

77543

Working Paper No.183

On the Validity of NSS Consumption data

J. Vaidyanathan

Centre for Development Studies
Uloor, Trivandrum 695 011

December 1983

On the Validity of NSS Consumption data

A. Vaidyanathan

1. For nearly three decades the National Sample Survey (NSS) has been collecting, every year upto 1973-74 and at five year intervals thereafter, data on the level and pattern of consumption from a representative sample of households in rural and urban areas of different parts of the country. It is by far the most comprehensive source of information on the subject. No other source provides such detailed continuous and apparently comparable data over as long a period. Naturally they have been used extensively for studying a variety of questions such as the sources of variation in per capita consumption levels (Visaria 1980); responsiveness of consumer demand patterns and demand for particular commodities to changes in incomes and prices (See among others Ganguly et al 1960; Iyengar 1967; Rudra, 1972); inter-regional disparities in living standards (Bhattacharya and Chatterjee 1974); inequality in consumption, incidence of poverty and changes therein (See e.g., Srinivasan and Baroan eds, 1974; Dandekar and Rath, 1971; Rajaraman, 1974; Ahluwalia, 1978). While these studies provide many valuable insights the question of the validity of the NSS data keeps surfacing in the discussion particularly in the context of the debate on trends in incidence of poverty.

2. Several important questions concerning the quality of NSS consumption data have been raised in the literature, prominent among them being: the chances of systematic understatement of consumption by upper income groups arising from the relatively small number of rich households in the sample

(Dandekar and Rath 1971); the possibility of inaccuracies in the information obtained from the sample households (Bhattacharya and Chatterjee, 1975); the appropriateness of current NSS design to get a reliable estimate of frequency distribution (Murthy 1977); comparability of estimates over time (Mukherjee, 1981; Bhattacharya and Chatterjee 1975); and the reasons for and significance of divergence between the NSS and the official estimates in respect of level and composition of consumption (Kansal, 1965) their behaviour over time (Mukherjee 1969; Mukherjee and Chatterjee 1972; Srinivasan et.al 1974) as well as its implication for inferences on changes in the incidence of poverty (Vaidyanathan 1974). Discussion of these questions has, however, been rather fragmentary and inconclusive. Considering their importance for the on-going debates on various aspects of growth and distribution, it seems worthwhile to review systematically the various possible sources of "bias" and "error" in the NSS data on consumption expenditure and its distribution by classes and examine, with such evidence as is available, whether and in what manner they in fact affect the reliability of NSS estimates. This is what the present paper seeks to do. While we are in no position to settle all the questions definitively, the evidence suggests that there is strong reason for being circumspect in using NSS estimates as the basis for judging trends in consumption and in the incidence of poverty.

Sources of Inaccuracy in NSS data

3. The reliability of NSS data on consumption (in the aggregate and by particular commodity groups) at a given point of time has to be judged basically in terms of how closely they correspond to the "true value" of these characteristics. The sample survey estimate may differ from the true

value because of sampling error, defects in sample design, inaccuracies in the information obtained from sample households or a combination of these. For a study of changes in consumption, the magnitude of the sampling error in the estimate relative to the true rate of change is clearly relevant: The larger the former relative to the latter, the less reliable the sample estimate for assessing change. Other biases and errors need not vitiate use of sample data for this purpose and for assessing changes in inequality so long as their magnitude relative to the true value in different regions/classes does not change systematically over time.

Sampling Errors

4. Any attempt at estimating the 'true value' of a characteristic of a population from observations relating to a sample thereof is necessarily subject to a "sampling" ^{error.} The NSS reports seldom give the sampling error of its estimates for consumption and its components. However, until the mid seventies, the NSS design had inter-penetrating sub-samples and the reports invariably gave sub-sample-wise estimates which provide a rough indication of the sampling error for various estimated magnitudes.^{1/} On this basis it would seem that the all India estimates of per capita consumption (in the aggregate, by major commodity groups and also by expenditure classes) were subject to a relatively small margin of sampling error.^{2/} The errors are naturally larger in the case of state level estimates and estimates for detailed commodity groups. For this reason, and given the relatively slow pace of change during the last 2-3 decades, even if the information were otherwise accurate and unbiased, the NSS is likely to be less reliable for measuring the magnitude of differences in levels of

expenditure (aggregate and even more by commodities) across regions and classes and of the magnitude of changes in them. However it is not always essential to have accurate estimate of levels or of percentage change; useful insights on several questions can be obtained with information on relative positions of region/classes in respect of total consumption, consumption of particular commodities and the composition of consumption.

Sampling Design

5. Sample households for consumption enquiries are selected essentially on the basis of probability proportional to population without reference to the level of their income or consumption. In this process the number of "rich" households likely to be included in the sample is necessarily small and consequently the estimate of the upper income groups consumption will have a higher margin of sampling error compared to lower income groups. For this reason some people have questioned the suitability of the present NSS design to get a reliable estimate of frequency distribution. Murthy (1977:192) observes:

"since the NSS estimates are based on general purpose design with emphasis on point parameters, the estimates of tails, in which the users are specifically interested, are subject to larger sampling error and hence not amenable to deeper analysis"

If we are interested in measures of relative inequality, this is indeed a serious limitation. However it has been shown (Srinivasan et.al.1974) that the larger sampling error (or for that matter non-sampling errors and bias) in the upper tail does not necessarily vitiate estimates of the incidence of poverty in the population or changes therein.^{3/}

6. Over the years the size of NSS sample for consumption enquiries has progressively increased^{4/} and so the sampling errors must have fallen. However

even as the NSS sampling design has remained basically unchanged, the basis for defining the strata, the number of strata, the procedure for selecting the sample households and the size of the sample for the consumption survey have all undergone changes. Thus the number of regions, which defines the strata for selecting sample villages, has risen from 48 in the early 1960's to 66 in 1972-73. The number of sample households surveyed for consumption expenditure has risen from 1 per sample village between 1958-59 and 1960-61 to 2 in 1961-2, 6 in 1970-71 and 12 in 1972-73; in 1973-74 it was again reduced to 2 per village (Gupta and Ramaratnam, 1975). It seems likely that these changes affect only the variance of the estimates without affecting comparability in any other way. However, the NSS also stratifies households within each sample village by household size, and means of livelihood or land possessed before sample selection. The basis of stratification and the number of strata vary (in 1964-65 there were as many as 6 while in 1970-71 there were only 3).^{5/} It is not clear how this stratification enters into, and affects, the selection of sample households for canvassing the consumption schedule especially in years (e.g., 1960-61 and 1961-62) when the number of households per village is less than the number of strata. Nor is it clear whether and to what extent these changes may affect the comparability of the estimates over time. These are technical questions which can best be clarified by specialists in sampling.

Bias in Sampling

7. Then there is the question of 'bias' in the NSS design. It has been alleged that the NSS has an in-built "bias" to under-estimate the true value of consumption per capita and the extent of inequality in distribution.

because (a) the sample has relatively few households from the 'rich' classes who account for a disproportionately high percentage of total consumption; and (b) it does not fully capture seasonal variations in consumption. Both these propositions are erroneous. As Srinivasan et.al (1974) have pointed out the fact that the NSS Sample has few "high" income households essentially increases the variance of the estimate of the average for the population and for the upper tail"; it does not imply any in-built "bias" in the sampling procedure toward systematic over or under estimation of the true value in any class. Again the practice of spreading the collection of data relating to the sample households over the year effectively captures seasonal variations in consumption. Some have argued however that the present procedure of canvassing different sets of households during each sub-round may exaggerate the degree of inequality.^{6/} There is hardly any other basis to expect "bias" in the sampling design itself.^{7/} Systematic errors can arise for reasons independent of sample designs: For instance a higher proportion of upper income sample households may refuse to respond; informants, or some sections of them, may systematically over- or under-state their actual consumption; and inaccurate reporting of information^{7a/} could arise from recall lapses, the way questions are framed and communicated or simply negligence.

Non-Sampling Errors

8. The incidence and magnitude of these "non-sampling errors" could be randomly distributed among the sample household in which case their effect on the estimate is similar to that of sampling errors though the magnitude of overall error in the estimate will naturally be larger than sampling error. Much the same would be the case even when the non-sampling errors are systematic so long as all households tend to over - or under-state the

true value of the characteristic to a uniform degree. In both cases, the errors in the estimates are affected but, subject to the caveat mentioned by Murthy, the estimated distribution around the mean should not be affected. However, when the magnitude of non-sampling errors relative to the true value varies systematically in the sample — this would be the case for instance when the degree of over or under reporting is greater in the upper income households than in poorer classes — the accuracy of both the average and its distribution are affected.

9. The nature and incidence of non-sampling errors may also change over time due to changes in concepts, the way questions are framed, or in the investigation procedure. Clearly if the systematic errors in different classes can be shown to have no particular trend over time, the sample data can be used with greater confidence to speak about changes in per capita total consumption and its distribution. On the other hand any evidence that these errors themselves vary in some systematic way over time would throw doubt on the usability of NSS data for evaluating changes over time. The rest of our discussion is largely concerned with the nature and direction of non-sampling errors in NSS consumption data. (Questions concerning sampling design are left to the specialists). It is based partly on a comparison of the NSS estimates of per capita consumption levels (in the aggregate and by major product categories) with corresponding estimates of the CSO, and partly on evidence internal to the NSS. In both respects this paper seeks to extend earlier work by using more recent information.

Comparison of NSS and CSO estimates

10. More than a decade back, Mukherjee (1969) had emphasised the close correspondence between CSO and NSS estimates of per capita consumption as

a basis for confidence in the NSS data. The force of this argument was however weakened by the increasing divergence between the two estimates from the early 1960's (Mukherjee and Chatterjee 1972, Srinivasan et.al., 1974). Since then the CSO has published a detailed time series of consumption (aggregate and by commodities) at current and constant prices. The CSO series differs markedly from those of the NSS not only in the aggregate but also in respect the major components categories of consumption; more importantly the two series show quite disparate trends.^{8/} (The basic data are given in Table 1). Thus the divergent trends between the two series of percapita total consumption noticed in first half of the sixties has been reversed after 1967-68 when the difference has progressively narrowed. However while the NSS shows a significant decline of PCE upto 1967-68 and a steep rise thereafter, total PCE in 1973-74 was no higher than in 1960-61. The CSO series on the other hand shows a sustained rising trend over the entire period, the level of PCE in 1973-74 being some 13 percent higher than in 1960-61.

11. The relative behaviour of the two series of real consumption in respect of major categories of goods and services can be seen from Chart 2. In the case of food other than foodgrains, though the CSO estimates are considerably higher than NSS the two series move more or less the degree of difference shows ^{but} a mild tendency to rise over time. Neither series shows any significant rise in per capita real consumption level over the period as a whole. The CSO estimates of per capita consumption of non-food items as a group are not only higher than NSS but shows a strong rising trend. The NSS shows a decline upto 1967-68 followed by a rise. While this has brought the two series closer, the level of per capita consumption in 1973-74 was only marginally higher than in 1961-62 according

Table 1: Per Capita Consumption expenditure estimates by major commodity Groups, India, 1960-61 to 1973-74

	Rs. per annum in constant (1960-61) prices												
	1960-61	1961-2	1962-3	1963-4	1964-5	1965-6	1966-7	1967-8	1968-9	1969-70	1970-71	1972-73	1973-74
Total	276	270	273	291	278	282	294	299	303	314	302	313	286
Foodgrains	279	273	254	270	265	257	246	257	256	263	279	286	
Food Excluding foodgrains	78	77	75	82	69	72	81	83	83	88	78	81	111
	110	113	104	103	106	105	98	97	96	98	101	111	
Non Food													
	68	87	86	91	88	85	87	85	71	96	92	97	73
	74	71	67	60	69	65	61	65	70	75	66	73	
	110	106	112	119	120	125	126	131	129	126	132	134	94
	94	90	84	85	83	79	78	83	87	86	97	94	

Notes: 1. CSO estimates taken from National Accounts Statistics 1960-61 to 1974-75 (New Delhi, October 1976). The publication gives estimates at current price and at 1960-61 prices. The implicit deflator for each category of consumption is used to derive per capita consumption at 1960-61 prices corresponding to the NSS series. The deflation is done for total consumption and for each of the categories separately. Therefore, the sum of deflated expenditure in component categories do not always add up to the figure under "total".

2. The NSS estimate have been taken from publication entitled "Tables with notes on consumer expenditure" for various rounds. The reports, identified by numbers, from which the figures have been estimated are listed below

1959-60	1960-61	1961-62	1963-64	1964-65	1965-66	1966-67	1967-68	1968-69	1969-70	1970-71	1972-73	1973-74
104	101	135	142&191	192	209	230	216	228	235	231	284	240

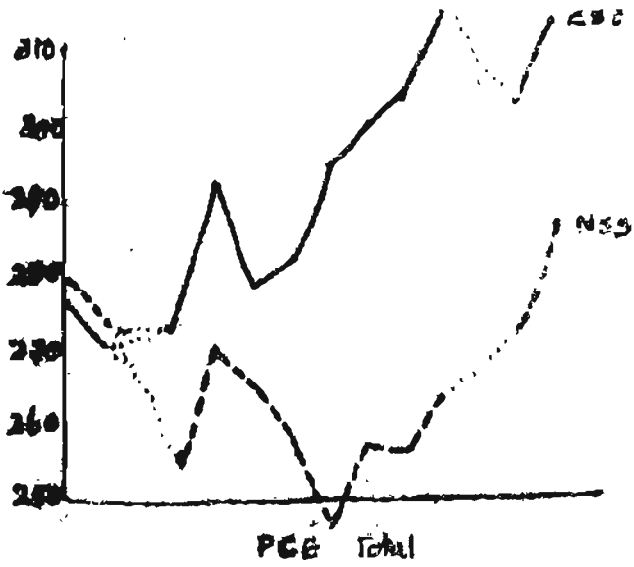
NSS Report No.

The figures relate to total population and have been estimated as the weighted average of rural and urban per capita figures using the proportion of population in the two sectors as weights. These proportions are based on the Census data for 1961 and 1971 with the values for other years obtained by linear interpolation.

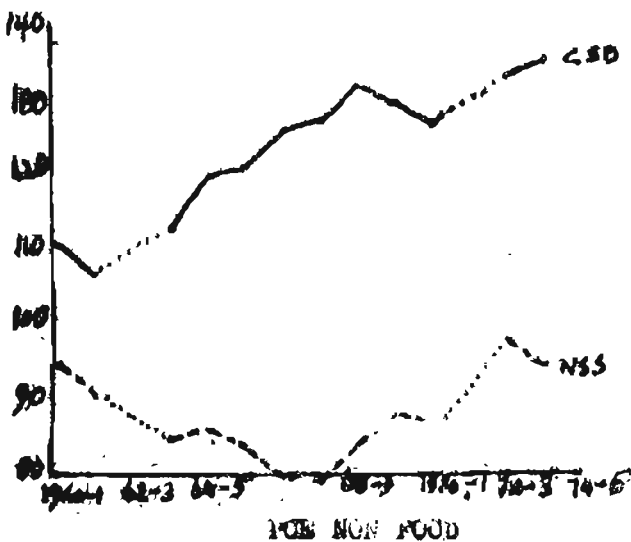
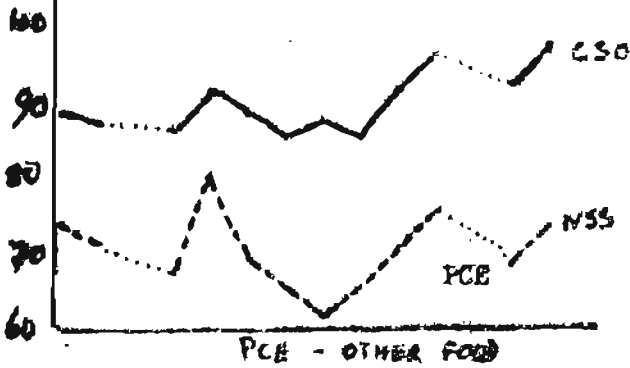
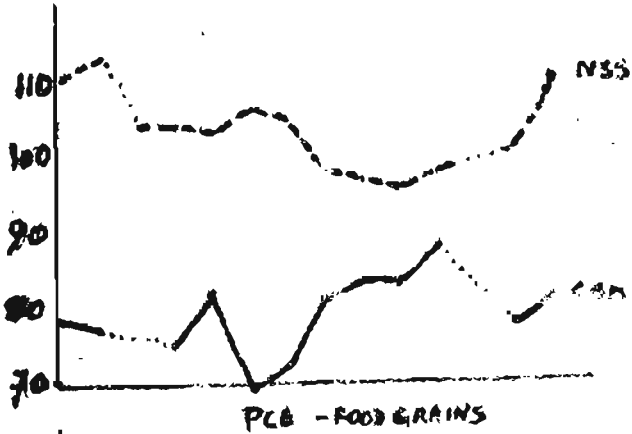
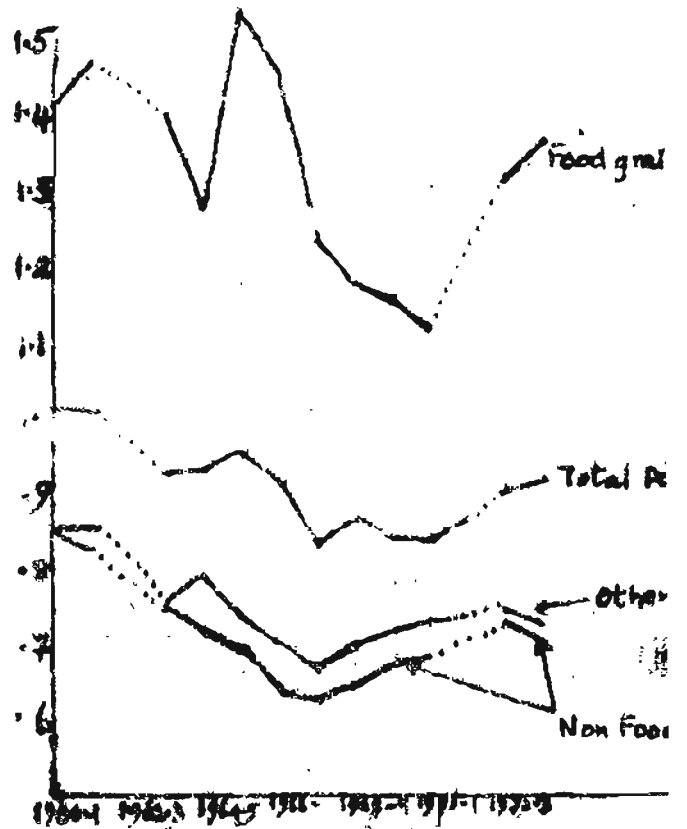
to the NSS, while CSO series show a 19% rise. But the most dramatic difference is in respect of food grains. Between 1960 and 1970, the NSS shows a sustained decline of nearly 13-14%, but estimated consumption rose sharply in 1972 and 1973 bringing it close to the 1960-61 level. The CSO on the other hand records erratic fluctuations upto 1970 followed by a steep decline in 1972 and 1973. In relative terms the difference between the two estimates fell during the sixties but has again increased in the early seventies.

12. There are ofcourse well known reasons why the two estimates of private consumption cannot be expected to agree. In the first place the NSS estimates of private consumption relate strictly to direct consumption by households while those of the CSO are derived by deducting estimated usage for intermediate inputs in the process of production, for capital formation, public consumption and exports from the estimates total availability of each commodity and service (production plus imports). The residual measures not only the direct consumption of households, but also "consumption" by institutions (hostels, temples, charities, etc.). However, the effect of this should be to make the CSO estimate generally higher than that of NSS. The second reason is that the CSO estimate is derived from estimates of availability and of absorption for uses other than private consumption all of which are subject to errors. Errors in the estimate of any of these magnitudes (which arise in part from lack of reliable data and in part to errors in data) necessarily affect the accuracy of the estimate of private consumption. Indeed for several items the CSO estimates are based on such weak data that one would be inclined to rely more on the NSS whose estimates are based on direct enquiry from consumers. This especially the case where "household consumption" is relatively easy

Estimating PCE in Rs for Income
at 1960-61 prices, NSS and CSO



Ratio of NSS to CSO estimate
of PCE



to define in clear terms and the informants are not required to remember a great many details.

13. Nevertheless it is instructive to compare the NSS estimate for specific commodities which are mostly used by households for consumption with corresponding CSO estimates especially when the latter are known to be based on data collected systematically and independently of NSS. If the NSS estimate is lower and if the two series move more or less parallel to each other, confidence in the reliability of NSS data for evaluating trends would be strengthened. The CSO and the NSS estimates of per capita real consumption of cloth, sugar and gur, edible oil are set out in Table 2. In all these cases the NSS estimate in most years is lower than that of the CSO; but they show quite dissimilar patterns of change over time. In the case of edible oil and clothing the two series often differ even in the direction of year-to-year change the overall trend is also dissimilar in the case of sugar-gur and even more in the case of clothing. In the latter case the NSS series point to a steep decline upto the mid-sixties followed by a rise with the per capita consumption in the early seventies being somewhat below the 1960-61 level; the CSO series points to a more or less steady rise of some 35% over this period.

14. In the case of foodgrains, we have already noted the disparate time pattern of change in per capita real consumption (i.e. value of consumption at constant prices). A more direct comparison of the differences in terms of quantities consumed is possible in the case of the 3 major cereals. Table 3 gives estimates of the per capita consumption of rice, wheat and other cereals and total cereals derived from three different sources namely, (i) the official estimates; (ii) the estimates from the NSS consumption

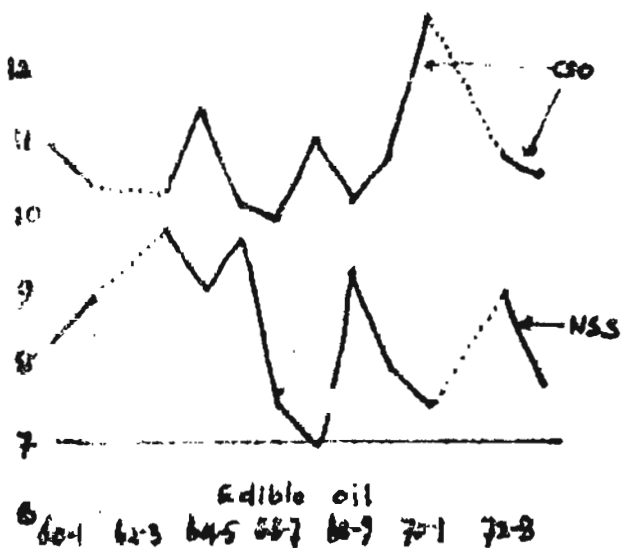
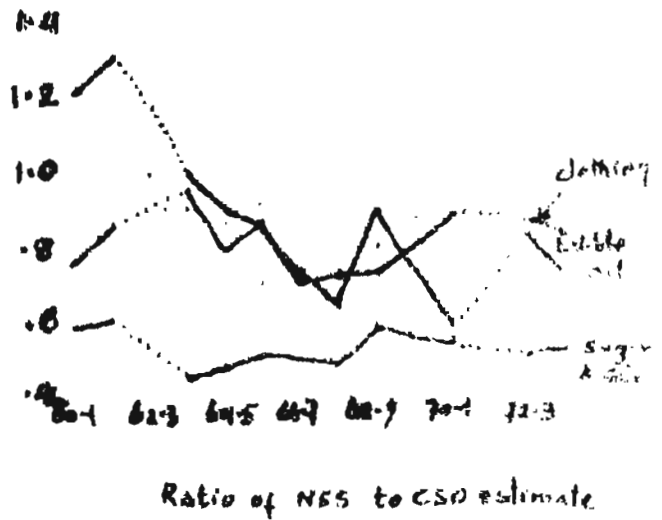
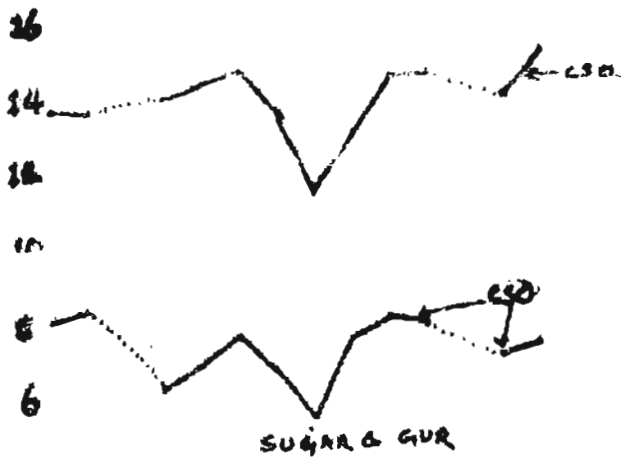
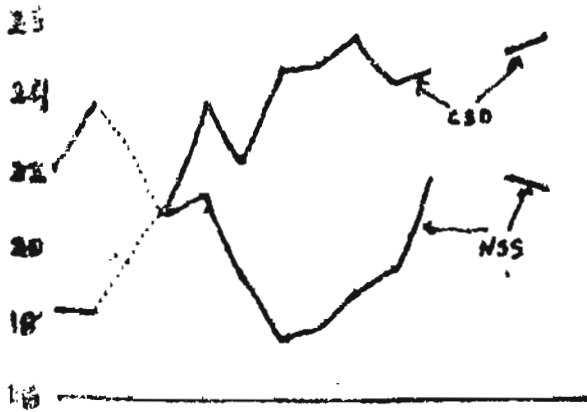
Table 2: Indices of per capita real consumption of fooded items as estimated by CSO and NSS

(1960-61 = 100)

	1960-61	1961-62	63-64	1964-65	1965-66	1966-67	1967-68	1968-69	1969-70	1970-71	1972-73	1973-74
	Rs/ Caput											
Edible Oils												
CSO	10.9	95.3	94.5	104.6	92.7	91.7	100.9	93.6	98.2	115.6	98.2	97.2
NSS	8.3	106.6	112.1	108.6	104.1	89.1	64.3	110.6	95.7	85.0	107.5	93.7
Sugar and Gur												
CSO	13.7	100.1	102.2	105.1	108.	100.	83.9	95.6	106.6	107.3	102.9	111.7
NSS	8.1	102.2	77.2	85.9	94.6	84.5	68.9	94.8	101.2	99.2	89.0	98.3
Clothing												
CSO	18.3	99.5	115.3	130.1	121.3	135.0	135.5	139.9	132.2	134.4	137.2	139.9
NSS	22.1	107.5	94.7	96.7	86.0	78.8	80.5	84.4	87.7	95.3	98.4	96.9

Source: See Table 1

PER CAPITA CONSUMPTION
(RS PER ANNUM AT 1960-1 PRICES)
OF SELECTED ITEMS



surveys; and (iii) estimator based on the NSS Land use and Crop Cutting Surveys. Since the relevant data are not available for all years in the case of NSS, we have compared averages for groups of years at different points between 1958 and 1973. They show that in all three estimates the direction of change in respect of total cereals as well as individual cereals is generally similar. For instance all three series point to a progressive rise in wheat consumption per capita over the period and a general decline in the case of rice and other cereals. However the two NSS based estimates show a much steeper and more sustained decline (or a much elower rise) than the official estimates.

Table 3: Alternate Estimates of per capita consumption of different cereals
in India

	(kg per annum)											
	1/ 1959-61			2/ 1965-64			3/ 1966-1968			4/ 1970-73		
	Off	NSS(P)	NSS(C)	Off	NSS(P)	NSS(C)	Off	NSS(P)	NSS(C)	Off	NSS(P)	NSS(C)
Rice	70.6	82.3	93.5	71.0	74.2	93.0	60.8	56.8	na	68.3	na	78.6
Wheat	28.7	37.6	36.2	31.0	37.4	38.1	34.3	40.0	na	42.4	na	43.8
Others	44.0	62.6	71.2	41.6	46.5	63.0	42.0	47.3	na	36.2	na	53.1
All Cereals	143.4	181.1	200.9	143.5	158.0	194.1	137.1	144.2	na	146.8	na	175.7

Off - Official Estimates

NSS(P)- Estimates based on NSS LU-CCE Survey

NSS(C)- Estimate based on NSS Consumption Surveys.

1. 1959-60, 1960-61, 1961-62

2. 1963-64, 1964-65

3. 1966-67, 1967-68, 1968-69

4. 1970-71, 1972-73, 1973-74

(Official year one calendar year 1959-60 take as 1959 etc.)

Source: Various NSS reports

The Nature of Non-Sampling Errors in NSS

15. The mere fact that NSS estimates diverge significantly from the official figures does not tell us much about the relative reliability of either. However there are independent reasons to believe that NSS estimates in respect of foodgrain and clothing are subject to rather large non-sampling errors, that these errors have been changing systematically over time, and further that the incidence of these changes is not the same in all classes.

16. The NSS rightly defines "household consumption" of foodgrains (and of other items) in terms of the consumption of the members of the household; it explicitly excludes any payments in grain for production activities, payment in exchange for consumer service (priests, barber washermen etc.) as well as transfers and gifts. The instruction to field workers for the 25th round (1970-71) states

"the data....should be strictly restricted to the non-productive consumption of the household including that of pet animals....While recording entries for household consumption care should be taken not to include any transfer payments in kind like loans advanced, charities, gifts and other payments in kind...But any consumption out of transfer receipts in kind like borrowings, gifts, charities, perquisites received by the household free collection and other receipts in kind if any will have to be considered" (GOI, NSS, 1970)

Payment in the form of grains or cooked food to hired workers for farm or non-farm enterprises of the household should thus be excluded from the estimate of its consumption. Consumption out of receipts in kind as wage must however figure in the schedule of the recipient household as part of its consumption. As for payments in grain for consumer services (priests, barber etc.) the Instructions require that such payments should be shown

in terms of imputed value against the appropriate item under "Miscellaneous goods and services"^{2/}. Payment for services in cooked meals is to be recorded at their imputed value against the particular service in the schedule of the paying household and as consumption of cooked meals in that of the recipient household.^{10/} However, the quality of the data actually reported in the schedule depends on whether investigators attempt to get, and succeed in getting, detailed information on these aspects.

17. The "consumption" of foodgrains by a household in a given period consists of the quantity cooked at home and eaten by members of the household as well as cooked food obtained from employers or as gifts, transfers and loan. The grain used in the former may have been obtained from several sources, namely past accumulations of grain stocks, fresh receipts in the form of uncooked grain for wages or as gifts, and fresh purchases. But since part of available grain may be used for other than household consumption, the estimation is complex.

$$\begin{aligned} \text{Total quantum of grain consumed} &= \text{Quantum cooked at home} + \text{Quantum} \\ &\quad \text{received as cooked food from others} \\ &\quad - \text{Quantum of cooked food given to} \\ &\quad \text{others} \end{aligned}$$

$$\begin{aligned} \text{Quantum of grain cooked at home} &= \text{Opening stock of the grain} + \text{Kind} \\ &\quad \text{receipts (own farm, payment for} \\ &\quad \text{services, gifts)} + \text{Purchases} - \\ &\quad \text{payments of uncooked grain for} \\ &\quad \text{wage, gifts etc.} - \text{sales} - \text{closing} \\ &\quad \text{stock} \end{aligned}$$

18. Given that few households keep any systematic accounts and that the foodgrain block in the NSS Schedule lists 25 to 40 items, getting reliable information on these aspects is a formidable task even for the best of investigators. That the schedule is not designed to reconstruct the sources and disposition of foodgrains in a systematic manner with built in checks

on internal consistency, adds to the difficulties. Thus the foodgrain block seeks information only on total consumption, consumption out of homegrown stock, purchase and, in some years, receipts in kind in exchange for goods and services. Consumption in the form of cooked meals is separately. Stocks are beyond the scope of the schedule and value of payments in exchange for goods and services are left to be recorded against the latter items.^{11/}

19. Under these circumstances, there is clearly considerable scope for errors in estimating household consumption of foodgrains. The most obvious possibility is the failure to fully capture cooked food received from others as part of the consumption of the recipient household and from failure to fully net out payment in cooked and uncooked grains under various categories listed above. The former seems likely to be relatively more important in poorer households (especially wage labourers and service castes) than among the rich and payments (of both cooked and uncooked food) on the other hand are apt to figure more prominently in the richer households who after all are the principal employers of attached labourers, domestic help and service castes. Much the same is true of food consumed at feasts. There is thus a real possibility of double counting of foodgrain consumption and a systematic tendency for the NSS estimate to be higher than the true value of consumption.^{12/}

Not only does the NSS tend to over-state the consumption of foodgrains but, if as seems plausible the quantum of kind payments for wages and services and by way transfer, gifts and loan relative to household consumption among upper income households is higher than among the poorer classes, the incidence of the over-estimation is likely to be greater among the former. This may well be the explanation for the improbably high level of foodgrain consumption in the top deciles of the population.^{13/} It would also mean that NSS would

tend to over-state the inequality in foodgrain consumption.

20. These problems would also affect estimates for other food items but to a lesser degree because they do not figure as prominently as cereals in payments to workers and artisans and because a larger proportion of household consumption is purchased. On the other hand, the use of these products outside the household sector (mostly by restaurants, eating houses and manufacturers of prepared food and, in the case of edible oil, for non-food uses) seems likely to be much larger than in foodgrains.^{14/} In the case of clothing, since non-household consumption is relatively small one should expect a much closer correspondence between the two estimates. But they differ partly because of the difficulties in getting accurate data on receipts and payments in kind and gifts (especially at marriage and other ceremonial occasions) and partly from the differences in the concept of consumption. Unlike the CSO, the NSS defines clothing consumption in terms of the time when piece of clothing is brought into use rather than in terms of purchase. These factors could lead to NSS estimate being lower than the CSO

Changes in non-sampling errors

21. There are also reasons to believe that the magnitude of non-sampling errors has not remained constant over time and may will be changing systematically. One indication of this is that both in the aggregate and in all the commodity groups reviewed, except foodgrains, there is a growing divergence between the two series upto the mid-sixties followed by a sustained tendency to converge. In the case of foodgrains the pattern is reversed. A second indication is the apparent change in the incidence of non-response. In 1959-60, the number of households actually surveyed was reported to be 1.5

percent fewer than the number allotted according to the sample design; in 1960-61, this difference was less than 1%. Thereafter there seems to be no difference between the number of allotted and surveyed households. Third, the NSS has changed its method of enquiry from 1964-65: Upto that point schedules relating to different socio-economic aspects of households were canvassed in relation to different sets of sample households. Thereafter, the procedure was replaced by the Integrated Household Survey (IES) whereby information on all socio-economic and demographic aspects selected for enquiry in a particular round were collected from the same set of sample households. The IES provided the means to incorporate closer checks on the internal consistency of responses. This could have made a difference to the estimates of both the level and the pattern of consumption expenditures.^{15/} But to our knowledge no attempt has been made to ascertain whether in fact this change made any significant difference to the data compared to those obtained by the earlier method.

22. There have also been significant changes in the level of detail of item classification under various categories of consumption. Thus in the case of cereals, schedules of the early sixties listed 31 items; the IES schedule for 1964-65 lists 26 in the detailed version and 10 in the abridged version; the 1973-74 schedule lists 10 items under this category. In the case of clothing the consumer expenditure schedule in 1960-61 and 1961-62 listed 56 items under clothing; in 1964-65, when consumption enquiry was merged into the integrated household survey, the number was reduced to 16; and subsequent round listed 23 items including footwear. The number of items under miscellaneous goods and services was also reduced from 95 in 1960-61 and 1961-62 to 59 in 1964-65. Thereafter 66 items are listed.

The level of aggregation and the names used to denote various items seems likely to make a difference to the response. The more detailed the itemisation, the lesser the chances of omission, and the more accurate the estimate of total consumption which is obtained by adding expenditure on individual items. In view of the evidence of systematic inaccuracies in the reported consumption of such important items as foodgrains and cloth, we cannot take it for granted that recall and reporting errors for various items are random in nature and tend to cancel out. These aspects however, have not been systematically investigated.^{16/}

23. The format for recording data on consumption of food and clothing has undergone marked changes. For all items under these two categories the 16-17th round schedules sought information separately in respect of (a) receipts in exchange for money; (b) receipts in exchange for goods and services; (c) consumption out of home-grown stock; and (d) consumption out of gifts, loans etc. The 18th-20th round schedules distinguish only two categories, namely cash purchase and consumption out of home grown stock. The category of receipts in exchange for goods and services was reintroduced in the 25th round for cereals and cereal substitutes only to be given up again in 1973-74. In the case of clothing this category was apparently not reintroduced. As for durables, where again the concept is "purchase", the 17th round included value of second hand purchases in total consumption; the 1965-66 schedule did not even seek the information while in 1967-68 the information is obtained but it is not clear whether it is counted as part of total consumption; the item on second hand purchase was dropped in 1970-71 only to be reintroduced in 1973-74, again, however, without clarifying whether or not it is to be counted as part of consumption.

Effects of non-sampling Errors

24. We have no way of knowing exactly what effect these changes may have had on the estimated level of per capita consumption and its composition. However it seems likely that the estimates of foodgrains and clothing would be significantly affected especially by the changes indicated in para 20 and, in so far as investigators try to balance total consumption against the sum of the quantities and values by source of supply, also by those mentioned in the previous para. There are in any case independent reasons to believe that there have been significant and systematic changes in the extent of error in the NSS estimates. That there is a pattern in the relative profiles of the two series in most cases is one indication of this.

25. In the case of foodgrains we have two ^{further} reasons for this surmise. Firstly, unlike in NSS Consumption surveys the techniques of estimation underlying the official series of foodgrain production have remained by and large constant over the last 2 decades, at any rate between 1960 and 1973 to which our comparisons relate. These estimates also seem more in line with the recorded increase in cropped area and in the major inputs (namely irrigation, fertilisers and HYV).^{17/} Secondly, close examination of the NSS series shows that the decline in the per capita food intake is not uniform across all segments of population. In fact, between the late fifties and early seventies, the NSS estimates suggest that the per capita intake of cereals of the poorest quartile of the rural population has been more or less stagnant while that of the richest quartile shows a sustained and steep reduction (Table 4). This seems altogether implausible

Table 4: Per Capita Cereal Consumption of poorest and richest quartiles of rural populationAll India

	1958-59	1959-60	1960-61	1961-62	1964-65	1965-66	1970-71	1972-73	1973-74
kg/Capita/30 days									
First Quartile									
Total	11.29	11.97	12.2	12.77	11.25	10.30	11.61	10.99	11.14
Rice	4.45	5.04	5.17	5.61	5.49	4.47	4.24	4.66	4.91
Wheat	1.09	.79	.87	1.03	1.21	1.35	1.33	1.96	1.63
Other	5.75	6.14	6.16	6.13	4.55	4.48	6.04	4.37	4.60
Fourth Quartile									
Total	23.3	23.55	22.26	22.26	21.05	20.00	19.04	19.27	18.91
Rice	10.35	9.97	10.64	10.75	10.08	8.87	8.97	7.77	8.58
Wheat	4.58	4.83	5.80	5.15	4.97	5.19	4.61	6.40	5.05
Other	8.37	8.75	6.14	6.36	6.	5.94	5.46	5.1	4.68

Estimated by interpolation from estimated per capita intake of total cereals rice and wheat in different per capita expenditure class

Source: Various NSS reports



26. In the case of clothing the official statistics are based on a fairly elaborate system of direct reporting of production, inventories and foreign trade, / the definition of consumption in terms of the time when they are brought into use might lead to an under-estimation in the NSS. The omission of receipts by way of gifts since the 20th round could account for the steep drop in reported consumption in the mid sixties. These two factors would probably result in the reported aggregate expenditure on cloth lower than "consumption" (in the usual sense of purchase plus net imports less inventory changes). But it is not obvious why the NSS estimate was much higher than the CSO's in the early sixties nor why they should lead to a divergence of the two series in the earlier half of the period and a convergence in the second half.

Conclusion

27. The above discussion points to the following conclusions on the reliability of NSS Consumption data:

1. It is not possible to say anything definite on the accuracy of the NSS estimate of the level of per capita consumption. While the sampling error seems to be quite small, and the sampling design unbiased the scope for non-sampling errors is considerable. There are indications that NSS generally over-states food grain consumption and that the degree of over-estimation is higher in the upper income group. In the case of clothing it is difficult with available evidence to be sure of the direction of the error. While we have no basis for evaluating the direction of non-sampling errors in respect of other items, the scope for such errors here would seem

to be generally less than in foodgrains and clothing; in any event the NSS data, being based on direct enquiry, would seem superior in many cases to the official estimates. There is no basis to assume that non-sampling errors in individual items will tend to cancel out. Therefore to the extent that a major item like foodgrains is over-stated, the estimate of total consumption is also likely to be overstated. And in so far as the degree of over estimation is greater in the upper income groups, it would also tend to over-state the disparities in consumption.

2. There is however more compelling evidence of systematic changes in the degree of non-sampling errors of estimates relating to foodgrain and clothing arising from identifiable changes in the design of schedules and in concepts. In both cases, there is reason to believe that the official estimates (which point to a mild rising trend) are much more reliable indicators of changes in consumption than the NSS (which shows a significant reduction during the sixties). Consequently even if we assume that the NSS estimates for other items are more reliable than the official series for assessing changes, the NSS series would tend to underestimate the growth (or exaggerate the decline) in percapita total consumption.

3. There is reason to believe that the NSS tends to overestimate the consumption of foodgrains and cloth at the beginning of the period and that the degree of overestimation has fallen during the sixties. The incidence of this phenomenon seems to be much more pronounced in the upper income groups than among lower income sections at least in the case of foodgrains. If this surmise is correct, the NSS would tend to understate the extent of deterioration (or exaggerate the improvement) in inequality in the distribution of consumption.

4. In so far as the NSS understates the growth in average pce, (or exaggerate its decline), it would tend to exaggerate the extent of increase in the incidence of poverty (measured by the proportion of population falling below a specified poverty line). However, its tendency to understate the worsening in inequality (or overstate the improvement) would have the opposite effect. It is not possible to say, with the evidence at hand, what the relative strength of the two effects are.

28. All this suggests the need for far greater awareness among analysts of the various sources of error and non-comparability in NSS data while using them for studying consumption levels, inequalities in living standard, incidence of poverty and changes in all those respects. Given the circumstantial and patchy nature of evidence available, the above inferences are necessarily tentative. But it is hoped that they will provoke persons with first hand knowledge of NSS and its working as well as persons who have made use of the data to corroborate or to join issue with the arguments, evidence and conclusions presented here so that the uses and limitations of NSS consumption data for the above purposes may be better understood by users. Hopefully it will stimulate greater interest in, and concern for, improvements in the design of the schedules, the methods of enquiry and ensuring comparability over time.

Acknowledgement

This is a revised version of a draft prepared some 2 years back. In revising it, I have greatly benefitted from discussions and comments from Nirmal K Chandra, N. Krishnaji, Chandan Mukherjee, T.N. Srinivasan and S.D. Tendulkar. While acknowledging my debt to them, responsibilities for any errors remain mine.

Notes

1. This practice was given up after the NSS was reorganised into a separate, agency under the government. Sampling errors of estimates under the revised system have not been published.
2. For instance the standard error of the 13th round estimate for per capita total consumption in rural India has been calculated at barely 1% for food 1.2% and non-food 2.5%. Gupta and Ramaratnam (1975)
3. More specifically, their conclusion was:

".....for the limited question of estimating the incidence of poverty all one needs to be concerned with is the degree of bias (and its variation over time) at the chosen normative poverty line provided that stated consumption is an increasing function of true consumption is satisfied. But if one is interested in quantifying the magnitude of income transfers from the rich to the poor required to reduce the incidence of poverty, it is essential to know the extent of bias at all levels of consumption" (Srinivasan et.al, 1974:161)

Note that the word "bias" is used here in a loose sense to cover both "sampling bias" and inaccuracies in information.

4. The number of sample households surveyed for consumption has risen substantially. Thus in rural areas it rose from 2600 in 1958-59 to 8400 in 1969-70, 16,800 in 1971-72 and over 1,00,000 in 1972-73. See Gupta and Ramarathnam, 1975:88 for details.
5. For a general description of the evolution of NSS design see Bhattacharyya, (1981). Tendulkar in a personal communication refers to a change in the definition of household: Upto the 25th round, the size of the HH was defined as normal residents plus temporary guests less temporary stay aways. Since the 25th round it is defined as normal residents less temporary guest plus temporary stayaways. Whether this has any effect on the comparability of averages and/or distribution is again unclear.

6. For instance Bhattacharya and Chatterjee (1975:103) argue:

"...the programme of field work of any NSS round is organised in the form of sub-rounds so that the dates of interview for individual sample households are evenly spread over the duration of the round. This eliminates the effect of seasonality almost completely as far as the averages of consumption based on the entire round are concerned. But as different households are interviewed on different dates seasonal variation is superimposed on the true variation between households and the distribution of population by size classes of PCE exaggerates the true extent of inequality. What is more serious, seasonality may be distorting the engel curves and engel elasticities based on NSS budget data".

7. If non-response is systematically more in some classes than others, the estimated average and distribution would be affected. Unfortunately there is little information on the magnitude of non-response or the characteristics of the non-responding households.

7a. Chatterjee and Bhattacharya (1975:105-6) note that there are "a good number of absurd entries" in the schedules, in terms of absurdly low (or high) consumption level and/or prices arising from errors in conversion from local units to standard units, difficulty of finding true measures of local units, confusion between number and length (in the case of some items of clothing) and failure to distinguish between purchase/consumption. They advocate systematic editing and even rejection of "absurd entries" from the estimates.

8. Srinivasan et.al (1974) also attempted to compare the behaviour of NSS series of consumption for a few selected commodities (namely foodgrains, clothing, edible oils and sugar) with a series constructed from official estimates of per capita availability of these items adjusted by their respective wholesale price index. This was necessary because detailed commodity break down of the National accounts estimate of private consumption were not available at that time.

9. The treatment of payment made in kind to priests, barbers, domestic servants barbers etc. for consumer services rendered by them to the sample household is however not simple. Thus Instruction 5.8.1 stipulates that "the imputed value of the items offered should be entered under column 4 of the block 8 against the particular service consumed in cases where the payment is made in terms of blocks 5,6 and 7 (where the concept is that of consumption). When such payments are made in terms of block 8 (miscellaneous goods and services) or 9 (durable goods for domestic use) (where the concept is that of purchase) the payment in kind should not be shown in block 8 against the service consumed. In this case, the value of purchase of items offered (provided such purchase was made in the reference period) should be entered in block 8 or 9 against the items themselves".(Instructions to field staff, 25th round, mimeo 1970).

It may also be noted that household consumption is supposed to exclude

"any transfer payments in kind like loans advanced, charities gifts and other payments in kind.....But any consumption out of transfer receipts in kind like borrowings, gifts, charities, perquisites received by the household, free collections and other receipts in kind, if any, will have to be considered" (Instruction 5.5.6)

10. The 25th round instructions (No.5.6.15) to investigators says:

"Meals received by an employee as perquisite from an employer's household should be accounted for in the employers household in terms of their constituent items like cereals, pulses, vegetables etc. The same meal should also be accounted for in the employee household as an item 96 i.e. cooked meal".

This seems somewhat ambiguous about the treatment of the item in the schedule of paying household. Obviously ingredients of cooked meals given to workers should be netted out in estimating the household consumption of employers in order to avoid double counting.

11. The Instructions in fact do not say whether total consumption of foodgrain is to be estimated independently or as the sum of the purchases, use of own stock, kind receipts for goods and services, and transfers. In either case there are difficulties of ensuring internal consistency in the foodgrain receipts and disposition accounts.
12. The following observations by Chatterjee and Bhattacharya (1975:101), who have intimate knowledge of NSS organisation and procedures, are relevant to this context:

"We feel that employees may not be excluding such meals etc. when reporting their respective household consumption so that there is risk of double counting as well as a tendency to exaggerate the consumption standards of the rural rich. This may be partly responsible for the high figure of cereals consumption obtained from the NSS. Ceremonials may also provide a partial explanation; while the host household reports the entire quantum of cereals needed for the feast in its budget, most, if not all, the invitees may forget that they had a meal outside when reporting for their respective households".

They feel that a similar problem may exist in respect of animal feed where household consumption is supposed to include only what is fed to pet animals but not the amount given to livestock used for productive purposes.

13. In 1960-61, the top decile of the population was estimated to consume 28 kg of cereals per head per month or nearly 1 kg per head; (GOI, p6, 1969:95). The corresponding figure for 1972-73 is estimated at 21 kg.
14. If CSO estimates are "reliable" (or at least as reliable as the NSS), this means that private consumption derived from the former should be generally larger than household consumption obtained from the latter.

While this seems to be the case, the levels of intake of non-cereals items in the top income levels seems implausibly high as can be seen from the following estimate of calorie content of different sources of food in 1972-73.

Per Expenditure Rs. 30/- days	Total	Calorie intake/Cons. Unit/day				1972-73
		I	II	III	IV	V
0-15	1493	1300	51	30	49	63
15-21	1957	1672	81	50	72	82
75-100	4574	2820	316	450	395	596
110+	6181	3167	441	608	620	1145
All	2924	2078	187	217	191	251

- I - Cereal, Cereal substitutes, potato, sugar jaggery
- II - Pulses, nuts and seed
- III - Milk, meat, egg, fish
- IV - Edible oil
- V - Fruits, Vegetable, Processed food

Source: Sarvekshana Vol. II No. 3 January 1979:116-7

Note that in the highest expenditure group, the average daily calorie intake from non-cereal sources is over 3000 per consumption unit; in per capita terms it will be considerably higher. The desirable intake from a nutritional view point, is said to be 2300 calories per capita from all sources.

15. According to Chatterjee and Bhattacharya (1975:99) the incorporation of the consumption questionnaire as part of a comprehensive questionnaire covering productive activity employment and other aspects, was a major change. They also refer to some differences between the Consumption block of IHS and earlier schedules.
16. The following estimates of per capita consumption in 1964-65 based on the detailed and the abridged schedules (IH 15 and IH17 respectively) suggests that aggregation could make some difference:

	RS/30 days Rural	
	IHS16	IHS17
Foodgrain	12.87	12.26
Other food	6.42	5.56
Non-food	7.15	7.01
Total	26.44	24.83

Source: GOI, NSS, 189:8

The sub sample estimates for IHS16 (26.2.-26.66) are closer than for IHS17 (24.2-25.5)

17. On this see among others Vaidyanathan (1978) Sarma et.al (1979)

- Mhluwalia, M.S (1976) Rural Poverty in India 1956-57 to 1973-74, India: Occasional Papers. World Bank Staff Working Paper No.279
- Bardhan P.K (1974) Pattern of Income Distribution in India, in Srinivasan T.N and Bardhan P.K (1974:103-138)
- Chatterjee G.S and Bhattacharya (1975) Some Observations on NSS Household Budget Data in Dandekar V.M and Venkataramiah.P ed (1975:97-113)
- Bhattacharyya A.K. (1981) "Evolution of the Sample Design of the National Sample Survey" Sankhya, Ser C Vol.38
- Dandekar M M and Rath N (1971) Poverty in India (Indian School of Political Economy, Lonavala)
- Dandekar V.M and Venkataramiah.P (eds) (1975) Data Base of Indian Economy II, Role of Sample Survey (Statistical Publishing Society, Calcutta)
- Gupta S.K and Ramaratnan D.S (1975) "Consumer Expenditure Surveys and Family Budget Enquiries", in Dandekar and Venkataramiah eds (1975)
- Ganguly A et.al (1960) Studies on Consumer Behaviour (Asia:Bombay)
- Government of India, National Sample Survey Reports Numbers 101, 104, 135, 142,191, 192, 209, 216,226,230,231,235,240,284 (Draft)
- Government of India, CSO (1976) National Accounts Statistics 1960-61 to 1974-75 (New Delhi)
- Iyengar N S (1967) Some estimates of Engel elasticities based on National Sample Survey Data, JRSS 4,13
- Kansal S.A (1965) Preliminary Estimates of Total Consumption Expenditure in India: 1950-51 to 1963-64. (ISI, memo.)
- Mukherjee M (1969) National Income of India: Trends and Structure (Statistical Publishing Society, Calcutta)
- Mukherjee M and Chatterjee G.S (1972) On the validity of NSS Estimates of Consumption Expenditure, Arthavijnan July
- Murthy M.N (1977) Use of Empirical studies in Evaluating Sample Design for estimating frequency Distribution Full Int.Stat,Inst XLVII,3

- Bhajaravan I (1975) Poverty, Inequality and Economic Growth: Rural Punjab 1960-61 to 1970-71, Journal of Development Studies (July)
- Rudra, A (1972) Relative Rates of growth: Agriculture and Industry", in Prmit Chaudhury ed. Readings in Indian Agricultural Development, (George Allen, and Unwin, London)
- Srinivasan T.N. and Radhakrishna P.N., and Vaidyannathan, . (1972) "Data on Distribution of Consumption Expenditure in India: An Evaluation" in: Srinivasan and Bardhan eds (1972).
- Sarma J.S., Roy Shyamal and George P.S (1979) Two Analysis of Indian Foodgrain Production and Consumption Data (IFPRI, Washington DC)
- Vaidyanathan, J. (1974) Some aspects of Inequalities in Living Standards in rural India, in Srinivasan and Bardhan ed (1974)
- Visaria, Pravin (1980) "Poverty and Living Standards in Asia" Population and Development Review Vol.6 No.2