Paper Prepared for the Special IARIW-SAIM Conference on
"Measuring the Informal Economy in Developing Countries"

Kathmandu, Nepal, September 23-26, 2009

Surveys of Informal Sector Enterprises – Some Measurement Issues

Kaushal Joshi, Rana Hasan and Glenita Amoranto, Asian Development Bank

This paper is posted on the following website: http://www.iariw.org

Surveys of Informal Sector Enterprises – Some Measurement Issues

Kaushal Joshi, Rana Hasan and Glenita Amoranto[1]

1. Introduction

The informal sector represents an important part of the economy and the labor market in many countries, especially developing countries, and plays a major role in employment creation, production and income generation (OECD 2002). The informal sector as defined in the resolution of the fifteenth International Conference of Labor Statisticians, January 1993, refers to economic activities i.e., production and distribution of goods and services by the operating units of the households which essentially differ from the formal sector in terms of technology, economies of scale, use of labor intensive processes, and virtual absence of well maintained accounts. Variety of terms have been in vogue within the administrative setup and statistical systems of countries to describe enterprises satisfying one or more similar characteristics, such as "unregistered", "unorganized", micro-enterprises etc.

The informal sector represents a substantial portion of economic activity, especially in developing and transition countries. Estimates (Charmes 2000) show that the sector accounts for more than two third of total employment and more than one third of the total GDP of the non-agricultural sector in Asia. In view of its estimated size the sector invites high policy interest in many parts of the world. Given their high potential for job creation and income generation in developing economies, the informal sector is gaining attention of policy makers. Therefore data on various characteristics and operations of these enterprises, output generated, and employment provided, constraints faced and their relationship with the formal sector are needed. Unfortunately, due to their very nature of smallness, invisibility, high rates of entry and exits, informal sector enterprises do not enter the business registers or the list frames usually maintained by national statistical offices (NSOs). Thus, many enterprise surveys conducted by NSOs in developing countries usually target only enterprises beyond a threshold, generally measured in terms of size of employment, which are available in the official lists. As a result informal sector enterprises escape official data collection systems and their contribution to the GDP is often understated in the official national accounts data despite their significant contribution.

In this paper we will be concerned with some measurement issues that are faced in the process of collection of data from these enterprises through sample surveys. We will be using the terms informal and unorganized interchangeably in this paper. One of the reasons for using the term unorganized is that the data that we use to look into some of the measurement issues and derive our conclusions relates to the NSS 56th round survey of unorganized manufacturing sector of India.

2. Measurement of Informal Sector

Several problems are associated with the collection of data on informal enterprises for use by statisticians, economists, researchers and other users to answer a variety of questions such as

---

[1] The authors are working respectively as Statistician, Principal Economist and Consultant with the Asian Development Bank. The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the Asian Development Bank.

their contribution to the level and growth of real GDP, employment and wages, impact on poverty and inequality and other questions of interest. Accurate measurement of various characteristics, including flow and stock variables is therefore critical. Most of these enterprises are household enterprises where money and goods are fungible between household and enterprise. Further, absence of written records of transactions leads to relying on the recall by the informant. Designing appropriate methods to capture accurate data is a challenging task for the survey statisticians. Additionally, there are issues relating to deliberate misreporting as the owners may be reluctant to reveal their accurate incomes and expenses fearing the use of information for tax purposes.

The method of measuring the informal sector depends upon what questions the users of data collected want to answer. A household labor force enquiry is useful if the main interest is measuring the employment in informal sector. However, if the need is to study the production, size of employment and condition of employment, goods and services produced, raw materials and inputs used in production, fixed assets and capital, credit and interest, relationship with the formal sector etc. then a survey of enterprises is required. "In this case, there are two basic survey design options, namely an enterprise survey or a mixed household-enterprise survey. The choice depends upon data requirements, the organization of statistical systems, and the resources available."(for more details see OECD2002). Such a survey will seek to collect data from the enterprise owners as respondents. Apart from the objectives of the enquiry, the data collection strategy and inclusion of data items in the survey instruments depends on many other factors. It depends on the relative amounts of sampling and non-sampling errors that would creep in any survey design and in survey operations, and many times on the availability of financial resources. NSOs would like to undertake nation wide sample surveys but are very often constrained by resources and consequently the design is guided by availability of resources both financial and skilled manpower.

The sampling errors would depend upon the sampling design and sample size, the non-sampling errors would depend to a large extent on the design of survey instruments, data items included number of visits and reference period (length of recall), availability of records etc. While designing a survey questionnaire the options could be (i) using detailed set of questions to collect data at disaggregated item level, or (ii) using a short questionnaire with very few direct questions. While there has been experimental research in developing countries to compare the two approaches to capture household expenditure, and which suggests that asking more detailed questions leads to more accurate estimates of household consumption (Deaton and Grosh 2001) such experiments for surveys of enterprises are rather limited.

Further, data for the reference period could be collected at a single point of time or at multiple points of time through interviews by repeated visits to the enterprise. Each of these methods have implications on costs and on the sampling and non sampling errors for a given sample size. "If resources for investigation are fixed, increasing the frequency of interviews will necessitate reducing the sample size and consequently tend to increase the sampling error. On the other hand, reducing the frequency of visits may tend to increase the amount of nonsampling errors, such as those due to measurement and response inaccuracies, particularly if significant amounts of memory recall are involved" (Liedholm, 1991, GEMINI). The larger the reference period, the more likely the enquiry to be affected by recall errors. Casey and Lury (1981) contend that, in

developing countries, non-sampling errors are relatively more important than the sampling ones. Reinterview studies have shown the presence of "alarmingly high levels of response errors even on the simplest of survey questions" (Scott, 1985, p. 15), and in some Indian surveys non-sampling errors were probably six times the sampling errors. (Casey and Lury, 1981,p. 87)[2].

In this paper we will examine the differences in the measures of (i) profits of an enterprise obtained from a detailed set of questions on incomes and expenses of the enterprise vs profits of the enterprise obtained through a single direct question and (ii) gross value added (GVA) obtained using *production approach* as difference of output and intermediate consumption of the enterprises from the detailed set of questions on incomes and expenses vs GVA from *income approach* obtained by asking few direct questions on the factor incomes of the enterprises of which profits of the enterprise is an important factor. To examine this we use the data from the 56th round survey of unorganized manufacturing conducted by the National Sample Survey Organization (NSSO) conducted during the period July 2000-June 2001. We also examine if the differences vary with the characteristics of the enterprises and suggest further empirical research to develop suitable tools for providing accurate measurements of informal sector enterprises.

3. NSS Surveys on Unorganized Manufacturing
We use the data from the fifty-sixth round of National Sample Survey (NSS) of India dedicated for collection of data on economic and operational characteristics of small manufacturing enterprises in the unorganized sector of the Indian economy conducted during the period July 2000-June 2001. Unorganized enterprises in India are mainly owned by household proprietary or partnership enterprises and belong to the unregistered factory sector. The registered factory sector is covered annually in a survey Annual Survey of Industries for manufacturing enterprises, registered under Sections 2m (i) and 2m (ii) of the Factories Act, 1948 (i.e. enterprises employing 10 or more workers using power and those employing 20 or more workers without using power). Surveys of unorganized manufacturing enterprises are conducted once in 5 years to meet the data gaps from this sector especially for national accounts purposes to estimate their contribution to GDP through indirect methods, using the benchmark-indicator procedure. In this procedure, the benchmark gross value added (GVA) estimates are initially prepared at detailed economic activity level for the base year of national accounts series as a product of estimated workforce engaged in the unorganized manufacturing and the value added per worker (VAPW) obtained through the quinequenniel survey. For subsequent years, the procedure followed for estimating the GVA annually, is to extrapolate the base year GVA estimates with the growth observed in the index of industrial production (IIP), at each detailed industry level. For the current price estimates of GVA for unregistered manufacturing, the relevant wholesale price indices (WPIs) are super-imposed on the constant price GVA estimates, at detailed industry level. The share of unorganized manufacturing sector in the NDP of the manufacturing sector was 37.5% in the year 1999 which came down to 32.1% in 2005-06. Given that the survey is used to provide benchmark data for contribution of unorganized manufacturing sector to the Indian economy, accuracy of data collected becomes very important.

4. Data Description
Data in the survey were collected for a reference period of one month, from more than 150000 enterprises in rural and urban India based on oral enquiry. From the unit level data, for our

---

[2] Scott, 1985, p15 and Casey and Lury, 1981, p.87 as quoted in GEMINI, 1991, p. 3

intended analysis we have dropped the enterprises with 20 or more workers for controlling the results that may be affected by outliers as some of the enterprises in the original sample were quite big even with more than 100 worker. A few records were corrected by supplying missing values based on other characteristics of the enterprise, and some records with unacceptable entries against some enterprise characteristics were dropped. From the remaining set we further dropped those enterprises for which the percentage difference in the profits obtained for an enterprise based on the two approaches (we describe later) was more than or equal to five hundred. This left us with a sample of 150775 enterprises. Basic characteristics of the sample used in our analysis are given in Table1.

Table 1: Summary Statistics of the Sample

	Rural	**Urban**	**Total**
Number of enterprises	59,797	90,978	150,775
Mean total workers	2.41	3.02	2.78
Mean hired workers	0.72	1.37	1.11
No. of Proprietary enterprises	97.7%	95.4%	96.3%
Male owned	79.0%	75.2%	76.7%
Female owned	18.7%	20.3%	19.6%
Own account enterprises	78.9%	59.4%	67.1%
Enterprise maintaining business accounts	2.04%	5.07%	3.87%
Enterprises located within household premises	64.0%	46.8%	53.6%
Enterprises registered with any local authority	12.5%	26.9%	21.2%
Mean total receipts (output) (in INR)	137,906	308,679	240,951
Mean total expenses (in INR)	107,904	255,535	196,984
Mean intermediate consumption (in INR)	94,054	219,015	169,456
Mean compensation of employees (in INR)	11,392	30,504	22,924
Mean profits(derived) (in INR)	30,003	53,144	43,966
Mean profits(direct) (in INR)	28,049	49,096	40,749
Mean GVA(P) (in INR)	43,852	89,664	71,495
Mean GVA(I) (in INR)	41,898	85,615	68,277

Table 1 shows that majority of enterprises in the unorganized manufacturing segment are very small in size and do not maintain books of accounts. In the sample 67.1% enterprises had no hired labor and the mean employment was 2.78 with hired employment being 1.11 per enterprise. Less than 4.0% enterprises maintained books of accounts and more than half of the enterprises operated from within the household premises.

5. 56th Round NSS Questionnaire
As stated before, the survey is main instrument for estimating the GVA per worker for unorganized manufacturing activity in India. Box 1 provides the details of items of data that were captured in the 56th round survey to calculate the GVA.

Box 1: Data Items in the Survey Questionnaire to calculate GVA

Total Intermediate Consumption	Total Output

• Total Raw materials consumed (data collected for 5 major raw materials and others during the reference period)
• Commodities purchased for resale
• Other Operating expenses
  • Electricity, fuel and lubricants
  • Raw materials consumed for own construction
  • Minor repair & maintenance of fixed assets
  • Rent on P&M
  • Service charges paid
  • Travelling, freight and cartage expenses
  • Communication expenses
  • Consumables, packing materials etc.
  • Paper and printing
  • Licence fees and local taxes (exclude indirect taxes)
  • Other expenses (consumer entertainment, performing rituals, etc.)
  • Distributive expenses
	• Receipts from goods manufactured (value of finished five major products and by products and others during the reference period)
	• Change in Stock of Semi-finished goods
	  • Opening stock of semi-finished goods
	  • Closing stock of semi-finished goods
	• Receipts from commodities traded
	• Change in stock of trading goods
	  • Opening stock of trading goods
	  • Closing stock of trading goods
	• Other receipts (from any other economic activity)
	• Other Receipts
	  • Commission charges received
	  • Market value of own construction
	  • Goods/services produced or traded goods used for own or employees consumption
	  • Rent receivable on P&M and other fixed assets
	  • Donations received (including govt. grants)
	  • Other receipts (incentives received by enterprise, scrap sale receipts, etc)

The data collected on above items from an enterprise enables calculation of Gross Value Added (GVA) by using the production approach as below:

GVA(P) = Total Output – Total Intermediate Consumption……………………………...(I)

where intermediate consumption are the products used in the production of final products of the manufacturing unit and lose their identity in the final product. Total output relates to the value of goods manufactured against the intermediate products consumed in the production process during the reference period of one month.

The NSS questionnaire also provides items to collect data on factor payments such as compensation paid to the employees (salary, wages and other benefits in cash & kind) during the reference period, rent payable on fixed assets and interest payable on loans outstanding. Given this information, the profits of an enterprise can be derived using the following identity:

Profits(derived) = Total Output – Total Intermediate Consumption (as above) –
Compensation of employees – rent payable – interest payable ……………………………(II)

Or

Profits(derived) = Total Output – Total Expenses……………………………………(II-A)

The NSS questionnaire seeks sufficiently detailed data on expenses and receipts of an enterprise. There were 37 items of expenditure and 18 items of receipts or incomes of enterprise in the questionnaire. Experiments by de Mel, Mckenzie and Woodruff (2008) in experimental surveys of micro-enterprises in Sri Lanka have shown that data on receipts and expenses of micro-enterprises collected based on interview method tend to improve as more and more detailed questions are asked to adjust for unreported categories such as, business goods used for home consumption, business revenue used for household expenses but not included in revenues (or in profits), firms inputs received as gifts but included as business expenses, which reflect the fungibility of resources between the household and business.

Upto its last survey on the unorganized manufacturing sector in the 51st round in 1994-95, NSS collected data for calculating the GVA at the enterprise level using the production approach. In the 56th round survey however, it included some additional question to collect direct data on profits of each enterprise. A new data item was included – *net surplus* (including home consumption of raw materials/goods/services produced or traded by enterprise). *Net surplus* of an enterprise was defined as the amount which the owner / partner(s) get out of the entrepreneurial activity after making payments to the workers (individual or as group benefits), rent on land and building and interest on outstanding loan for the reference month. For getting a complete idea of the net surplus or earning of an enterprise, withdrawals (if any) from enterprise earnings by the household or partners and home consumption during the reference month, other than the surplus visible at the end of the reference month were taken into account. Payments made to paid family members who are treated as hired workers of the enterprise, were excluded from the net surplus and included under the compensation paid to employees.

This additional data item to capture profits i.e. profits(direct), of the enterprise provided additional measure of calculating the GVA using the *income approach* by simple addition of factor incomes of the enterprise during the reference month as follows:

GVA(I) = Net surplus + Interest + Rent + Compensation to employees…………………(III)

Thus while profits could also be derived using (II-A) from the already collected detailed data items on receipts and expenses of an enterprise, the direct question on *net surplus* provided another value for the profits made by the enterprise during the reference period. The 56th round questionnaire thus provided for getting the GVA and profits of the same enterprise following two different approaches, which are summarised in Box 2.

Box 2: Approaches to measuring GVA and profit of an enterprise

Approach -1 (Production Approach)	Approach -2 (Income Approach)
GVA(P) = Gross Value Added = Total Output – Total Intermediate consumption	GVA(I) = Net Surplus + Interest + Rent + Compensation of Employees
Profits(derived) = Total Receipts – Total Expenses	Profits(direct) = Net surplus

In Approach 1, profits and GVA of an enterprise can be calculated by using detailed data collected on inputs, other operating expenses, expenses on salaries and wages, interest, and rent of the enterprise. There were 37 items of expenditure and 18 items of receipts or incomes of enterprise in the questionnaire. On the other hand, Approach 2 captures profits (net surplus) by a direct question and the GVA by adding only four data items Net surplus, rent, interest and compensation of employees to calculate GVA(I). In Approach 1 data on Interest, Rent and Compensation of Employees is used to calculate the profits(derived) of the enterprise, whereas in Approach 2, these items are used to calculate GVA(I). The 56th round survey provides an opportunity to compare the data collected from the two approaches for the same set of enterprises.

6. Results of Data Review

In the following sections using the data of the NSS 56th round survey of 59797 rural enterprises and 90978 urban enterprises in the unorganized manufacturing sector we examine some of results based on the two approaches and compare the results. We will define:

Gap(profits) = [Profits(direct) – Profits(derived)]/ mod [Profits(derived)] *100
Gap(GVA) = [GVA (I) – GVA(P)]/ mod [GVA (P)] *100

There were a few observations with zero profits(derived) and GVA(P) affecting the denominator, these records were dropped from the calculations. It was also observed that almost all such enterprises had zero revenues and expenses and their impact on overall results is expected to be negligible.

Gap(profits) and Gap(GVA) measure the percentage difference between estimates obtained from two approaches. A negative Gap indicates that the profits or GVA of the enterprise based on Approach 2 (direct question on profits) is lower than the profits or GVA based on Approach 1 (detailed disaggregated questions on incomes and expenses).

Table 2A: Mean GVA(P), GVA(I) and Gap(GVA)

		All	Sector		Whether maintaining accounts or not?	
			Urban	Rural	Yes	No
GVA (P)	Mean	71,495	89,664	43,852	357,477	59,992
	S.D.	222,003	227,841	209,820	856,171	135,524
	Median	25,944	37,680	16,740	207,966	24,660
	No. of obs	150,775	90,978	59,797	5,830	144,945
	% negative	0.30	0.29	0.33	0.69	0.29
GVA (I)	Mean	68,277	85,615	41,898	330,733	57,721
	S.D.	186,025	189,802	176,898	682,071	120,026
	Median	25,200	36,000	16,800	201,438	24,000
	No. of obs	150,775	90,978	59,797	5,830	144,945
Correlations	Pearson	0.9021	0.9032	0.8974	0.8605	0.9471
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000
	Spearman	0.9893**	0.9897**	0.9864**	0.9724**	0.9897**
Gap(GVA) (%)			**-4.5**	**-4.5**	**-7.5**	**-3.8**

Table 2B: Mean Profits(derived), Profits(direct) and Gap(profits)

|		All	Sector		Whether maintaining accounts or not	
			Urban	Rural	Yes	No
Profits(derived)	Mean	43,966	53,144	30,003	204,453	37,511
	S.D.	179,945	181,875	176,055	773,588	92,405
	Median	22,320	29,700	15,480	95,154	21,408
	No. of obs	150,775	90,978	59,797	5,830	144,945
	% negative	0.70	0.71	0.68	2.21	0.64
Profits(direct)	Mean	40,749	49,096	28,049	177,708	35,240
	S.D.	137,228	135,361	139,065	581,112	72,275
	Median	21,600	30,000	15,000	90,000	21,480
	No. of obs	150,775	90,978	59,797	5,830	144,945
Correlations	Pearson	0.8471	0.8445	0.8503	0.8263	0.8835
	p-value	0.0000	0.0000	0.0000	0.0000	0.0000
	Spearman	0.9794**	0.9760**	0.9818**	0.9416**	0.9810**
Gap(profits) (%)		**-7.3**	**-7.6**	**-6.5**	**-13.1**	**-6.1**

Table 2A reports the mean, standard deviation, median, of GVA(P) and GVA(I) calculated using the two approaches separately for rural and urban enterprises. It also gives percentage of enterprises with negative GVA and correlations. Table 2B reports similar results from the two approaches for profits(derived) and profits(direct). The two tables show that the observed correlations are very high, although on an average across rural and urban sectors, GVA(I) and profits(direct) which are based on one shot question on net surplus are lower than GVA(P) and profits(derived). At the aggregate level mean GVA(I) is about 4.5% lower than mean GVA(P), and mean profits(direct) are lower by 7.3% as compared to mean profits(derived). While there is no rural urban differential in the Gap(GVA), based on the two approaches, the Gap(profits) is much higher for urban enterprises than the rural. The differences in the two approaches do not appear to be very high at the aggregate level, yet Approach 2 to obtain profits and therefore GVA through a direct question on profits, yielded lower estimates of mean profits and consequently lower estimates for mean GVA compared with Approach 1. We further examine this issue and look into the estimates of the two variables across various enterprise characteristics which are discussed below.

6.1 Books of Accounts
The unorganized sector enterprises in India usually do not keep books of accounts and therefore information is collected through oral enquiry depending to a large extent upon the recall of the informant. Less than 4.0% of the enterprises in the entire sample maintained books of accounts and provided the data from them. As the data collected for enterprises maintaining books of accounts was based on written records, it is expected to be free from recall errors or errors of deliberate under or over reporting on the part of the respondent which are very likely in an oral inquiry. However, the enterprises maintaining books of accounts were also asked direct one shot question on profits. Table 2A reports the mean, standard deviation, median, of GVA(P), GVA(I) and Table 2B reports mean, standard deviation, median, profits(derived) and profits(direct) separately for the enterprises based on accounts maintained. It is observed that, the mean profits(derived) and the GVA(P) were higher than the mean profits(direct) and GVA(I) for

enterprises irrespective of whether the accounts are maintained or not. However, Gap(profits) was much higher at (-)13.1% for enterprises with books of accounts (record based enquiry), while this gap was only (-)6.1% for enterprises with no books of accounts (oral enquiry). This implies that when the direct question on profits was asked, enterprises with books of accounts reported profits which were lower by 13.1% from the profits derived as a difference of incomes and expenses using the books of accounts. The mean Gap(GVA) was (-)7.5% for enterprises with accounts as compared to (-)3.8% for enterprises without accounts (oral enquiry). This gives evidence to the argument that enterprises tend to underreport profits if asked directly, even if they maintain books of accounts. (see Figure 1)

Figure 1
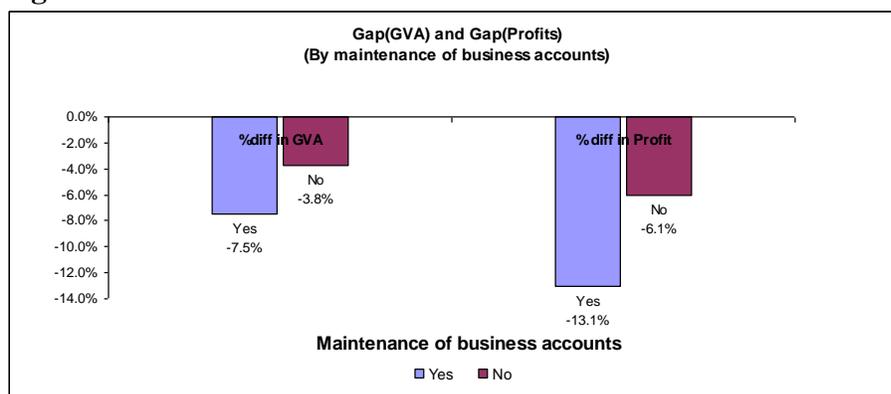
Enterprises that maintained books of accounts were relatively larger enterprises. They have higher number of workers, are more likely to be registered, located outside the household premises and with much higher expenses, receipts, profits and GVA (see Table 3). It is seen that the mean receipts, expenses, profits(derived), profits(direct), GVA(P), GVA(I) and mean value of P&M are almost 80-90% lower for enterprises that do not maintain books of accounts. Obviously, the larger the operations of an enterprise it is more likely the accounts are maintained to track the transactions.

**Table 3: Basic characteristics of enterprises by Accounts maintained or not**

	Accounts maintained?		Total	% Difference between Yes and No
	Yes	**No**		
Number of enterprises	5,830	144,945	150,775	
Mean total workers	6.1	2.6	2.8	
Mean hired workers	4.3	1.0	1.1	
No. of Proprietary enterprises	77.7%	97.1%	96.3%	
Male owned	67.3%	77.0%	76.7%	
Female owned	10.4%	20.0%	19.6%	
Own account enterprises	15.3%	69.2%	67.1%	
Enterprises located within household premises	17.6%	55.2%	53.7%	

Enterprises registered with any local authority	76.6%	19.0%	21.2%	
Mean total receipts (output)	1,778,602	179,103	240,951	-89.9
Mean total expenses	1,574,149	141,592	196,984	-91.0
Mean intermediate consumption	1,421,124	119,111	169,456	-91.6
Mean compensation of employees	118,739	19,070	22,924	-83.9
Mean Profits(derived)	204,453	37,511	43,966	-81.7
Mean Profits(direct)	177,708	35,240	40,749	-80.2
Mean GVA(P)	357,477	59,992	71,495	-83.2
Mean GVA(I)	330,733	57,721	68,277	-82.5
Mean size of plant and machinery	183,950	19,680	26,032	-89.3

6.2 Response Code

In any survey, the quality of data reported depends upon a very large extent the type of response of the informant. NSS captures this information through respondent codes: informant (i) cooperative & capable (ii) cooperative but not capable (iii) busy (iv) reluctant and (v) others. Out of all the enterprises in the sample 78.7% of the enterprises were coded as informant cooperative & capable, which is quite an encouraging number, 16.8% were cooperative but not capable, 1.8% were coded as busy and only 2.4% were reported as reluctant. Although the numbers appear to be quite encouraging, none of these codes would capture if there is any deliberate misreporting by the respondents. Figure 2 shows that while all respondents on an average reported lower profits(direct), the largest Gap(profits) of (-)14.2% was for response code 'others', (-)11.8% for informants coded as 'busy' and (-)8.6% for 'reluctant informants'. It was lowest at (-)7.0% each for recoded as 'cooperative & capable' and 'cooperative but not capable'.

Figure 2
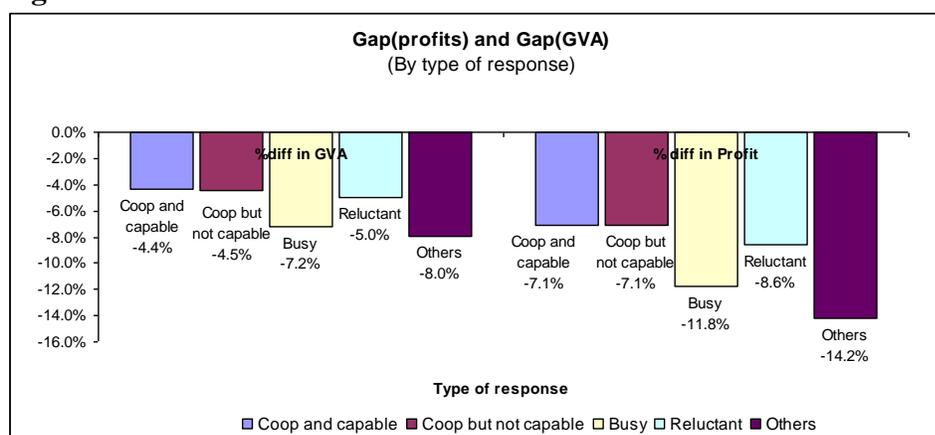
6.3 Informant

Normally in respect of a proprietary/partnership enterprise with no written records of transactions, it is the owner/partner of the enterprise who is expected to have the best information about the transactions of the enterprise. For 91.3% of the enterprises, the respondents were owner/partners in the observed sample, 2.8% were Managers of enterprise and 5.9% were others.

Mean GVA(P) and profits(derived) were higher as compared with mean GVA(I) and profits(direct). The observed Gap(profits) was largest for Managers at (-)8.9% as compared to (-) 7.0% for Owner/Partner. Thus on an average if the respondent was the owner of the enterprise, the observed differences in the profits and GVA from the two approaches were lower. (see Fig.3)

Figure 3

Gap(profits) and Gap(GVA)
(By informant)
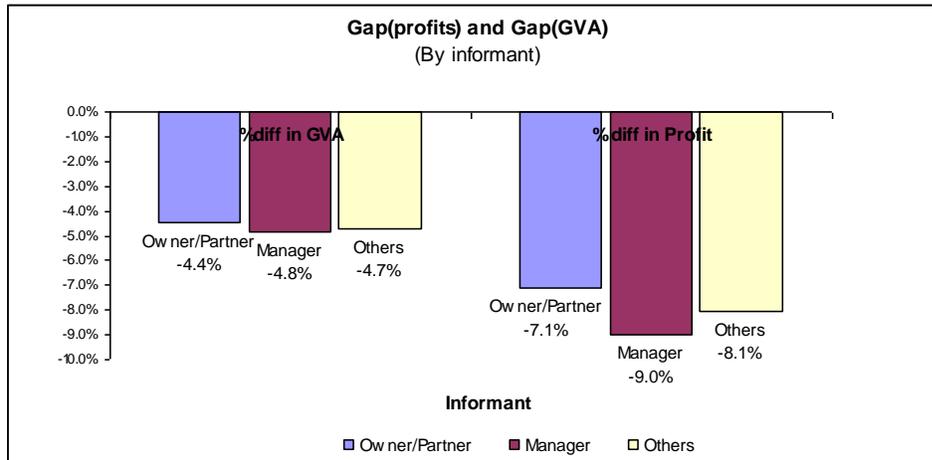
6.4 Enterprise Type

The NSS classifies enterprises into three types (i) Own Accounts Manufacturing Enterprises (OAE) – are the enterprises run by household labour i.e. with no hired labour (ii) Non-directory Manufacturing Enterprises (NDME) have less than six workers with atleast one hired worker and (iii) Directory Manufacturing Enterprises (DME) are the enterprises with six or more workers and atleast one hired worker. In our sample 67.1% enterprises are OAMEs, 21.8% are NDMEs and 11.1% are fairly large enterprises i.e. the DMEs. Figure 4 shows that the mean profits(direct) is lower than mean profits(derived) across all the three enterprise types. However, the lowest Gap(profits) from the two approaches are observed in case of OAMEs at (-)3.9%, increasing to (-)7.3% for NDMEs and a high of (-)10.3% for DMEs. Similar trend is for the Gap(GVA) but to a lesser extent.

Figure 4

Gap(profits) and Gap(GVA)
(By type of enterprise)
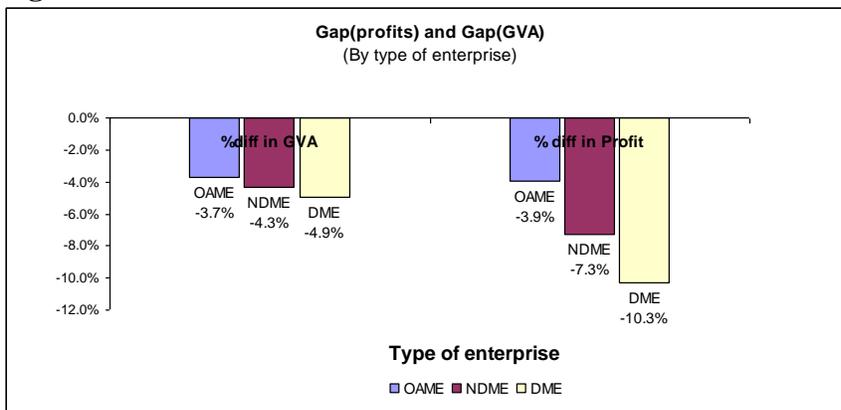
6.5 Size of Employment

The average employment of an enterprise in the observed sample is 2.8, with 2.4 for rural enterprises and 3.0 for urban enterprises. Nearly 65.0% enterprises were small with less than 2 workers and 84.4% enterprises had less than 5 workers. For rural enterprises these were 74.4% and 90.0% respectively. Irrespective of the size of employment mean profits(derived) and mean GVA(P) were higher than mean profits(direct) and mean GVA(I) respectively. It is observed that Gap(profits) increased substantially with increase in the size of employment. Thus while the observed Gap(profits) was (-)3.1% for enterprises with only 1 worker, the Gap(profits) increased to (-)8.4% for enterprises with 4 workers and was highest at (-)11.2% for enterprises with 10-19 workers. However, the Gap(GVA) did not increase to that levels and was around (-)5.0% for employment size of 4 and above. (see Figure 5)

**Figure 5**
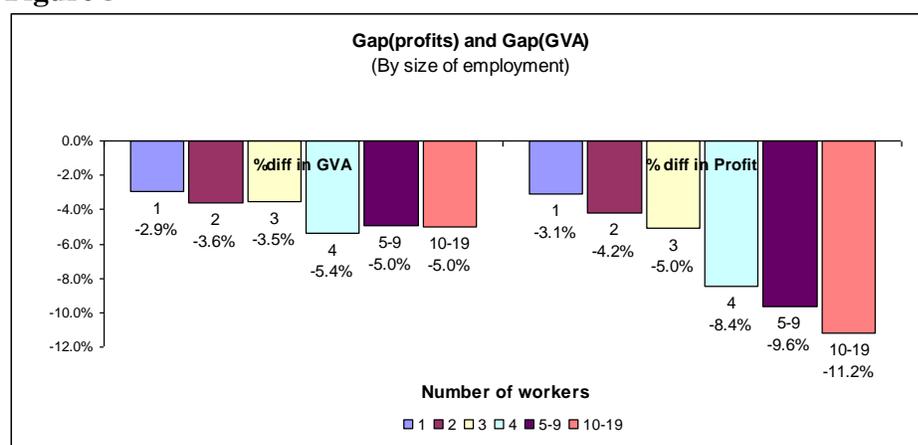
6.6 Size of Plant and Machinery

Given that the survey captures the unregistered manufacturing sector, 74.2% of the enterprises in the sample had plant & machinery (P&M) valued at less than INR 10000 (or roughly US$200). For rural enterprises this was 81.0%. Data also suggests that across rural and urban sectors and across all classes of size of P&M, mean GVA(P) and mean profits(derived) are higher than average GVA(I) and profits(direct). Figure 6 shows the mean GVA(P), GVA(I), profits(derived) and profits(direct). The Gap(profits) increased from (-)4.9% for the lowest category of P&M value of less than INR1000, to high of (-)10.5% for enterprises with P&M above INR 50000. Although, the observed percentage difference dropped for the 2nd size class of INR 1000-5000 it generally increased with the size of P&M. The Gap(GVA) also increased with the increase in the P&M size except for the 2nd size class of INR 1000-5000.

Gap(profits) and Gap(GVA)
(By size of plant and machinery)
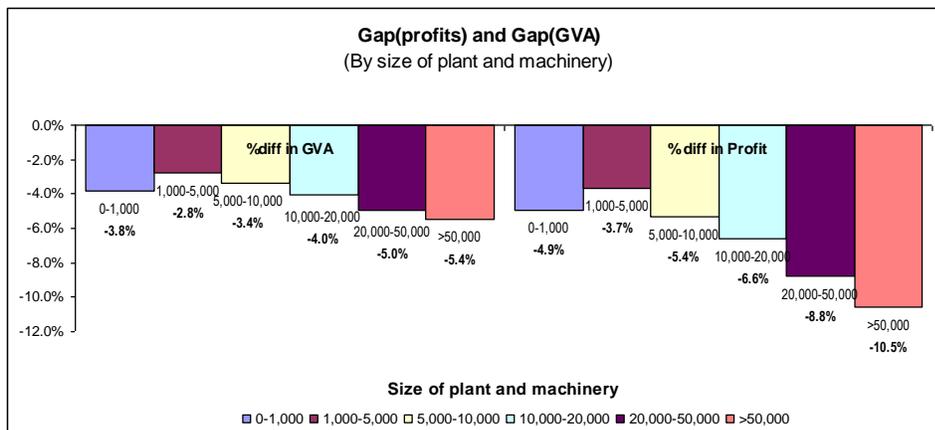
6.7 Registration

The survey covered enterprises which are not registered as factories under the Indian Factories Act. However, information related to registration with local authorities was collected from each enterprise. The registration of an enterprise even with a local authority is an indicator of larger operations as compared with an unregistered enterprise. Only about 21.2% enterprises were registered with one or more of the local agencies. From Figure 7 it can be see that mean GVA(P) and mean profits(derived) are higher as compared with mean GVA(I) and mean profits(direct). The Gap(profits) was however (-)10.0% for enterprises with some local registration as compared to (-)4.7% for enterprises with no local registration.

Figure 7

**Gap(profits) and Gap(GVA)**
(By whether registered or not)
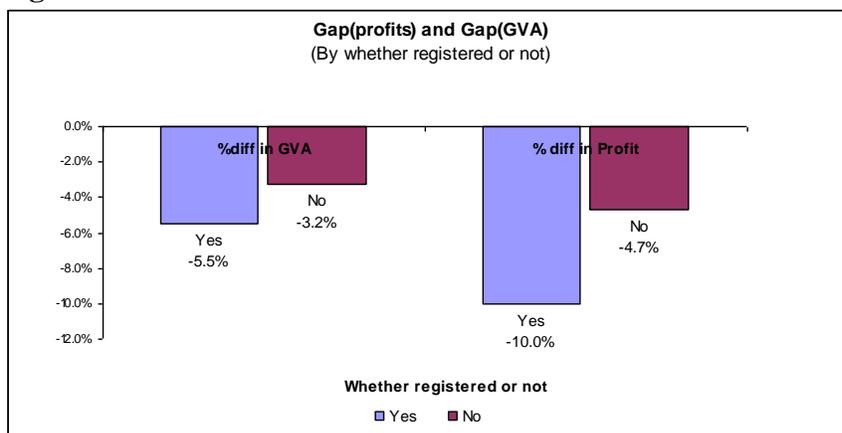
6.8 Location of Enterprise

Most of the enterprises in the unorganized sector because of their small nature of operation, even in the manufacturing sector, operate from the household premises. 53.7% of the enterprises in the observed sample operated from within the household premises, with this percentage at 64.1% for rural sector. Data on mean GVA(P), GVA(I), profits(derived) and profits(direct) for the enterprises shows that the GVA and profits earned are higher for enterprises located outside the premises of the household. However, consistently across rural and urban, on an average GVA(P) and profits(derived) are higher than GVA(I) and profits(direct). The Gap(profits) is much higher

at (-)8.1% for enterprises located outside household premises as compared to (-)5.3% for enterprises located inside the household premises. (see Figure 8)

Figure 8

Gap(profits) and Gap(GVA)
(By location of enterprise)
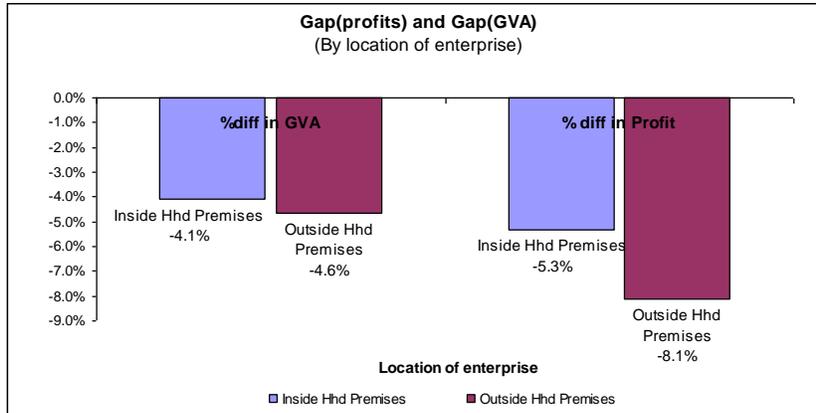
6.9 States

We also examined the sample of enterprises across - states/UTs of India for mean GVA(P), GVA(I), profits(derived) and profits(direct) to examine variations due to geographical location of enterprises within the country. Results can be seen in Figure 9. It is obvious that except for the small UT of Pondicherry, in all other states/UTs across rural and urban sectors, mean GVA(P) and mean profits(derived) were higher than GVA(I) and profits(direct) respectively. Among the major states, the three southern states of Kerala, Tamil Nadu and Karnataka had lowest Gap(profits) of (-)1.4%, (-)1.7% and (-)2.3% respectively with the 4th southern state of Andhra Pradesh also reporting a low difference of (-)3.5%. Among bigger states Gap(profits) above (-)8.0% are observed in Maharashtra (-)9.1%, Haryana (-)12.5%, J&K, (-)14.9%, Assam (-)15.3%, Punjab (-)16.0% and Uttaranchal (-)17.8%. Some smaller states and UTs fell in the category Gap(profits) exceeding (-)10.0%.

Figure 9

Gap(profits) and GAP(GVA)
(By state)
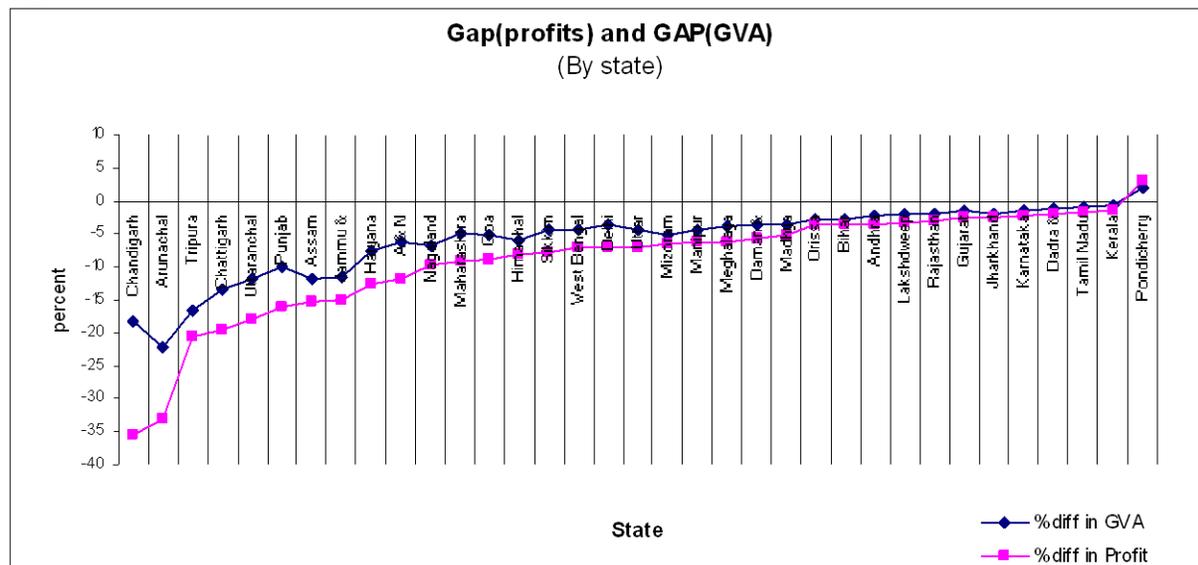
7. Profits(derived) and Profits(direct): Profile of enterprises

It is observed from the data that that the mean profits(derived) are higher than mean profits(direct) for the sample of enterprises. This however, does not imply that all enterprises had profits(derived) greater than profits(direct). We looked into the characteristics of the enterprises based on whether the profits(derived) were less than, greater than or equal to the profits(direct). More than 60.0% of the enterprises had profits(derived) higher than profits(direct), 28.0% reported profits(direct) higher than profits(derived) and for nearly 12.0% the two measurers of profits were surprisingly equal. The basic characteristics of the three types of enterprises are given in Table 4. The enterprises for which profits(derived) is less than profits(direct), have more or less similar general characteristics in terms ownership, local registration, maintenance of accounts, and location of enterprise within household premises. However, the two types of enterprises differed in terms of mean value of employment, total receipts, total expenses, profits(derived), profits(direct) and GVA(P) and GVA(I). The enterprises with profits(derived) higher than profits(direct) had much higher mean employment, P&M size, total receipts, mean total expenses and mean profits(derived) and profits(direct), pointing that these were relatively larger enterprises in terms of business operations. The enterprises where the two measures of

Table 4: Characteristics of enterprises with profits(direct) less than, greater than or equal to profits(derived)

	Profits (direct)< Profit (derived)	Profit (direct) > Profit (derived)	Profit (direct) = Profit (direct)	Total
Number of enterprises	91,369	42,328	17,078	150,775
Mean total workers	2.89	2.75	2.27	2.78
Mean hired workers	1.19	1.09	0.70	1.11
No. of Proprietary enterprises	96.1%	96.4%	97.2%	96.3%
Male owned	78.9%	76.8%	64.6%	76.7%
Female owned	17.2%	19.6%	32.6%	19.6%
Own account enterprises	65.4%	65.8%	79.6%	67.1%
Enterprise maintaining business accounts	3.9%	3.8%	4.0%	3.9%
Enterprises located within household premises	52.3%	52.3%	64.1%	53.6%
Enterprises registered with any local authority	22.9%	21.0%	12.8%	21.2%
Mean total receipts (output) (in INR)	269,707	212,242	158,258	240,951
Mean total expenses (in INR)	218,478	179,235	125,986	196,984
Mean intermediate consumption(in INR)	188,736	152,294	108,840	169,456
Mean compensation of employees (in INR)	25,027	21,908	14,192	22,924
Mean Profits(derived) (in INR)	51,229	33,007	32,273	43,966
Mean Profits(direct) (in INR)	43,806	37,569	32,273	40,749
Mean GVA(P) (in INR)	80,971	59,948	49,418	71,495
Mean GVA(I) (in INR)	73,547	64,510	49,418	68,277

profits were equal were much smaller in size and production compared to the other two categories. The enterprises that reported profits(direct) lower than profits(derived), had profits(direct) lower by around (-)14.5%. On the other hand enterprises which reported profits(direct) higher than mean profits(derived), had profits(direct) higher by around (+)13.8%. There is again some evidence that enterprises with larger incomes tend to under report profits when confronted with a direct question to reveal their profits.

An interesting observation from the data (Tables 2A and 2B) is that while **0.7%** of enterprises in the sample had profits(derived) less than zero, none of these enterprises reported profits(direct) as negative. In other words none of the sampled enterprises reported a loss when asked a direct question. de Mel et.al. (2009) report similar results in their Sri Lankan experiments with microenterprises. In our sample there are around **0.7%** enterprises that report zero profits(direct). Thus while enterprises reported zero profits (or no profit and no loss) when asked the direct question on profits, none of them reported a loss (or negative profits), whereas some enterprises had profits(derived) negative when profits were derived using detailed data on their receipts and expenses.

8. What do we conclude from above?
We started by reviewing whether the Approach 1 with a long questionnaire and detailed items to capture income and expenses of an unorganized (informal) enterprise captures profits and GVA data more accurately as compared with the Approach 2 with a single shot question on profits. The results from the Indian experience of NSS 56th round shows that Approach 1 yields on an average a measure of profits and GVA which is higher than the measure of profits (and GVA) from Approach 2. From the results reviewed above we conclude the following:

1. On an average the profits(direct) were lower than the profits(derived). In other words Approach 1 of interviewing the enterprise with detailed sets of questions on receipts and expenses of the enterprise during the reference period gave higher profits as compared with the Approach 2 of asking profits of the same enterprise through a single shot question. As GVA(P) and GVA(I) depend upon Approach 1 and 2 respectively, on an average GVA(P) was higher than GVA(I).
2. This was true across various enterprise characteristics – such as rural or urban, response code, informant, ownership, enterprise type, maintenance of accounts, registration, location of enterprise, employment, size of P&M, industry, and state in which the enterprises are located.
3. The correlations between the profits(derived) and profits(direct) as also between GVA(P) and GVA(I) were very high and positive.
4. The Gap(profits) and Gap(GVA) were lower if the respondents were cooperative as compared with other respondents who were busy or reluctant. Similarly, the Gap(profits) and Gap(GVA) were lower if the respondents were the owners themselves. Compared with other major states, the Gap(profits) and Gap(GVA) were much lower for the four southern states of Kerala, Karnataka, Tamil Nadu and Andhra Pradesh as compared with the other states.
5. It is observed that as the size of enterprise increases in terms of employment, or size of P&M, the Gap(profits) and Gap(GVA) also increase implying that with the increase in size the directly reported profits are much lower than the profits derived from incomes

and expenses of the enterprise. This is also true for all characteristics which indicate a higher level of operation of an enterprise such as registration, location, accounts maintained etc. Putting this in other words, the measure of profits and GVA using Approach 2 are much more likely to be closer to measure of profits and GVA obtained from Approach 1 for smaller enterprises as compared to larger enterprises.

6. This also suggests that if the target enterprises for a researcher are very small enterprises in terms of employment and size of plant and machinery (for manufacturing enterprises), a short questionnaire with a few direct questions would yield results closer to what would be obtained by using a questionnaire with detailed disaggregated data items. This is particularly useful when there are resource constraints and the researcher decides that some compromise could be made on the non-sampling errors in the interest of saving time and resources. Well trained field interviewers are the key in controlling the recall errors in any survey enquiry.

7. One interpretation of this could be that as smaller enterprises have much simpler operations and less number of transactions as compared to a larger enterprise, the recall lapses are much less and therefore the two approaches give much closer results for smaller enterprises. Another explanation in the Indian context could be that a large number of very small enterprises do not fall within the income-tax threshold. For such enterprises there is little incentive in suppressing incomes or profits. However, enterprises which are close or above the threshold income limits and are not paying taxes will have the perverse incentive to under report incomes and profits. The NSS data however, does not captures the information whether the enterprises paid any taxes on incomes to study the difference in behaviour of tax payers and non tax payers. It is also sensitive about including such a question in such enquiries as this would make the enterprise further suspicious. Reasons for underreporting of profits and revenues could also be due to expectation of benefit under some government scheme.

8. Although Approach 1 yielded higher profits and GVA as compared with Approach 2, still there are possibilities that the enterprises underreported revenues and/or overstated expenses and even though the profits(derived) on an average were higher compared to profits(direct), they on an average might still be underreported due to underreported revenues and overstated expenses. However, there is no way to test this from the NSS 56th round data.

9. Suggestions for further methodological work

Considering that in Indian context the enterprise surveys of NSSO are used to provide benchmark estimates of GVA per worker for estimating the contribution to GDP and a single shot question provides lower estimates of GVA, there are severe implications in resorting to a single shot questionnaire approach. Although as observed above, the short questionnaire approach could be useful when the researcher is interested in collecting data on very small informal enterprises with low levels of investments and employment. However, it would be interesting for NSSO to undertake further pilot studies to test several approaches to get an indication of overestimation or underestimation of profits and GVA even within the existing approach. Additional questions could be included to get indirect estimates of misreporting of incomes, expenditures and profits in the current approach of data collection. This was attempted in the 56th round by asking the perception of the interviewer whether the enterprise underreported its direct(profits) and further asking the range of the profits of the enterprise as assessed by the interviewer. Unfortunately, the data on these questions has not been provided to

researchers. Although it may be possible for the interviewer to judge whether the enterprise misreported its profits, it is very difficult for the interviewer to get a perception of the profits of an enterprise in a 60 minutes interview, which was attempted in the 56th round. Approaches used in de Mel et.al. (2008) experiments in Sri Lanka, which aim at getting such information through indirect questions, could provide more meaningful understanding of the extent to which enterprises over report expenses or under report revenues and profits. de Mel et.al (2008) also find fears of income tax in their experiments with firms in Sri Lanka as reasons for misreporting, which might be quite true for India. Further methodological research could to be undertaken to test (i) a short questionnaire vs a long questionnaire (ii) testing a long questionnaire to collect data for the reference period in a single visit to the enterprise as against multiple visits to the enterprise (iii) introduction of diaries to the enterprises to record daily transactions. Multiple visits and diary method although will increase the costs of enquiry, are expected to reduce recall errors and yield more accurate estimates.

References:

- Central Statistical Organisation (2007): National Accounts Statistics - Sources and Methods, Ministry of Statistics and Programme Implementation, Government of India.
- Charmes, J. The contribution of Informal Sector to GDP in Developing Countries: Assessment, Estimates, Methods, Orientation for the Future, presented at 4th Meeting of the Expert Group on Informal Sector Statistics (Delhi Group) Geneva 28-30 August 2000
- Deaton, Angus and Grosh, Margaret (2000). "Consumption" pp 91-133 in M. Grosh and P. Glewwe (eds.), *Designing household questionnaires for developing countries: lessons from fifteen years of the living standard measurement study, Vol. 1*. Washington, DC, World Bank.
- de Mel, S., McKenzie, and Woodruff , Christopher, Measuring Microenterprise Profits: Must we ask how the sausage is made?, Journal of Development Economics 88 (2009) 19-31
- Kolli, Ramesh, The Informal Sector in the National Accounts of India, presented at International conference organized by the International Association for Research in Income and Wealth (IARIW) and the National Bureau of Statistics (NBS) of China,.Beijing, China September 18-21, 2007,
- Kulshreshtha, A.C (1998): Informal Sector in India: Conceptual and Estimational Issues in the Context of UN System of National Accounts. *Indian Journal of Labour Economics*, Vol. 41, No.3.
- Kulshreshtha, A.C and Singh, Gulab (1998): Contribution of Unorganised Sector in the Indian Economy. *Manpower Journal*, Vol. XXXIV No.3, pp. 45-67
- Liedholm, Carl (1991) "Data Collection Strategies for Small-scale Industry Surveys", GEMINI Working Paper No. 11
- Liedholm, Carl (1987) "Small Scale Industries in Developing Countries: Empirical Evidence and Policy Implications", Paper No. 9, MSU International Development Papers
- OECD, *Measuring the Non-Observed Economy, A Handbook,* 2002