period method). There could be random errors and biases in the data entries due to recall periods and wilful distortions. Also there is a possibility that the keeping of diaries results in behavioural change of the responding households. There is evidence that the respondents have a tendency to over-state some items of consumption and under-state the consumption of others. The diary keeping method is a good way of collecting day-to-day expenditure data. Yet, there is no way of getting over the response biases completely. Also, there is no evidence that a shorter reference period would necessarily ensure greater accuracy of the data.

From the third series, the ABS has adopted the acquisition approach which means the full cost of goods and services is recorded at the time of acquisition (not at the time of consumption) with a reference period of time, two weeks for diary keeping, and three months to two years for infrequent purchases depending on the item. Average weekly expenditures on various goods and services were then recorded on the basis of being evenly spread over the reference period. This creates a problem, because some households may not spend anything on certain items during the reference period or diary-keeping period at the time the survey was conducted. For example, the 1988-89 HES showed that 19% and 22% of households did not purchase anything on clothing and furniture respectively during the survey period. On the contrary, some households purchased more than usual during the survey period. As a result, in some cases their spending may well exceed their

income. The 1988-89 HES Unit Record Tape showed that 57% of households spent more than their net income over the reference period.

Experience suggests that as time passes, more and more Australian families are purchasing bulk household commodities including some food items such as flour, rice, meat, and seafood etc. As a result, they purchase household commodities at less frequent intervals. This could probably happen due to widespread use of cars, freezers etc., because of changing life-styles as well as to save time due to the greater participation rate in employment by couples. The Australian Institute of Family Studies (AIFS) Bulletin (no. 6, p. 17) shows that the employment participation for couples with children increased to 53.2% in July 1988 from 42.8% in July 1979; the corresponding figures for couples without children are 61.8% and 50.7% respectively.

The Australian Bureau of Statistics collects expenditure data with the main purpose of providing weights for the Consumer Price Index (CPI). However, data on quantity (in physical terms), which are very important for nutritional analysis, are not collected. Also, the availability of both quantity and expenditure information would provide a valuable crosscheck on the reliability of the information. By computing the price as a ratio of these two, it is possible to eliminate extreme observations due to errors of recording. It is thus hoped that the ABS will collect both quantity and expenditure information in future HESs.

Because quantities are not collected, it is not possible to estimate quality elasticity, an important indicator of the quality of living of the households. Quantity data can be obtained by dividing the expenditure estimates (value term) by consumer price indices. However, the consumer price indices available in the country produced by the ABS are often found to be inadequate for such purposes, as they relate to particular capital cities and to particular commodities. In addition, there is a problem of inter-temporal and inter-regional differentials in price variation. Hence, one should have different sets of consumer price indices for different income groups of the population. This is important for examining the expenditure pattern of households in real terms, because Haque (1991c) has shown that there is a clear time trend in average weekly household income and expenditure of the population in different income groups in different states.

We now come to the effect of seasonality on HES data. Spreading interviews evenly over the 12 months of the survey year eliminates the effect of seasonality almost completely as far as the averages of expenditures based on the entire survey year are concerned. However, as different households are interviewed at different points in time, seasonal variation is superimposed on the true variation between households, and the distribution of population by size-classes of household income exaggerates the true extent of the inequality. More seriously, seasonality may be distorting the estimates of the Engel functions and Engel elasticity based on the HES data.

Many analyses based on HES data assume that the biases (if any) in the data are uniform over time. This may not be true even under the most favourable situations, because the biases may change abruptly in some unusual situations, for example when prices change very sharply. Also, biases may fall or rise systematically through time. However, one might assume that estimates represent consumption

expenditures obtained from large samples. But, in the case of the HES data, the situation was weakened by the changes in the data collection procedure mentioned in the preceding section.

One of the major weaknesses of HES data may be the 'unduly high' estimate of per capita intake of 'meals out and take-away food', and alcohol and non-alcohol beverages in the highest quintile income group, as the level of consumption of such items in physical terms is one of the better indicators of living standards in Australia. It is unfortunate that this problem has not been given the attention it deserves. Variations in consumption standards may be reflected more or less consistently over time and across regions. However, the absolute levels of these estimates are quite doubtful and should be checked more thoroughly. 'Ceremonial parties' may provide a partial explanation. The host households report the entire quantum of meals in hotels and restaurants, drinks and alcohol needed for the function in its budget, while most, if not all, of the invitees may forget that they had a meal outside when reporting for their respective households. The same difficulties arise when one is offered drinks or biscuits, etc., while visiting friend's or relative's houses on special occasions like Christmas Day or Mother's Day. Haque (1989a, 1989b) has observed that the Engel elasticity obtained for meals out and take-away food is high and sensitive.

Another criticism of the HES data may be the possible under-enumeration of the highest and lowest income bracket households. The HES excludes inmates of hospitals, jails, military cantonments, etc., and also is possibly under-enumerating the homeless and destitute people. Although, there is no clear evidence of such a claim, it is desirable to keep a systematic record of approximate characteristics of these households. It would be excellent if a representative sample of these casualty households could be pursued and covered for the consumer expenditure enquiry.

On the whole, the HES income data appear to agree fairly well with the corresponding figure derived from other household based surveys, such as the Survey of Employed Wage and Salary Earners, the Income Distribution Survey, the Housing Survey, etc. This agreement is quite remarkable considering the completely different data sets. However, when the HES and the Australian Taxation Office (ATO) data are compared, it appears that incomes have been significantly underestimated in the HES data. In a recent study, the Australian Bureau of Statistics (1991) has shown that the average annual per capita income for those individuals with taxable income over the tax threshold based on the 1988-89 HES was \$20,007, compared to \$24,374 based on the 1988-89 ATO estimates for those individuals who lodged tax returns. This obviously does not guarantee that the ATO income data are correct and should be preferred over HES income data. Substantial research work should be carried out to justify the superiority of the ATO income data over the HES income data.

As regards the pattern of household expenditure, the picture emerging from the HES differs systematically from that based on the private final consumption expenditure of the Australian National Accounts (ANA) estimates. These differences are quite reasonable on the grounds of the significant differences in scope, concepts, sources and methods used for these two data sets. The ANA estimate is designed to provide an aggregate of the diverse transactions in the

economy, and the various components (viz. household sector) involved in those transactions. While, the HES deals with the levels of private expenditure of households in a narrowly defined population, one still does not know which of the two estimates approximates more closely to reality. By and large, the HES estimates are lower than the ANA estimates across commodity groups.

In view of the importance of the HES data, it is indeed surprising how little technical work has been done in the country to cross-examine the data to improve its quality. For example, there is no research to justify the 'accusation method' over the 'payment method', which has been adopted since the third version of the HES. It would be expected that a sample survey organisation would carry out various types of field studies before embarking on a regular large-scale survey and also to publish the results of such studies. This has hardly been done so far. As a result, there have been a very few field studies, which throw light on or justify the procedures followed in the HES. It is possible that sudden changes in the questionnaire and data collection procedures introduced a break in the time series of HES estimates and increased discrepancies between the HES estimates and the Official National Income and Expenditure estimates. In the absence of appropriate field trials, one cannot easily rule out this possibility. Similar observations may be made regarding the change in the method of imputation of consumption expenditure out of 'self-done jobs' and/or homegrown foods or vegetables. At this stage, we emphasise the need for systematic and continual verification of the HES data at the technical level with a view to solve the questions frequently raised by various users concerning the validity and the need for additional information for future HES data collection.

#### 4. OTHER OBSERVATIONS ON THE HES DATA

One of the most obvious drawbacks of the HES data is the delay in the publication of results, although it has become quicker in the later series. Yet, it took about 15 months to get any usable data on income and expenditure from the HES from the completion of data collection. Further, the ABS failed to release any information on income and expenditure data for unemployed families at a time when the country was seriously discussing the 'problem of unemployment'. It is important to know the degree to which the community is suffering in terms of a fall in quality of living due to unemployment.

Household expenditure on different commodities and income data, have been published by weekly household gross income deciles, since the 1984 HES. However, the concept of income and its statistical measurement are questionable. Moreover, data on expenditure and income per family cannot give the real picture of the expenditure pattern of the commodity. This is because Cramer (1973a) has mentioned that

large families tend to have high incomes, and conversely. This is not due to a direct causal link between the two variables (income and family size) but to fortuitous characteristics of our social structure.

Hence, there is a strong positive correlation between income and family size. For this reason, it is advisable to present the HES data for different per-equivalent

adult (or at least per capita) total expenditure deciles instead of household gross income deciles. Literature in favour of total expenditure versus income is abundant [see Cramer (1973a), Friedman (1957), Prais and Houthakker (1955), Podder (1971), Vickrey (1947) and Haque (1984), Dayal, Gomulka, Milford, Southerland and Taylor (2000), Geoffrey and Ferraro (1996), and Thesia and Blanciforti (1994)].

Theoretically, one is interested in the relationship between expenditure on a particular item (Y) and income (X) and hence the same reference time interval (i.e., from July 1 to July 21; not some households in July, some in December, and some in March, etc.) for all households. A moving two or three weeks reference period introduces transitory (seasonal) components in both X and Y. Successive transitory components are correlated. Under these circumstances, it is quite possible that the Engel elasticity for some luxury items may be over-estimated and those of some necessary items underestimated. This subtle problem has been relatively unnoticed in the literature [see Liviatan (1961)]. In this regard, the following steps are suggested to clarify the position.

- (i) Commodity type analysis may be undertaken by collecting budget data for the same reference year, through repeated visits. These data can be analysed to investigate the biases. If any bias is found in the currently collected HES data, a decision may then be taken for future data collection, to convert to a moving annual reference period for clothing and other categories of items where purchases are infrequent and marked by pronounced seasonality.
- (ii) The ABS may also collect data on infrequent purchases of two or three items such as clothing with respect to an annual reference period and the data be analysed to clarify the problem.
- (iii) One may also obtain a better measure of consumption than per capita expenditure by excluding unusual purchases such as housing, cars, health or ceremonial expenses, and use these as explanatory variables in the Engel curve analysis. In this connection, it would be of interest to collect the data from those households who spent a large sum of money for housing, cars, health or ceremonial purposes, by asking how much the household normally spends on such items when they are not making the infrequent, large expenditures on those items. Those data would then be added with other expenditures to get the total expenditure, which can then be used as one of the explanatory variables in a regression model. Total expenditure elasticity for all the consumption items would then be accurately estimated.

Most researchers need some idea of the sampling errors of the ratios, indices or coefficients. Such needs cannot be met by computing the standard errors (S.Es) of a few global estimates. Even if S.Es are presented for every estimate appearing in the 'Tables with notes', the researchers may not be able to compute the S.Es of the ratios etc. In such situations, one may need sampling co-variances of various pairs of estimates. In this regard, one may consider the following list of statistics frequently estimated from HES data:

- (1) cumulative percentages of population living below specified levels of per capita expenditure (X);
- (2) shares in aggregate consumer expenditure enjoyed by various groups of the population;
- (3) Lorenz ratios for the distribution of population by size of (x);
- (4) the analogues of (2) and (3) for specified items of consumption like cereals, etc;
- (5) Engel elasticity of consumption for various items; and
- (6) indices of inter-regional variation (in consumer prices or the like) in average per capita household consumer expenditure.

Scrutiny of statistical data is an 'art' and it is difficult to put objective rules which would be acceptable to all experts. Scrutiny of filled-up HES 'consumer expenditure' questionnaires is essential; because some defects may arise from the failure to distinguish between purchase, expenditure and consumption. Without editing the raw data from the field, the technical section has problems with some absurd entries. The data for households can be scrutinised on a computer, using limits set up after preliminary examination of the data. A number of sample households can be rejected on the grounds of defective entries. A household can be rejected altogether if its per capita expenditure on some item, say for example consumption of cereals exceeds \$30 or 10 kg per week, or the expenditure on cereals appears to be less than 50c or higher than \$50 per kg. Similar rules can be applied to other items.

There are doubts about the utilization of multi-stage probability sampling techniques to measure and control non-sampling errors creeping in during the field work. It is really difficult to find biases that are common to both parties of investigators. However, the technique is extremely helpful in controlling errors entering at the processing and tabulation stages. Unfortunately, the importance of this is not clear to persons who do not have wide experience of survey data processing.

It is standard practice to publish mass individual household expenditure data in the form of cell means for a given classification of sample households. As a result, most analyses are based on grouped cell means, even though individual data are available. This is probably because of high variability in the individual data, which also fails to answer some basic questions such as what is the average expenditure on food made by the poor people of the community, is there any significant difference between the average expenditures on food by the poorer and richer sections of the community. Grouped cell means can answer these questions very accurately, if the distribution of expenditures on those commodities within groups is symmetrical.

Inspection of scatter diagrams of expenditures of different commodities by income groups of a number of individual households shows that the dispersion is so large that any Engel function will give an equally poor fit. A good fit of some specified form can emerge when the within group cell means are used. Cramer (1973a) has shown that the coefficient of determination ( $R^2$ ) is usually low for

individual data and a higher  $R^2$  can be achieved by using the grouped data for the same phenomenon, leaving the properties of the regression estimates largely intact.

However, cell means are sometimes meaningless, since mean values are highly affected by the extreme values. In fact, the 1988-89 HES Unit Record Tape shows that the distribution of expenditure on some commodities, particularly infrequently purchased items, is highly skewed even within the gross income deciles. For some items there are multiple modes. In these circumstances, it is very difficult to accept cell means as the representative values for those groups. Moreover, cell means are dependent on the particular income class limits adopted. The choice of an optimal classification, which maintains homogeneity within each group, is a difficult task. The case of multiple regression grouping may well affect the correlation among the explanatory variables and hence lead to a serious multicollinearity problem, although this does not occur in the individual observations when the least squares method is used. Until now, very little attention has been paid to this problem. Prais and Aitchison (1954) mentioned that an investigation of optimal grouping in multiple regressions should begin from matrix algebra. We urge that the ABS should do some research work on the optimum classification and present the data accordingly instead of in gross income deciles. It is our feeling that the data would be better represented by fractile groupings [Vide, Mahalanobis (1960)], with households having continuous ranks when they were ranked in ascending order of income (say either gross or net). Then the group denoted by 0-5 per cent would comprise the bottom 5 per cent of the estimated number of households, the group represented by 5-10 per cent would comprise the next 5 per cent of the estimated number of households, and so on.

Currently, the ABS publishes the arithmetic cell means within each gross deciles income groups for different consumption items. Aitken's (1934) Generalised Least Squares (GLS) method can be used to obtain unbiased estimates of regression coefficients, using the within group arithmetic cell means if the linear function is chosen. However, Haque (1988, 1989b), Podder (1971), Bewley (1982) and many others have shown that the linear function is not superior over other functions for the Australian HES data. In fact, almost all published works are based on non-linear Engle functions such as semi-log, double-log, and log-log inverse. In such cases, Aitken's GLS method needs the within group geometric/harmonic means for the logarithmic/inverse functional forms, in order to estimate unbiased regression parameters and hence income elasticity. A bias is introduced when the within group arithmetic cell means are used as a proxy for the geometric/harmonic means in estimating non-linear Engel functions. Both Kakwani (1977b) and Haque (1990a) showed that these biases are significantly higher for some items, using the estimated within group geometric/harmonic means based on the concentration curves. It is thus urged that the ABS should produce at least one set of within group geometric/harmonic means for some broad groups of items. See more about the uses and validity of the HES data in the United Nations (1989), Taylor, Gomulka and Sutherland (2000), Sutherland, Taylor and Gomulka (2001) and Mario (1995).

## 5. RECOMMENDATION AND CONCLUDING REMARKS

Some recommendations and conclusions are made in this section. To examine the validity and to improve the quality of the HES data, the following recommendations are made.

Substantial technical work should be done by the ABS and elsewhere to verify the HES data both for internal consistency as well as against external evidence.

Fieldwork should be undertaken to investigate:

- (a) the effects of various changes made in the questionnaire over time in collecting the HES data;
- (b) the validity of the estimates for meals out and take-out food and drinks;
- (c) the effects of seasonality on the budget data;
- (d) the effects on data collection of the 'acquisition' method; and
- (f) the appropriateness of the month reference period for clothing and some other durable items.

Improved procedures of scrutinizing the field data should also be evolved and standardised.

- A systematic record of expenditures and incomes and other appropriate characteristics of the households at different income level particularly the lowest and the highest income groups should be kept.
- Expenditure, quantity as well as price information for various goods and regarding changes in the quality of living and in disparities of level of living in real terms cannot be answered satisfactorily with the present data. The HES data of various 'series' cannot be compared in the absence of appropriate Consumer Price Indexes for different income groups (deciles groups) of the population.
- It is important to publish estimates of consumption in quantity for a number of food items along with average prices. These are important for studies of nutrition. Also, Hofsten (1952) has shown that quality variations have an economically significant effect on the calculated value of the index number.

If the HES data have to meet the growing and varied needs of researchers and policy makers then the following recommendations should be given special considerations.

- All estimated statistics should be published promptly. They must include expenditure and quantity for detailed food items and other important groups of expenditure items.
- Statewide households and persons based data in different size-classes (deciles groups) of per capita expenditure data (if not per equivalent adult expenditure) should be published.

- A table for broad groups of items kept consistent for every 'series' should be published on the basis of per capita (if not per equivalent adult) fractile groupings of the total expenditure.
- It is also argued that the ABS should produce one set of within group geometric/harmonic means for broad groups of items based on per capita fractile groupings of the total expenditure.
- Estimates of standard errors and co-variances of some commonly estimated statistics mentioned in Section 4 should be published.
- The 'Unit Record Tape' should be released quickly to the users of the HES data.
- We finally would like to make a suggestion to the ABS to collect the same HES information repeatedly from version to version for some specified households (say 1500) who are socially disadvantaged such as aborigines, migrants and low income households (viz., single mothers, retired elderly people, etc., whose per capita annual income say is less than \$10,000). This will help us to investigate how the living standards of those families are affected due to many changing situations such as economic recession, government interventions and changes in social and economic environments. Short and long term income elasticity of some consumption items using such **panel data** on households would be of great interest in making proper social policies. The demand for intake of nutrients such as Bhargava (1991), Strauss and Thomas (1998), and National Nutrition Surveys can also be analysed using time observations on individuals of households. It is hoped that these data will support the view that increases in household incomes will in turn improve the intakes of nutrients, which the Australian National Health and Medical Research Council (NHMRC) tries to establish for health reasons. This suggestion is quite consistent with an ABS Australian Statistician, Castles (1991), who wrote as:

Our mission is to assist and encourage informed decision-making, research and discussion within governments and the community by providing a high-quality, user-oriented and dynamic statistical service....

He also mentioned that in introducing the Australian Bureau of Statistics Bill 1975, the Honourable Minister L. F. Bowen, MP pointed out that:

There is no need for me to argue the virtues of statistical information in providing a generally informed society; in providing a firm basis for decision-making in providing a basis for the development of programs and a means of measuring their progress over time.

In conclusion, we would like to mention that the stratified multi-stage probability sampling procedure should be continued in the collection of HES data. In future, the data including tabulation and analyses of the HES as recommended in this study, should be supplied to the various researchers. It is important to bear in mind that the consumer expenditure enquiry over time for some specified households would be a major feature for future HES data collection. However, we may state that the risks of non-comparability of data from successive enquiries of the same nature would be minimised if the enquiries were carried out continuously. Also if one carries out

budget enquiries at regular intervals, say once in every four or five years, the fluctuations in some consumption products may prevent us from examining gradual trends in levels of living. It appears to be wise to distribute the resources and effort over many medium-type enquiries to be carried out every year instead of conducting a large-survey at a regular interval of say every five years. Household Expenditure Survey data of this type is important for future development and social policies, and can be found in Carnegie and Walker (2001), Dasgupta (2001), Deaton (1997), Dolton (2002), and Lechner and Pfeiffer (2001).

Income Elasticity and Economic Development  
Methods and Applications

Haque, M.O.

2005, XVII, 259 p.,

ISBN: 978-0-387-24344-3