WORK IN PROGRESS

A NOTE ON HOUSEHOLD INCOME SURVEYS IN INDIA

Aparajita Bakshi Indian Statisitical Institue

STUDYING VILLAGE ECONOMIES IN INDIA A COLLOQUIUM ON METHODOLOGY

December 21 to 24, 2008



A NOTE ON HOUSEHOLD INCOME SURVEYS IN INDIA

Aparajita Bakshi, Indian Statistical Institute

1. INTRODUCTION

Income poverty is a major feature of deprivation and underdevelopment in India and a major cause of lack of social opportunities. Income poverty in India is mainly studied with respect to and in terms of data on consumption, more specifically per capita consumption expenditure. Unlike data on consumption, workforce and land holdings, there are no serial, large-scale, survey based data on household incomes in India.

In many countries of the world including developed countries like United States of America and United Kingdom and less developed countries like Sri Lanka, China and Malaysia, household income data is collected through household surveys. Estimates of household incomes (or income accruing to the household sector) can also be derived from National Accounts Statistics (NAS). The merit of household survey data on incomes compared to NAS based estimates is that household survey data allow for inter-personal comparison of incomes and the analysis of sources and patterns of income generation.

Although India has a long and well-established system of statistical data collection, there are no regular surveys on incomes, rural or urban. The National Sample Survey Organisation (NSSO), which conducts large-scale socio-economic surveys, had made a few attempts to conduct household income surveys. The NSSO could not evolve an acceptable methodology for household income estimation through these surveys. The National Council of Applied Economic Research (NCAER) also conducted some large-scale household income surveys. This paper attempts to describe and critically evaluate the few large-scale surveys on household incomes in India.

2. PROBLEMS OF HOUSEHOLD INCOME SURVEYS

Household incomes are difficult to measure with accuracy because income is a derived variable. Hence it is important to identify and measure each component of income accurately to derive reliable estimates of income. Income estimates from household surveys are thus subject to certain limitations.

The starting point for estimating any parameter is to define the parameter. There is no uniform definition of household income accepted and applied in all countries in the world. The problem of defining household incomes arises from the fact that households produce goods and services for sale in the market as well as for self-consumption. The monetised part of household incomes can be accounted for, but it is often difficult to value the non-monetised part of household incomes. There is no consensus on the treatment of different kinds of monetary and nonmonetary benefits received by households in the estimation of household incomes. Smeeding and Weinberg (2001) provided a framework for estimating household incomes by specifying the different components that constitute household incomes. Their analysis of income survey data from 25 countries reveals that no single income survey collects information on all components of the income summary measure (ibid). International comparisons of household incomes are difficult because of the differences in definitions and methodologies used in income surveys. The Luxembourg Income Study is trying to harmonise income survey data from OECD countries to enable cross-country comparisons (Atkinson 1990, Smeeding and Schmaus 1990). The International Expert (Canberra) Group of Household Income Statistics is trying to move nations to adopt a more uniform and comparable definition of household income.¹

It is a common feature in most countries that income estimates from household surveys are lower than estimates obtained from NAS. At the same time it is difficult to statistically check the divergence in the two estimates because the margin of error varies across income groups and income components. It is generally understood that higher income groups tend to under-report incomes and that income from certain activities are more difficult to measure than others. In a comparison of household income estimates from household surveys and NAS in six developed countries Sawyer (1976)² found that the divergences between the two estimates varied between 15 and 30 per cent. The degree of under-reporting varied between income categories and was high for incomes from investments and self-employment (*ibid*). Anand (1983) compared estimates of personal income and average household income derived from the Post Enumeration Survey in Malaysia for the year 1970 and concluded that household income estimates from NAS by 25 from the Post Enumeration Survey were below the personal income estimates from NAS by 25

¹ Final Report and Recommendations of the International Expert (Canberra) Group of Household Income Statistics (2001)

² Cited in Anand (1983)

to 30 per cent. Atkinson et. al. (1995)³ analysed income survey data from seven countries and compared the data with comparable fiscal measures and concluded that wages and salaries are fairly accurately reported in income survey data, but on the whole, income surveys underreported incomes. Even for developed economies like the United States it was found that income estimated from farm and non-farm sources show wide divergence between estimates derived from the Current Population Survey and those derived from independent fiscal sources (Deaton 1997). Deaton states that it is difficult to obtain accurate estimates of income for self-employed households engaged in agriculture or family business because personal and business incomings and outgoings may be confused. However, according to Deaton, such problems can be minimised by using an accounting framework for estimation of household incomes from self-employed sources (*ibid*).

The third problem that arises in income surveys is that often income estimates derived from surveys are often lower than estimates of consumption, even when national income estimates show high rate of household savings (Deaton 1997). According to Deaton (1997), "although there are often good reasons to doubt the absolute accuracy of national income figures, the fact that surveys repeatedly show large fractions of the poor people dis-saving, and apparently doing so consistently, strongly suggest that the surveys underestimate savings." Anand and Harris (1994) on the other hand do not agree that chronic dis-saving can be reason enough to regard household income estimates as "implausible". Drawing from Fisher's Permanent Income Hypothesis, the authors point out that short-term incomes and consumption are not strictly comparable. Income may well fall below consumption as individuals save or borrow to bridge differences between income and consumption. The authors found from the Sri Lanka Consumer Finance Survey data that if annual incomes are compared with non-durable expenditure data, the bottom 90 per cent of the population showed net dis-saving. When income data were compared with the aggregate of non-durable and durable expenditure data, only the bottom 40 per cent of the population is found to dis-save. When income data is compared with only food expenditure data it was found that only the bottom decile showed negative savings.

4. HOUSEHOLD INCOME SURVEYS IN INDIA

Given the complexity of occupational patterns in the informal economies of the countryside in less developed countries, the conceptual issues involved in the analysis of household incomes is

108

_

³ Cited in Smeeding and Weinberg (2001)

correspondingly complex. In India, these problems are further mediated by specific social conditions – including caste and gender relations and regional and agro-ecological factors – that obtain in the countryside. There are very few studies on household incomes in India primarily because of the dearth of data on incomes. The only nation-wide surveys that have directly collected data on income are those of the National Council of Applied Economic Research (NCAER) and the attempts by National Sample Survey Organisation (NSSO) in the 1960s and the recent Situation Assessment Survey of Farmers in 2002-03.

4.1 NCAER surveys

The major large-scale household income surveys of the NCAER were conducted between 1968 and 1971. The NCAER undertook a study to measure changes in income levels and income distribution and the consequent changes in consumption patterns, investment and saving behaviour of rural households over three consecutive years from 1968-69 to 1970-71 (NCAER 1987). As part of the study, data on household incomes, consumption and savings were collected from 261 villages during three consecutive years between 1968-69 and 1970-71. A sample of 5115 households were selected for the survey though the actual number of households surveyed was much less due to non-response. A multi-stage stratified sampling method was adopted for sample selection (*ibid*).

In 1981-82, the NCAER conducted a resurvey of households, which were surveyed in 1970-71 for a longitudinal study of the same variables that were covered in 1970-71 (NCAER 1987). A total of 5263 households were covered in this survey of which 3299 households belonged to the list of households surveyed in 1970-71. These 3299 households were the households taken up for the longitudinal study to study the changes in income, inter-class mobility, changes in incidence of poverty, consumption levels and patterns and changes in demographic behaviour.

In 1993-94, NCAER conducted another large-scale survey of 35,130 households from 1765 villages in 16 States in India. The data was mainly used for the "Human Development Profile of India" (Shariff 2001). According to Lanjouw and Shariff (2004), "a fairly comprehensive measure of rural household incomes" can be constructed from this data. Unlike other NCAER surveys, data on consumption, savings and investment behaviour were not collected in this survey.

According to Gaiha (1988) income estimates for each of the three years in the 1968-71 NCAER surveys were lower than Central Statistical Organisation (CSO) estimates. A critical evaluation of the NCAER 1993-94 survey data on incomes by Rawal, Swaminathan and Dhar (2008) reveals that detailed data on incomes and costs of production were not collected in the survey. The income estimates were based on either imputation, as in the case of income from cultivation, or on aggregate incomes reported by the household.⁴

4.2 NSSO surveys

The National Sample Survey Organisation conducts quinquennial surveys on consumption expenditure in India but the organisation does not conduct any survey on household incomes at regular intervals. In the initial years of the Organisation, the NSSO had attempted to collect information on household incomes along with the consumption expenditure rounds in 9th (1955) and 15th (1958-59) rounds (NSSO 1993). This was done on an "experimental basis" but "no proper methodology for systematic data collection on household income has yet been evolved" (*ibid*). Data on household incomes are not published in the NSSO reports on consumption expenditure surveys for the 9th and 15th rounds.

Later as part of its Integrated Household Surveys in the 19th (1964-65) and 24th (1968-69) rounds, NSSO collected data on receipts and disbursements that is data on income flows of households. However, these efforts were not continued as it was found that the estimates of income were lower than the estimates of consumption and savings put together (Joshi 1996, NSSO 1993).

The two most important surveys on household incomes conducted by the NSSO are the pilot survey of income, consumption and savings in 1983-84 and the Situation Assessment Survey of Farmer Households (SAS) in 2003.

4.2.1 Pilot survey of income, consumption and savings (1983-84)

In 1983-84 the NSSO attempted a pilot enquiry on household incomes in rural and urban areas in five States (Haryana, Maharashtra, Tamil Nadu, Orissa, Uttar Pradesh) and the metropolitan cities of Delhi and Calcutta. A hundred villages and 80 urban Blocks in the five States were

110

⁴ For a substantive critique of All India Consumer Expenditure Survey of the NCAER in the 1960s that also collected data on incomes, see Bardhan (1974).

covered under this pilot enquiry. The reference period of the survey was January to December 1983. The objective of this enquiry was to "explore the possibility of evolving an operationally feasible and technically sound methodology for collection of data on household incomes through household interview" (Joshi 1996).

The sample households in each village and urban block were divided into three sets. For the first set only income data was collected, for the second set data on consumption and savings were collected while from the third set data on all three variables were collected. The logic behind such a sample design was that data on incomes should be comparable with data on consumption and savings. A total of 24 households were selected from each village and urban Block. Each household was interviewed twice in the year, in June and December, and data on incomes and savings were collected for the periods January to June and July to December. Data on consumption was collected for the last 30 days on the assumption that consumption patterns were stable over the six-month period (NSSO 1993).

The two main findings from the pilot survey were; a) household incomes were under-reported in rural areas, and b) household incomes were lower than the aggregate of consumption and savings.

In all five States it was found that the average annual household income for rural areas was almost half of the corresponding urban incomes. This led the NSS to infer that rural incomes were seriously under-reported in the survey. Household incomes of metropolitan cities were much higher than urban incomes from the States (NSS 1993, Joshi 1996). There were also discrepancies in income estimates obtained from the three sub-sets in rural and urban areas. Though the discrepancies in the three income estimates were lower for rural areas compared to urban areas, income estimates from Set III (income, consumption and savings) were significantly higher than estimates from Set I (income only) for rural areas (NSS 1993, Joshi 1996).

Comparison of income data with that of consumption and savings showed serious discrepancies for rural areas. Income estimates were found to be 30 to 40 per cent lower than the sum of consumption and savings in rural areas. For urban areas also, estimates for consumption and savings were marginally higher than that of incomes but the differences were not statistically significant (NSS 1995, Joshi 1996). In metropolitan areas the estimates were almost equal (see Table 1).

Table 1 Estimates of average income per person, consumption added with savings by place of residence

	Average income		Average consumption and savings	
Place of residence	Set I	Set III	Set II	Set III
Rural	932	1004	1532	1434
Urban (excluding metropolitan cities)	1985	2020	2104	2052
Metropolitan cities	3111	3425	3179	3053

Source: Joshi 1996

To summarise, the NSS pilot survey could not solve the twin problems of underestimation of incomes and generating income estimates consistent with estimates of consumption and savings. Since all household surveys on incomes by the NSSO prior to the Situation Assessment Survey were experimental in nature, these studies have not been discussed and evaluated by the academia. I will focus on two issues in my critique of the pilot survey. The first relates to the measurement of household incomes and income component and the second relates to discrepancies on income and consumption estimates.

Issues in measurement of household incomes. It is a well-known fact that it is difficult to obtain income data for self-employed households and a major portion of rural households in India are self-employed in agriculture. It is important to use an accounting framework to estimate incomes of these households with any degree of precision. Such a framework was not applied in the questionnaire used for the pilot survey. The data collected on agricultural costs are not adequately disaggregated by crop or crop operation to facilitate recall by the respondent. The absence of an accounting framework might affect the accuracy of the estimates.

Discrepancy in income and consumption estimates. The NSSO has been critical of income estimates obtained from household surveys since estimated incomes are often lower than estimates of consumption expenditure and savings. But the consumption data collected in the survey may be inaccurate and affected by seasonality. The NSS Consumption Expenditure Surveys collect consumption expenditure data in four sub-rounds spread across the year to check seasonality in the data. The pilot survey was conducted in June and December and these two months in most parts of India are months that closely follow the harvest months when food stocks and realised incomes from agriculture are high. This is likely to affect the data on consumption in rural India, where consumption patterns and food availability are significantly affected by agricultural

seasonality. Thus, the consumption expenditure data collected in the pilot survey may be inaccurate. Secondly, the pilot survey used a 30 day recall for all items. In recent years, the NSS uses 30 day recall period for non-durable consumption items and 365 day recall period for durable consumption items, medical and educational expenditures for more accurate reporting of consumption expenditure. In short, the method of estimating consumption expenditure was different in the pilot survey and the regular NSS surveys.

There are also some conceptual problem in equating household incomes with consumption and savings. First, accrual of incomes and disbursal of incomes as consumption and savings are not synchronous processes. Consumption and savings in period t are both functions of income in period (t-1). Secondly, in situations where consumption is financed through borrowings the sum of consumption and savings will overestimate incomes, unless we consider 'net savings', that is, savings net of borrowings. Anand and Hariss (1994) have also shown that short-term incomes are closer to estimates of food expenditure than total consumption expenditure.

4.2.2 Situation assessment survey of farmer households (2003)

The most comprehensive and large-scale attempt by the NSS to collect data on household incomes through direct enquiry was as part of the Situation Assessment Survey of Farmer Households (SAS). In 2003 the NSSO conducted the SAS as part of the NSS 59th round. This survey concentrated on many issues related to farming households in India, including levels of income, consumption, farming practices, indebtedness, access to modern technology in farming, ownership of productive assets and livestock. Surveys were conducted in 51,105 households in 6634 villages in all States and Union Territories in India (NSSO 2005).

The SAS was limited to farmer households only. Farmer households are a small section of the total households dependent on agriculture because agricultural labour households and households earning rental earnings from leased out land are not included in the sample. Nevertheless, this is the single largest data source on rural household incomes in India at present.

The SAS income survey was no doubt a major improvement on previous efforts by the NSSO in terms of the survey methodology used. The reference period coincided with the agricultural year July 2002 to June 2003. The survey was conducted in two phases from January to December 2003. Each household in the sample was visited twice. The first phase of the survey was

conducted between January and August 2003 and information on cultivation of Kharif crops was collected in this phase. The second round visits to households were conducted in September to December 2003 to collect information on Rabi crops.

Issues in definition and measurement of household incomes. The SAS questionnaire was an improvement on the NSS pilot survey on that it used an accounting framework for deriving incomes from self-employment activities such as cultivation, animal husbandry and non-farm business. However, there may be errors in data quality on account of the following. First, costs on labour and other inputs are disaggregated by crops and not by crop operations or type of input used. This may lead to recall errors by farmers. In most cases, farmers have a tendency to overstate costs, which can be checked to some extent by disaggregating the cost components.

Secondly, the SAS used different recall periods for different kinds of occupations. The recall period for farming was an agricultural season, recall for non-farm business was 30 days and for wages and salaries it was seven days. The appropriateness of seven day recall for wages and salaries instead of 30 days may be questioned as salaries are generally received on a monthly basis. Wage employment is highly seasonal and uncertain in character. In spite of the fact that the survey was carried out over a year, which would reduce seasonal variations in the data, a 30 day recall may provide more stable data on wages especially when the objective is to estimate annual household incomes.

Rental incomes from agricultural land and machinery are considered as non-farm business income in the SAS and income estimates from these components are based on a 30 day recall. A 30 day recall will certainly underestimate incomes from these sources as rental incomes from agricultural land and machinery are received for an agricultural season or specific operations in each season. Like agricultural incomes, these types of rental incomes should also have a recall period corresponding to the previous agricultural season.

Thirdly, income definition used by the NSSO has two limitations.

 a. Income from cultivation does not strictly adhere to any commonly accepted definition of farm business income of the Commission for Agricultural Costs and Prices (CACP).
 Costs of owned animal labour, depreciation of agricultural machinery are not included in agricultural costs. b. Transfer earnings from remittances, pensions and scholarships are not included in the questionnaire. Interest earnings are also not included.

Differences in levels of income and consumption. The income data from SAS show discrepancies between income and consumption levels, even though data on incomes and consumption were collected in the same questionnaire. Average consumption expenditure per farmer household is significantly higher than income in all States in India, except Assam, Jammu and Kashmir and Jharkhand (Bhalla 2006, NSSO 2005). Bhalla (2006) has also shown that in 14 out of the 18 Indian States the income of farmer households owning up to 2 hectares of land was insufficient to meet consumption needs.

Unlike the pilot survey, the SAS did not use uniform 30 day recall period for all consumption items. The survey was conducted in two phases all through the year and thus consumption estimates would be less affected by seasonality than the pilot study. In fact, average monthly household consumption expenditure of farmer households was found to be the same as that of all rural households obtained from consumption expenditure survey in the same year (NSSO 2005). The SAS experience suggests that even when seasonality is checked and consumption data is collected together with income data, income estimates are still lower than consumption. As I have discussed in the previous section, this is a phenomenon observed in many income surveys in the world and there should not be a conceptual problem in this. Anand and Harris (1994) have shown that conceptually and empirically short-term income data is closer to data on food expenditure rather than total consumption expenditure. The published reports on SAS do not provide separate estimates for food expenditure, so it is not possible to make a comparison at this point. However, this issue can be taken up for further analysis with unit level data from the SAS.

4. CONCLUSION

There are a limited number of studies on rural incomes in India because of limited data. Most of the literature on rural incomes and income diversification is based on micro-level village studies. There are no serial large-scale household surveys on incomes in India. Some by government and non-government agencies to collect large-scale data on incomes through household surveys have failed to generate consistent and reliable estimates of household income.

The general criticism of household survey data on incomes in India focus on two issues. The first criticism relates to under-reporting of rural incomes in household surveys. The estimates of household incomes obtained from the NCAER surveys in 1968-71 were lower than estimates of personal incomes in National Accounts Statistics. The income estimates from the NSSO pilot survey were lower than the aggregate of household consumption and savings, specifically for rural areas. Thus it was concluded that income estimates from household surveys under-estimate rural incomes.

The second criticism is, in both the NSSO surveys (the pilot study and the SAS) estimated incomes were lower than consumption expenditure. Even though it is not explicitly stated in any government source, it seems that the reluctance of government agencies to conduct income surveys in India rest on these two basic methodological issues.

Data from National Accounts Statistics (NAS) are not sacrosanct as the estimates are derived from many sources and are often subjective in nature.⁵ It is not easy to locate the reason and origin of discrepancies in income estimates from survey data and NAS. But it is not reasonable to totally reject the estimates obtained from household surveys because of inconsistencies with NAS. Anand (1983), in his study on Malaysia compares the income estimates from Post Enumeration Survey (PES) in 1970 with the personal income estimates in NAS and states that "the discrepancy in estimates from the two sources is not easy to resolve, but it cannot be taken as a reason for rejecting the Post Enumeration Survey."⁶

The problem of under-estimation of household incomes through household surveys is not specific to India, it is a reality observed in most countries of the world including developed countries. Response errors and underestimation of household incomes can be minimised through meticulous design of questionnaire. Income is a derived variable. If all components and sources of income are identified accurately and an accounting framework is applied to estimate incomes from each source, fairly accurate estimates of income can be made. The accounting framework should disaggregate the items of costs and receipts in a way so as to facilitate recall and minimise chances of under-reporting.

⁵ See Rudra (1961) and Minhas (1988)

_

⁶ PES was conducted in Malaysia after the 1970 Census of Population and Housing to check under coverage and content of Census. Information on household incomes was also collected in this survey.

Discrepancies in income and consumption estimates from survey data should also not be taken as a reason to reject income surveys. The common assumption that large sections of the population cannot dis-save when the macro-economic indicators show positive savings rate for the economy as a whole may not hold true because there are definitional, conceptual and temporal differences in the variables.

Income surveys serve a purpose of their own in the understanding of important issues of income generation, income diversification and poverty. Aggregate data on personal incomes from the NAS or household survey data on consumption are not adequate to understand these issues since such data are collected with the only objective of quantification of levels of incomes and poverty. Thus, according to me, collection of household level data on incomes is of utmost importance and useful in the understanding of the dynamics of the household sector and income-poverty linkages, in spite of problems of estimation and consistency.

References

Anand, Sudhir (1983), Inequality and Poverty in Malaysia, Oxford University Press.

Anand, Sudhir and Hariss, Christopher J. (1994), Choosing a Welfare Indicator, American Economic Review, Vol. 84(2).

Atkinson, A. B. (1990), Introduction, in Smeeding, T. M. (ed.) *Poverty, Inequality and Income Distribution in Comparative Perspective: The Luxemberg Income Study, Harvester Wheatsheaf*, New York.

Atkinson, A. B., Rainwater, L. and Smeeding, T. M. (1995), *Income Dsitribution in OECD Countries: The Evidence from Luxembourg Income Study*, Organization for Economic Cooperation and Development, Paris.

Bardhan, P. K. (1974), The pattern of Income Distribution in India: a review, in T. N. Srinivasan and P. K. Bardhan (eds.) *Poverty and Income Distribution in India*, Statistical Publishing Society, Calcutta.

Bhalla, G. S (2006), Conditon of the Indian Peasantry, National Book Trust, India.

Deaton, Angus (1997), The Analysis of Household Surveys: a micro econometric approach to development policy, World Bank, Washington.

Expert Group of Household Income Statistics (2001), *Final Report and Recommendations*, available at: http://www.lisproject.org/links/canberra/finalreport.pdf.

Gaiha, Raghav (1988), Income Mobility in Rural India, Economic Development and Cultural Change, Vol. 36(2)

Joshi, P. D. (1996), *Household Expenditure and Income Surveys – An Indian Experience*, paper presented at Expert Group on Household Income Statistics, Canberra, 2-4 December. Available at: http://www.lisproject.org/links/canberra/canberrareport/joshi.htm.

Lanjouw, Peter and Shariff, Abusaleh (2004), Rural Non-farm Employment in India: Access, Incomes and Poverty Impacts, *Economic and Political Weekly*, Vol. 39(40).

Minhas, B. S. (1988), Validation of Large Scale Sample Survey Data: Case of NSS Estimates of Household Consumption Expenditure, *Sankhya: The Indian Journal of Statistics*, Vol. 50, Series B, Pt. 3.

NCAER (1987), Changes in the Structure of Household Income and Distribution of Gains in the Rural Household Sector – An All India Temporal Analysis, 1970-71 – 1981-82.

NSSO (1993), Report on the Pilot Survey of Income, Consumption and Savings, Part I and II, Department of Statistics, Delhi.

NSSO (2005), *Income, Expenditure and Productive Assets of Farmer Households*, Ministry of Statistics and Programme Implementation, Government of India.

Rawal, Vikas, Swaminathan, Madhura and Dhar, Niladri Sekhar (2008), On Diversification of Rural Incomes: A View from Three Villages of Andhra Pradesh, *Indian Journal of Labour Economics*, Vol. 51(2).

Rudra, A. (1961), National Income Estimates - Why not Discontinue them? Economic Weekly, Feb 4.

Sawyer, Malcolm (1976), Income Distribution in OECD Countries, OECD Economic Outlook, No. 19.

Shariff (2001), India human development report - A profile of Indian states in the 1990s, NCAER, Delhi.

Smeeding, T. M. and Schmaus, Gunther (1990), The LIS Database: Technical and Methodological Aspects, in Smeeding, T. M. (ed.) *Poverty, Inequality and Income Distribution in Comparative Perspective: The Luxemberg Income Study,* Harvester Wheatsheaf, New York.

Smeeding, Timothy M and Weinberg, Daniel H. (2001), Toward a Uniform Definition of Household Income, *The Review of Income and Wealth*, Vol. 47(1).