

INFORMAL SECTOR ENTERPRISES: SOME MEASUREMENT ISSUES

BY KAUSHAL JOSHI,* GLENITA AMORANTO AND RANA HASAN

Asian Development Bank

In this paper, we 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, versus profits obtained through a single direct question; and (ii) gross value added (GVA) obtained using the production approach as the difference between output and intermediate consumption from a detailed set of questions on incomes and expenses, versus GVA using the income approach by asking a few questions on factor incomes, and a single direct question on profits. We use data from the 56th round survey of unorganized manufacturing conducted by the National Sample Survey Organization of India 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.

JEL Codes: E01, E23, E26

Keywords: informal sector, unorganized sector, gross value added, enterprise profits, national sample survey organization

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 15th International Conference of Labor Statisticians (ICLS) held in 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. A 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,” and micro-enterprises.

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 thirds of total employment and more

Note: 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. This is a revised version of the paper presented at the Special IARIW–SAIM Conference on Measuring the Informal Economy in Developing Countries, held September 23–26, 2009 in Kathmandu, Nepal. The authors are grateful to the two anonymous referees and Mr. Derek Blades for their useful suggestions to improve the paper.

*Correspondence to: Kaushal Joshi, Senior Statistician, ERD/ERDI, Asian Development Bank, 6 ADB Avenue, Mandaluyong City 1550, Metro Manila, Philippines (kjoshi@adb.org).

© 2011 The Authors

Review of Income and Wealth © 2011 International Association for Research in Income and Wealth
Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St,
Malden, MA, 02148, USA.

than one third of the total gross domestic product (GDP) of the non-agricultural sector in Asia. In view of its estimated size, the sector continues to invite high policy interest in many parts of the world on account of its tremendous potential for job creation and income generation in developing economies. Therefore, the need for data on various characteristics and operations of these enterprises—output generated, employment provided, and constraints faced and their relationship with the formal sector—cannot be overemphasized.

National accountants recognized as early as the 1960s the importance of measuring its economic significance to the extent that estimates on what was then called the “traditional sector” had to be generated from whatever limited data were available (Charmes, 2000). In response, the international statistical community has carried out several pilot surveys of the informal sector in various countries which became the basis later for developing statistical methodologies to redefine national accounts and to permit special surveys. In 1993, the resolution concerning statistics of employment in the informal sector was adopted by the 15th ICLS which provided the first internationally approved technical guidelines for the development of statistics on the informal sector. Since then, various countries have attached to the occupational household survey questionnaires a module on informal sector through capacity-building projects of the International Labour Organization (ILO). As of 2000, more than 60 countries have already started to collect and publish data on employment and other characteristics of the informal sector (ILO, 2000). There has been extensive debate on having a uniform definition for the informal sector. However, consensus has been difficult to arrive at and a variety of criteria have been used by the countries. Considerable work has been undertaken on harmonizing these definitions in the Delhi Group on Informal Sector Statistics, and discussions on this can be traced to the various documents, papers, and ongoing work on the draft *Manual on Surveys of Employment and Informal Sector* under the Delhi Group at the website www.mospi.gov.in/mospi_informal_sector.htm.

Unfortunately, due to the very nature of their operation, i.e. small-sized, and characterized by high rates of entry and exit, informal sector enterprises do not enter the business registers or the list frames usually maintained by national statistical offices (NSOs). As a result, informal sector enterprises escape official data collection systems, and their contribution to GDP is generally poorly estimated, and often understated, in the official national accounts data despite their significant contribution to actual economic activity.

Besides achieving a more accurate measurement of national accounts, sound statistics on the informal sector are essential tools for evidence-based policy-making and better targeted social and economic policies toward improvement of the working conditions and legal and social protection of the persons employed in the informal sector; enhancing the productivity of informal sector activities; training and skills development; organization of informal sector producers and workers; development of appropriate regulatory frameworks; and urban development. Since many women and children are employed in the informal sector, it is closely linked with issues relating to the contribution of women to economic activities and concerning child labor (Suwal and Pant, 2009). Further, as most people who participate in the informal sector come from the poor, whose working

conditions more often are below decent work standards, accurate information on the dimensions of the informal sector and its contributions to employment and to GDP is essential for more robust poverty analysis (ECLAC, 2007).

Work on the informal sector undertaken by the ILO has helped in better understanding of the informal sector and in removing the usual misconception that the informal sector is totally unorganized or unstructured. Rather the urban informal sector is found to have its own mechanisms and networks of financial services, training, marketing, welfare schemes, and social safety nets. While most enterprises in this sector are run by individuals as own account enterprises, the sector also includes micro and small-scale enterprises, usually in the manufacturing sector, that hire labor on a fairly regular basis and have links with the formal sector enterprises through subcontracting (ILO, 2000). This has also been observed in the surveys of unorganized manufacturing enterprises conducted by the National Sample Survey Organization (NSSO) in India. These further reinforce the need for accurate and regular statistics on the structure, size, and contribution of this sector to employment and GDP.

In this paper, we will be concerned with some measurement issues that are faced in the process of collecting data from these enterprises through sample surveys. We will be using the terms “informal” and “unorganized” interchangeably. 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 from relate to the National Sample Survey (NSS) 56th round survey of the unorganized manufacturing sector in India.

2. MEASUREMENT OF THE 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 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 the household and enterprise. Further, absence of written records of transactions leads to relying on recall by the informant. This is problematic since the irregular nature of transactions of informal enterprises makes proprietors susceptible to recall error (Liedholm, 1987). 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 information may be used 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 inquiry is useful if the main interest is measuring employment in the 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, organization of statistical systems, and resources available (for more details, see OECD, 2002). Such a survey will seek to collect data from the enterprise owners as respondents. Apart from the objectives of the inquiry, the data collection strategy and inclusion of data items in the survey instruments depend on many other factors such as relative size of sampling and non-sampling errors that would creep into any survey design and in survey operations, and availability of financial resources. NSOs would like to undertake nationwide sample surveys but are very often constrained by resources, and consequently, the design is guided by availability of resources, both financial and skilled manpower.

Sampling errors would depend upon the sampling design and sample size. 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. In designing a survey questionnaire, the options could be: (i) using a detailed set of questions to collect data at the 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 (see discussions in Deaton and Grosh, 2000), such experiments for surveys of enterprises are rather limited.

Further, data for the reference period could be collected at a single point in time or at multiple points in time through interviews by repeated visits to the enterprise. Each of these methods has implications on costs and on the sampling and non-sampling errors for a given sample size. As Liedholm (1991) notes, “If resources for investigation are fixed, increasing the frequency of interviews will necessitate reducing the sample size and consequently, will 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.” The longer the recall period, the more likely the inquiry will be affected by recall errors. Additionally, Casley and Lury (1981) as quoted in Liedholm (1991, p. 3) argue that, in the context of developing countries, non-sampling errors are more significant than sampling errors. Re-interview studies have shown the presence of “alarmingly high levels of response errors even on the simplest of survey questions” (Scott, 1985, p. 15; as quoted in Liedholm, 1991, p. 3), and in some Indian surveys, errors may have been six times the sampling errors (Casley and Lury, 1981, p. 87; as quoted in Liedholm, 1991, p. 3).

Motivated by the fact that the accurate measurement of profits from micro enterprises is crucial for understanding the success of a variety of policy and programmatic interventions, including micro finance, a recent study tried different methods of data collection to better understand the problems that plague the accurate measurement of profits from micro enterprises. In particular, de Mel *et al.* (2009) conducted two panel surveys of Sri Lankan micro enterprises between 2005 and 2007. One of the surveys involved 618 micro enterprises with invested capital of around \$1000 or less, excluding investments in land and buildings, and were

engaged in retail trade and manufacturing operations. The other survey focused on 180 retail trade firms.

de Mel *et al.* (2009) carry out a number of experiments involving data on profits, revenues, and expenditures. To begin with, the authors find considerable discrepancy between a direct measure of profit (obtained by asking owners directly for their profits) and a measure obtained as the difference between reported revenues and reported expenses. For example, in the sample of manufacturing micro enterprises, the Pearson correlation between reported profits and reported revenues minus expenses is as low as 0.20 (the Spearman correlation coefficient is higher, around 0.42) once data outliers are dropped. Moreover, while no entrepreneur reports negative profits, over a quarter of micro enterprises have negative values for reported revenues minus expenses. Clearly, determining which of the two measures is more accurate is important.

The authors consider a variety of reasons for the discrepancy between reported profits and reported revenue minus expenses. These include unreported categories of expenses or forms of profit (for example, business goods and materials used for home consumption but recorded as business expenses); a mismatch in the timing of input purchases and sales; recall errors; and deliberate misreporting. The authors find that a large part of the differences between profits versus revenues minus expenses can be accounted for by the reporting of goods used for home consumption under business expenses, and the mismatch between revenues and expenses. Recall errors also occur so that monthly sales tend to be understated when recalled after four months versus one month. The use of diaries seems to reduce the recall error for both revenues and expenses. The analysis of the authors also suggests that firms underreport revenues by around 30 percent and that the average micro enterprise also underreports profits by around 20 percent.

A key conclusion drawn by de Mel *et al.* (2009) is that asking owners of micro enterprises directly for their profits yields a measure which appears “at least as reasonable as asking for all the ingredients in terms of detailed revenue and expenses.”

In this paper we will examine the differences in the measures of: (i) profits of enterprises obtained from a detailed set of questions on incomes and expenses of the enterprise versus profits of the enterprise obtained through a single direct question; and (ii) gross value added (GVA) as the difference between output and intermediate consumption of the enterprises obtained from a detailed set of questions on incomes and expenses, versus GVA using the income approach obtained as the sum of factor incomes of which profits (captured by a single direct question) is an important component. To examine this, we use data from the NSS 56th round survey of unorganized manufacturing conducted by the NSSO 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

The 56th round of the NSS conducted during July 2000–June 2001 is dedicated to collection of data on economic and operational characteristics of small

manufacturing enterprises in the unorganized sector of the Indian economy. The registered factory sector in India representing the organized manufacturing sector is covered annually through the Annual Survey of Industries (ASI). The ASI covers all enterprises in the manufacturing sector registered under the Factories Act 1948. These are enterprises employing 10 or more workers and using power, and those employing 20 or more workers without using power. The units not covered under the ASI are treated as part of the unorganized sector with respect to manufacturing activities. This concept of organized and unorganized enterprises in the manufacturing sector is consistent with the organized–unorganized dichotomy used in the Indian system of national accounts. Surveys of unorganized manufacturing enterprises are conducted once in five 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 GVA estimates are initially prepared at the 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 obtained through the quinquennial 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 at each detailed industry level. For the current price estimates of GVA for unregistered manufacturing, the relevant wholesale price indices are superimposed on the constant price GVA estimates at the detailed industry level (Kulshreshtha and Singh, 1998; Kolli, 2007; Kulshreshtha, 2009). The share of the unregistered manufacturing sector in the net domestic product (NDP) of the manufacturing sector was 37.5 percent in 1999–2000, which came down to 32.6 percent in 2007–08 (CSO, 2009). Given that the survey is used to provide benchmark data for the contribution of the unorganized manufacturing sector to the Indian economy, accuracy of data collected becomes very important.

4. DATA DESCRIPTION

The NSS 56th round adopted a stratified two-stage sample design with villages/urban frame survey (UFS) blocks as the first stage units and manufacturing enterprises in the unorganized sector as the ultimate stage units (USUs). Data in the survey were collected for a reference period of one month, from a sample of more than 150,000 enterprises in rural and urban India based on oral inquiry. From the unit-level data, we dropped the enterprises with 20 or more workers to control for results that may be affected by outliers, as some of the enterprises in the original sample were quite large—some had more than 100 workers. 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 (to be described later) was more than or equal to 1000. This is equivalent to the topmost 0.5 percent of the distribution. Further, more than 600 enterprises with zero entries in all sub-items of receipts, expenditures, and factor incomes were also dropped. The above data cleaning and

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 number of hired workers	0.72	1.37	1.11
By ownership of enterprise			
Proprietary enterprises ^a (%)	97.7	95.4	96.3
Male-owned (%)	79.0	75.2	76.7
Female-owned (%)	18.7	20.3	19.6
Non-proprietary enterprises ^b (%)	2.3	4.6	3.7
Own account enterprises ^c	78.9	59.4	67.1
Enterprise maintaining business accounts	2.0	5.1	3.9
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)	137,906	308,679	240,951
Mean total expenses	107,904	255,535	196,984
Mean intermediate consumption	94,054	219,015	169,456
Mean compensation of employees	11,392	30,504	22,924
Mean profits (derived)	30,003	53,144	43,966
Mean profits	28,049	49,096	40,749
Mean GVA(P)	43,852	89,664	71,495
Mean GVA(I)	41,898	85,615	68,277

^aProprietary enterprises are those where an individual is the sole owner of the enterprise.

^bIncludes the following: (a) *partnership within members of household*; (b) *partnership between members outside household*; (c) *co-operative society*; (d) *limited company*; and (e) *others*. *Partnership* is defined as the "relation between persons who have agreed to share the profits of a business carried on by all or any one of them acting for all." *Co-operative society* is a society that is formed through the co-operation of a number of persons (members of the society) to benefit the members. A *limited company* can be either private or public company. A *private company* means a company which by its Articles (a) RESTRICTS the right to transfer its shares, if any, (b) LIMITS the number of its members (not including its employees) to 50, and (c) PROHIBITS any invitation to public to subscribe for any shares or debentures of the company. *Public company* means a company which is not a private company.

^cAn own account enterprise is an undertaking run by household labor, usually without any hired worker employed on a "fairly regular basis." By "fairly regular basis" it is meant the major part of the period of operation(s) of the enterprise during the last 365 days.

filtering procedures yielded a final sample of 150,775 enterprises from the original sample of 152,494. Basic characteristics of the sample used in our analysis are given in Table 1.

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

5. NSS 56th ROUND QUESTIONNAIRE

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

Box 1. Data Items in the Survey Questionnaire to Calculate GVA from Production Side

Total Intermediate Consumption	Total Output
<ul style="list-style-type: none"> • Total raw materials consumed (data collected for five 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 and maintenance of fixed assets • Rent on plant and machinery • Service charges paid • Travelling, freight, and cartage expenses • Communication expenses • Consumables, packing materials, etc. • Paper and printing • License fees and local taxes (exclude indirect taxes) • Other expenses (consumer entertainment, performing rituals, etc.) • Distributive expenses 	<ul style="list-style-type: none"> • 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 plant and machinery and other fixed assets • Donations received (including government grants) • Other receipts (incentives received by enterprise, scrap sale receipts, etc.)

the GVA. The data collected on the above items enables calculation of GVA by using the following production approach:

$$(1) \quad \text{GVA (P)} = \text{Total Output} - \text{Total Intermediate Consumption}$$

where intermediate consumption represents the products used in the production of final products of the manufacturing unit and which 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 collects data on factor payments such as compensation paid to employees (salary, wages, and other benefits in cash and 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:

$$(2) \text{ Profits (derived)} = \text{Total Output} - \text{Total Intermediate Consumption (as above)} \\ - \text{Compensation of Employees} - \text{Rent Payable} \\ - \text{Interest Payable}$$

or:

$$(2a) \quad \text{Profits (derived)} = \text{Total Output} - \text{Total Expenses.}$$

The profits(derived) above is actually the balance left with the enterprise after paying for the inputs, other operating expenses, net taxes (on production), compensation of the hired workers (labor cost), the interest payable on outstanding loan, and rent payable on land and building for which data were collected in the NSS survey. This is different from the concept of the “operating surplus,” or “mixed income” in the case of unincorporated household enterprises, used in the UN System of National Accounts (SNA, 1993) which includes rent and interest. We have used this definition of “profits” because it corresponds to the item “net surplus” used in the NSS survey to obtain data on the profits of enterprise through a direct question so that valid comparisons can be done on the two measures.

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. The items included in the questionnaire and the instructions provided to the investigators are to ensure coverage of often unreported categories, such as business goods used for home consumption; business revenue used for household expenses but not included in revenues (or in profits); and firm inputs received as gifts but included as business expenses, which reflect the fungibility of resources between the household and business.

Up to 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 NSS 56th round survey however, 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 in the survey was defined as the amount that 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 the outstanding loan for the reference month. To get 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 family members who are treated as hired workers of the enterprise were excluded from net surplus and included under compensation paid to employees. The new item *net surplus* thus defined is equivalent to the profits (derived) in equations (2) and (2a) above. The data for *net surplus* was collected from the enterprise by posing a direct question after the disaggregated data on the different items of expenditures (including the compensation to employees, rent, and interest payments) and incomes had been collected orally from the owner (or respondent) for enterprises without books of accounts or from the records of the enterprise where books of accounts were available. Thus, while profits could also be derived using equations (2) or (2a) from the already collected detailed data items on receipts and expenses of an enterprise, the direct question on *net surplus* provided alternative estimate for the profits of the enterprise during the reference period.

Net surplus thus attempts to capture the balance left with the enterprise after meeting all kinds of expenditure and is denoted by “profits(direct)” in the paper.

Box 2. Approaches to Measuring GVA and Profit of an Enterprise

Approach 1	Approach 2
GVA(P) = Gross Value Added(production) = Total Output – Total Intermediate Consumption Profits(derived) = Total Receipts – Total Expenses	GVA(I) = Gross Value Added(income) = Net Surplus(Profit(direct)) + Interest + Rent + Compensation of Employees Profits(direct) = Net Surplus

This also provided an additional measure for estimating the GVA of the enterprise using the *income approach* by simple addition of factor incomes of the enterprise during the reference month as follows:

$$(3) \quad \text{GVA (I)} = \text{Net Surplus (or Profits(direct))} + \text{Interest} + \text{Rent} \\ + \text{Compensation to Employees.}$$

The NSS 56th round questionnaire thus provided a means for obtaining the GVA and profits of the same enterprise following two different approaches, which are summarized in Box 2.

In Approach 1, the GVA of an enterprise can be calculated following the production approach using detailed data collected on inputs and other operating expenses. On the other hand, under Approach 2, GDP of the enterprise can alternatively be estimated using the income approach by adding net surplus (profits(direct)), rent, interest, and compensation of employees—i.e. by adding incomes accruing to the factors of production. As far as profits are concerned, the same have been estimated by netting out all expenses of the enterprises including those on rent, interest, and compensation of employees under Approach 1, while Approach 2 captures profits (net surplus) by a direct question. The NSS 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 data from the NSS 56th round survey of 59,797 rural enterprises and 90,978 urban enterprises in the unorganized manufacturing sector, we examine some of results obtained from the two approaches and compare the results. We define:

$$\text{Gap(profits) for a set of enterprises} \\ = \frac{[\text{Mean Profits(direct)} - \text{Mean Profits(derived)}]}{\text{Mean Profits(derived)}} \times 100$$

$$\text{Gap(GVA) for a set of enterprises} = \frac{[\text{Mean GVA(I)} - \text{Mean GVA(P)}]}{\text{Mean GVA(P)}} \times 100.$$

Gap(profits) measures the deviation of mean profits(direct) from the mean profits(derived), expressed as a percentage of mean profits(derived). A negative

TABLE 2a
MEAN GVA(P), MEAN GVA(I), AND GAP(GVA)

		Section			Maintaining Accounts or Not	
		All	Urban	Rural	Yes	No
GVA (production approach)	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
	% negative	0.30	0.29	0.33	0.69	0.29
GVA (income approach)	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
Correlations	Pearson	0.9021***	0.9032***	0.8974***	0.8605***	0.9471***
	Spearman	0.9893***	0.9897***	0.9864***	0.9724***	0.9897***
Gap(GVA) (%)		-4.5	-4.5	-4.5	-7.5	-3.8
No. of observations		150,775	90,978	59,797	5,830	144,945

Note: ***p < 0.01.

TABLE 2b
MEAN PROFITS(DERIVED), MEAN PROFITS(DIRECT), AND GAP(PROFITS)

		Section			Maintaining Accounts or Not	
		All	Urban	Rural	Yes	No
Profit(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
	% negative	0.70	0.71	0.68	2.21	0.64
Profit (as reported)	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
Correlations	Pearson	0.8471***	0.8445***	0.8503***	0.8263***	0.8835***
	Spearman	0.9794***	0.9760***	0.9818***	0.9416***	0.9810***
Gap(profits) (%)		-7.3	-7.6	-6.5	-13.1	-6.1
No. of observations		150,775	90,978	59,797	5,830	144,945

Note: ***p < 0.01.

Gap(profits) indicates that the mean profits of the sample of enterprises based on Approach 2 (direct question on profits) is lower than the mean profits of the sample of enterprises based on Approach 1 (detailed disaggregated questions on incomes and expenses). Similarly, Gap(GVA) measures the deviation of mean GVA(Income) from the mean GVA(derived) expressed as a percentage of mean GVA(derived).

Table 2a reports the mean, standard deviation, and median of GVA(P) and GVA(I) calculated using the two approaches separately for rural and urban enterprises. It also gives the percentage of enterprises with negative GVA and correlations. Table 2b reports similar results from the two approaches for profits(derived) and profits(direct). It is important to note the large differences between the mean and median values due to the presence of enterprises with extremely large GVAs and profits. In fact, the aggregate GVA/profits of the top 10 percent of enterprises alone accounts for more than 50 percent of the total GVA/profits of all enterprises. The two tables show that the observed correlations between the two measures of

profits or the GVA are very high, although on average across rural and urban sectors, GVA(I) and profits(direct), which are based on a 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 percent lower than mean GVA(P), and mean profits(direct) are lower by 7.3 percent 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 rural enterprises. 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 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 inquiry, depending to a large extent upon the recall of the informant. Less than 4.0 percent 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 underreporting or overreporting 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, and median of GVA(P) and GVA(I); Table 2b reports the 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), respectively, for enterprises irrespective of whether accounts are maintained or not. However, Gap(profits) was much higher at –13.1 percent for enterprises with books of accounts (record based inquiry), while this gap was only –6.1 percent for enterprises with no books of accounts (oral inquiry). This implies that when the direct question on profits was asked, enterprises with books of accounts reported profits that were lower by 13.1 percent from the profits derived as a difference of incomes and expenses using the books of accounts. The mean Gap(GVA) was –7.5 percent for enterprises with accounts compared to –3.8 percent for enterprises without accounts (oral inquiry) (see Figure 1). Does this suggest that enterprises tend to underreport profits if asked directly, even if they maintained books of accounts from which the data was collected on detailed expenditures and receipts? The process of collecting data from enterprises with books of accounts for estimating the profits could be seen as a process of collecting the data from two different persons—the first being the disaggregated data on expenses and receipts obtained from the books of accounts (to calculate profits(derived)), and the second obtained from the enterprise owner through a direct question on net surplus (profits(direct)). This is somewhat different from the process of obtaining data from an enterprise where the disaggregated data on expenses and receipts has been

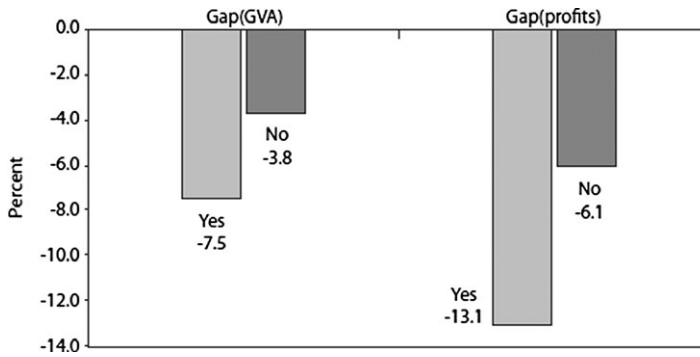


Figure 1. Gap(GVA) and Gap(profits) by Maintenance of Business Accounts

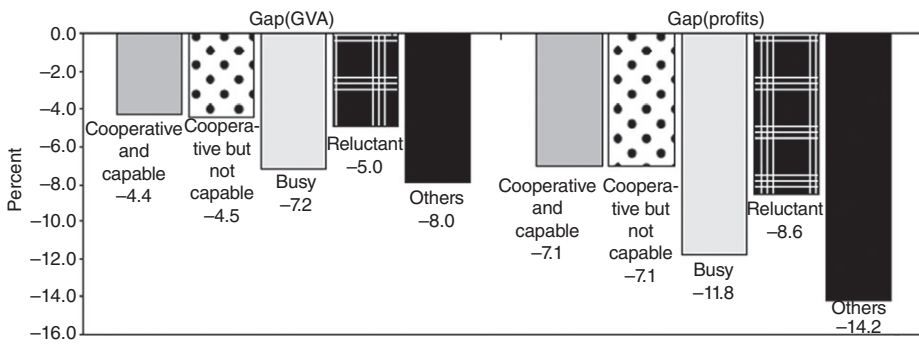


Figure 2. Gap(GVA) and Gap(profits) by Type of Response

collected orally from the owner (or respondent) before again posing a direct question on the *net surplus* (profits(direct)). Thus it could also be argued that smaller gaps derived for firms without books of accounts as compared to those with books of accounts could be a result of response contamination of the second measure of profits (*net surplus*) and therefore they are more likely to conform.

6.2. Response Code

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

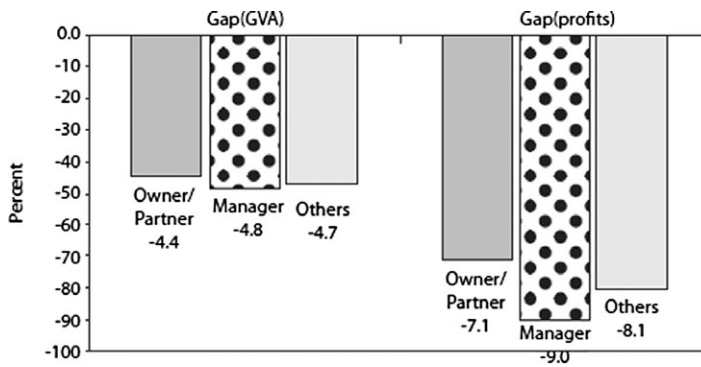


Figure 3. Gap(GVA) and Gap(profits) by Informant

percent for “busy,” and -8.6 percent for “reluctant informants.” It was lowest, at -7.0 percent each, for “cooperative and capable” and “cooperative but not capable.”

6.3. Informant

Normally, with respect to a proprietary/partnership enterprise with no written records of transactions, the owner/partner of the enterprise is expected to have the best information about the transactions of the enterprise. For 91.3 percent of the enterprises, the respondents were owner/partners in the observed sample; 2.8 percent were managers, and 5.9 percent were others. Mean GVA(P) and profits-(derived) were higher compared with mean GVA(I) and profits(direct), respectively (Figure 3). The observed Gap(profits) was largest for managers at -9.0 percent compared to -7.1 percent for owner/partner. Thus on average, if the respondent was the owner of the enterprise, the observed differences in the profits and GVA from the two approaches were lower.

6.4. Enterprise Type

The NSS classifies enterprises into three types: (i) own-account manufacturing enterprises (OAMEs) are enterprises run by household labor, i.e. with no hired labor; (ii) non-directory manufacturing enterprises (NDMEs) have less than six workers with at least one hired worker; and (iii) directory manufacturing enterprises (DMEs) are enterprises with six or more workers and at least one hired worker. In our sample, 67.1 percent enterprises are OAMEs, 21.8 percent are NDMEs, and 11.1 percent are fairly large enterprises, i.e. 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 the case of OAMEs at -3.9 percent, increasing to -7.3 percent for NDMEs, to a high of -10.3 percent for DMEs. A similar trend is observed for Gap(GVA) but to a lesser extent.

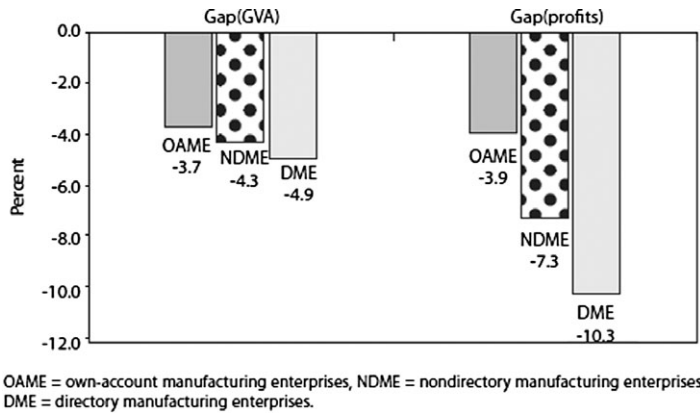


Figure 4. Gap(GVA) and Gap(profits) by Type of Enterprise

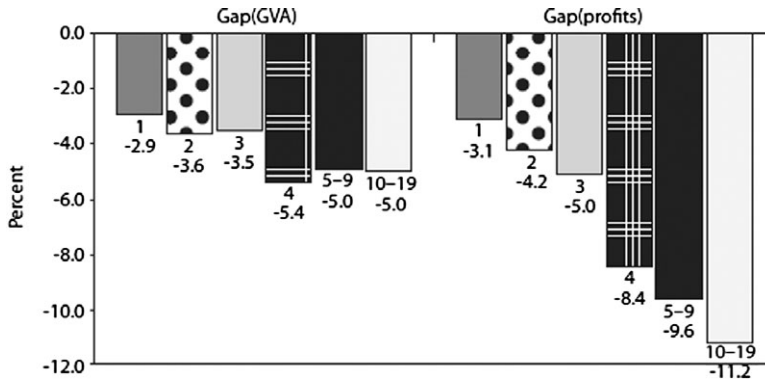


Figure 5. Gap(GVA) and Gap(profits) by Size of Employment

6.5. Size of Employment

The average employment of an enterprise in the observed sample is 2.78, with 2.41 for rural enterprises, and 3.02 for urban enterprises (Table 1). Nearly 65.0 percent of the enterprises were small with less than two workers, and 84.4 percent had less than five workers. For rural enterprises, these were 74.4 percent and 90.0 percent, respectively. Irrespective of size of employment, mean profits(derived) and mean GVA(P) were higher than mean profits(direct) and mean GVA(I), respectively (Figure 5). It is observed that Gap(profits) increased substantially with increase in the size of employment. Thus, while the observed Gap(profits) was -3.1 percent for enterprises with only one worker, the Gap(profits) increased to -8.4 percent for enterprises with four workers, and was highest at -11.2 percent for enterprises with 10-19 workers. However, the Gap(GVA) did not increase by as much and was around -5.0 percent for employment size of four and above.

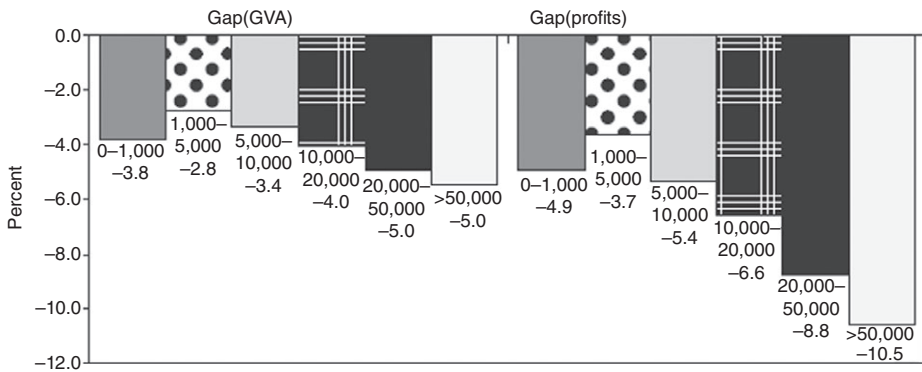


Figure 6. Gap(GVA) and Gap(profits) by Size of Plant and Machinery

6.6. Size of Plant and Machinery

Given that the survey captures the unregistered manufacturing sector, 74.2 percent of the enterprises in the sample had plant and machinery (P&M) valued at less than 10,000 rupees, roughly US\$200. For rural enterprises, this was 81.0 percent. Data also suggests that across rural and urban sectors (Tables 2a and 2b) and across all size classes of P&M (Figure 6), mean GVA(P) and mean profits-(derived) are higher than average GVA(I) and profits(direct), respectively. The Gap(profits) increased from -4.9 percent for the lowest category of P&M value of less than Rp1000, to a high of -10.5 percent for enterprises with P&M above Rp50,000. Although the observed percentage difference dropped for the second size class of Rp1000–5000, it generally increased with size of P&M. The Gap(GVA) also increased with an increase in the P&M size except for the second size class of Rp1000–5000.

6.7. Registration

The survey covered enterprises that are not registered as factories under the Indian Factories Act. However, information related to registration with municipal bodies and other government authorities (e.g. silk board, jute commissioner, etc.) was collected from each enterprise. The registration of an enterprise with other government authorities is an indicator of larger operations compared with an unregistered enterprise. Only about 21.2 percent of the enterprises were registered with one or more of the authorities. From Figure 7 it can be seen that mean GVA(P) and mean profits(derived) are higher compared with mean GVA(I) and mean profits(direct), respectively. The Gap(profits) was, however, -10.0 percent for enterprises with some registration compared to -4.7 percent for enterprises with no registration.

6.8. Location of Enterprise

Most of the enterprises in the unorganized sector, including those in the manufacturing sector, operate from the household premises because of the small

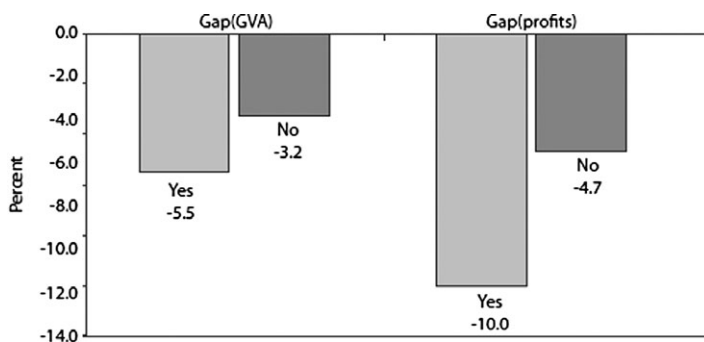


Figure 7. Gap(GVA) and Gap(profits) by Enterprise Registration

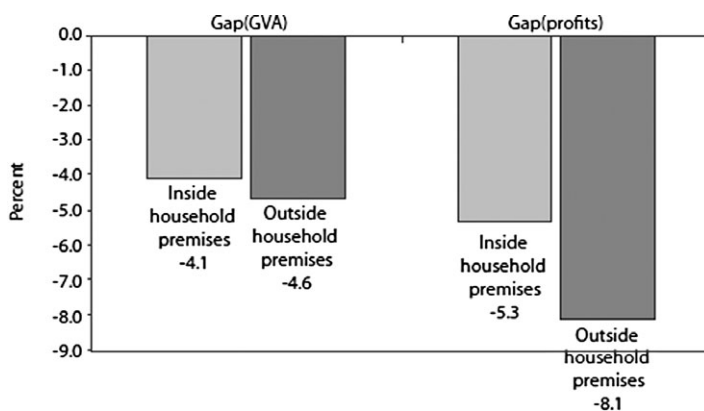


Figure 8. Gap(GVA) and Gap(profits) by Location of Enterprise

nature of their operation. Around 53.6 percent of the enterprises in the observed sample operated from within the household premises, of which 64.0 percent are from the 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 average GVA(P) and profits(derived) are higher than GVA(I) and profits(direct). The Gap(profits) is much higher at -8.1 percent for enterprises located outside household premises compared to -5.3 percent for enterprises located inside the household premises (see Figure 8).

6.9. States

We also examined the sample of enterprises across major states 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. The three southern states of Kerala, Tamil Nadu, and Karnataka had lowest Gap(profits)

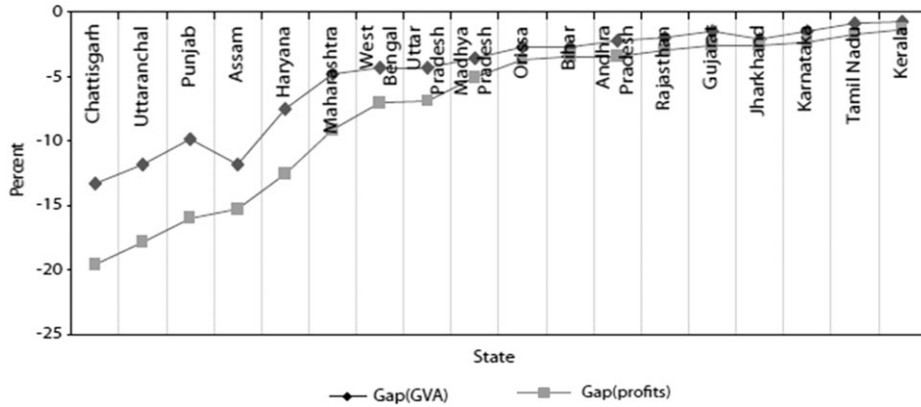


Figure 9. Gap(GVA) and Gap(profits) by Major States

of -1.4, -1.7, and -2.3 percent, respectively, with the fourth southern state of Andhra Pradesh also reporting a low difference of -3.5 percent. Gap(profits) above -8.0 percent are observed in Maharashtra (-9.1 percent), Haryana (-12.5 percent), Assam (-15.3 percent), Punjab (-16.0 percent), Uttaranchal (-17.8 percent), and Chattisgarh (-19.6 percent) (Figure 9). Other than the geographical feature just noted—i.e. the tendency for the southern states to have lower difference in Gap(profits) and Gap(GVA)—there does not seem to be any other systematic relationship between state characteristics and the distribution of the two Gap variables. Thus, for example, it is not the case that richer states show lower (or for that matter higher) differences. This may be seen by considering the cases of Gujarat and Punjab. These two states were among the richest in the 2000s (i.e. around the time of the enterprise survey). Yet, the differences in the Gap variables is large in the case of Punjab and small in the case of Gujarat. There is a similar lack of any relationship if we consider a socially oriented state characteristic such as literacy. Thus, while it is the case that Kerala had the highest literacy rate as per the 2001 Census (90.9 percent), and shows up with the lowest differences in the Gap variables, Jharkhand had one of the lowest literacy rates in India (53.6 percent) and yet shows up with fairly low differences in the Gap variables (and not too different from those of Kerala).

7. PROFITS(DERIVED) AND PROFITS(DIRECT): PROFILE OF ENTERPRISES

Data show 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 investigated the distribution of the Gap(profits) and the characteristics of the enterprises based on the size of Gap(profits) and these are presented in Table 3.

More than 60.0 percent of the enterprises had profits(direct) lower than profits(derived); 28.0 percent reported profits(direct) higher than profits(derived);

TABLE 3
CHARACTERISTICS OF ENTERPRISES BY SIZE OF GAP (PROFITS)

(1)	Profits(direct) < Profits(derived)			Profits(direct) = Profits(derived)			Profits(direct) > Profits(derived)			Profits (direct) < Profits (derived) (11)	Profits (direct) = Profits (derived) (12)	Profits (direct) > Profits (derived) (13)	All (14)
	>-20.0% (2)	[-20, -10%] (3)	[-10, -5%] (4)	[-5, 0%] (5)	0% (6)	(0, 5%] (7)	(5, 10%] (8)	(10, 20%] (9)	>20% (10)				
Number of enterprises	9,031	18,875	18,569	51,894	17,078	24,461	6,834	4,815	6,218	91,369	17,078	42,328	1,50,775
Percent of enterprises	6.0	7.9	12.3	34.4	11.3	16.2	4.5	3.2	4.1	60.6	11.3	28.1	100.0
Mean total workers	3.7	2.9	2.8	2.3	2.3	2.7	2.6	2.7	3.1	2.9	2.3	2.7	2.8
Mean number of hired worker by ownership of enterprise	2.0	1.3	1.1	1.1	0.7	1.0	1.0	1.1	1.6	1.2	0.7	1.1	1.1
Proprietary enterprises ^a (%)	93.4	95.7	96.6	96.5	97.2	96.7	96.5	96.7	94.4	96.1	97.2	96.4	96.3
Male-owned (%)	81.1	79.3	78.0	78.8	64.6	77.4	74.5	75.3	78.2	78.9	64.6	76.8	76.7
Female-owned (%)	12.4	16.4	18.6	17.8	32.6	19.4	22.0	21.0	16.2	17.2	32.6	19.6	19.6
Non-Proprietary enterprises ^b (%)	6.6	4.3	3.4	3.5	2.8	3.3	3.5	3.3	5.6	3.9	2.8	3.7	3.7
Own-account enterprises ^c (%)	48.9	62.9	66.9	68.4	79.6	69.2	68.6	66.0	49.3	65.4	79.6	65.8	67.1
Enterprises maintaining business accounts (%)	8.2	4.6	3.6	3.1	4.0	3.0	3.7	3.6	6.9	3.9	4.0	3.8	3.9
Enterprises located within household premises	39.6	50.5	53.1	54.6	64.1	55.3	53.6	51.1	39.9	52.3	64.1	52.3	53.6
Enterprises registered with any local authority (%)	38.2	25.2	21.7	20.1	12.8	19.2	19.3	21.0	30.1	22.9	12.8	21.0	21.2
Mean total receipts (output) (rupees)	618,743	299,610	227,641	217,174	158,258	209,368	181,777	191,540	273,061	269,707	158,258	212,242	240,951
Mean total expenses (rupees)	510,195	249,860	184,777	172,589	125,986	170,517	150,396	161,403	259,035	218,478	125,986	179,235	196,984
Mean intermediate consumption (rupees)	454,794	217,213	157,915	146,947	108,840	146,173	125,687	136,459	217,880	188,736	108,840	152,294	169,456
Mean compensation of employ (rupees)	46,365	27,740	22,699	21,526	14,192	20,583	20,704	20,189	29,775	25,027	14,192	21,908	22,924
Mean Profits(derived) (rupees)	108,548	49,750	42,864	44,586	32,273	38,851	31,381	30,137	14,026	51,229	32,273	33,007	43,966
Mean Profits(direct) (rupees)	54,124	42,686	39,792	43,703	32,273	39,522	33,608	34,352	36,730	43,806	32,273	37,569	40,749
Gap(Profits) (%)	-50.1	-14.2	-7.2	-2.0	0.0	1.7	7.1	14.0	161.9	-14.5	0.0	13.8	-7.3
Mean GVA(P) (rupees)	163,949	82,397	69,726	70,227	49,418	63,196	56,091	55,082	55,181	80,971	49,418	59,948	71,495
Mean GVA(I) (rupees)	109,525	75,334	66,654	69,344	49,418	63,866	58,318	59,296	77,885	73,547	49,418	64,510	68,277
Gap(GVA) (%)	-33.2	-8.6	-4.4	-1.3	0.0	1.1	4.0	7.7	41.1	-9.2	0.0	7.6	-4.5
Mean size of plant and machinery (rupees)	50,825	30,398	25,030	24,088	17,861	22,603	23,128	21,346	43,641	27,742	17,861	25,635	26,032

^aProprietary enterprises are those where an individual is the sole owner of the enterprise.

^bIncludes the following: (a) *partnership within members of household*; (b) *partnership between members outside household*; (c) *co-operative society*; (d) *limited company*; and (e) *others*. Partnership is defined as the "relation between persons who have agreed to share the profits of a business carried on by all or any one of them acting for all." Co-operative society is a society that is formed through the co-operation of a number of persons (members of the society) to benefit the members. A limited company can be either private or public company. A private company means a company which by its Articles (a) RESTRICTS the right to transfer its shares, if any, (b) LIMITS the number of its members (not including its employees) to 50, and (c) PROHIBITS any invitation to public to subscribe for any shares or debentures of the company. Public company means a company which is not a private company.

^cAn own account enterprise is an undertaking run by household labour, usually without any hired worker employed on a "fairly regular basis." By "fairly regular basis" it is meant the major part of the period of operation(s) of the enterprise during the last 365 days.

and for nearly 12.0 percent of the enterprises, the two measures of profits were surprisingly equal. The basic characteristics of these three types of enterprises are given in columns 11, 12, and 13 in Table 3. At the aggregate level, the enterprises with non-zero Gap(profits) have more or less similar characteristics in terms of ownership, local registration, maintenance of accounts, and location of enterprise within household premises. However, the enterprises with profits(direct) lower than profits(derived) had higher mean employment, mean P&M size, mean total receipts, mean total expenses, mean profits(derived), and mean profits(direct), as compared to the other two types of enterprises. If we investigate the characteristics of the enterprises across distribution of Gap(profits), it is seen that nearly 90 percent of the enterprises with Gap(profits) in the range of 0–20 percent on either side of the distribution have quite similar characteristics in terms of mean total workers, mean hired workers, type of ownership, location of enterprise, and levels of registration. However, the mean receipts, mean expenses, and mean size of P&M are increasing with the increase in the Gap(profits) for the enterprises with profits(direct) less than profits(derived). This is not the case for the enterprises on the other side of the distribution, where the enterprises with Gap(profits) in the range of 0–5 percent have higher expenses and receipts compared to enterprises with differences in the range of 5–20 percent. Now looking at the characteristics of the enterprises with Gap(profits) larger than 20 percent on either side of the distribution, these enterprises clearly have characteristics that are distinct from the rest and are bigger in size as compared with the rest. In particular, the enterprises with profits(direct) less than profits(derived) and with differences larger than 20 percent are clearly big in terms of all characteristics—size of employment, hired workers, total receipts, total expenses, mean profits, mean GVA, and mean size of plant and machinery and contributed most to the Gap(profits). There is some evidence to suggest that enterprises with relatively larger incomes tend to underreport profits when confronted with a direct question to reveal these profits. However, the fact that 28 percent of enterprises reporting higher profits(direct) than profits(derived) also suggests that, while deliberate misreporting of profits could be a reason for large Gap(profits), recall errors due to oral enquiry also contribute to inaccurate information. All these suggest the real challenges in obtaining accurate data through oral enquiries in surveys of informal sector enterprises that seek information in one visit for the reference period. The third category of enterprises for which the two measures of profits were equal were much smaller in size and production as compared to the enterprises that have differences in the two measures.

An interesting observation from the data (Tables 2a and 2b) is that while 0.7 percent 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 micro enterprises. In our sample, around 0.7 percent of enterprises report zero profits(direct). Thus, while some enterprises reported zero profits(direct), i.e. no profit and no loss, none of them reported a loss or negative profits(derived). On the other hand, some of the enterprises with positive profits(direct) had negative profits(derived) using detailed data on their receipts and expenses.

8. WHAT DO WE CONCLUDE ABOUT UNORGANIZED MANUFACTURING ENTERPRISES FROM THE ABOVE?

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

- (1) On 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 during the reference period gave higher profits compared with 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 average, GVA(P) was higher than GVA(I).
- (2) This was true across various enterprise characteristics such as rural or urban, response code, characteristics of the informant, enterprise type, maintenance of accounts, registration, location of enterprise, employment, size of P&M, and state in which the enterprises are located.
- (3) The correlations between the profits(derived) and profits(direct) as well as those 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 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 Andhra Pradesh, Karnataka, Kerala, and Tamil Nadu compared with other states.
- (5) Our analysis also suggests that if the target enterprises for a researcher are very small in terms of employment and size of P&M, a short questionnaire with a few direct questions would yield results close 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 recall errors by eliciting information through probing questions in any survey inquiry.
- (6) Assuming that the informant is honestly providing information, one interpretation of this could be that as smaller enterprises have much simpler operations and fewer transactions compared to a larger enterprise, the recall lapses are expected to be much less, and therefore the two approaches give much closer results for smaller enterprises.
- (7) 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 that are close to or above the threshold income limits

and are not paying taxes will have an incentive to underreport incomes and profits. The NSS data however, does not capture information on whether the enterprises paid any taxes on incomes in order to study the difference in behavior of taxpayers and non-taxpayers. It is also sensitive about including such a question as this would make the enterprise suspicious and wary of responding to the survey. Reasons for underreporting of profits and revenues could also be due to an expectation of receiving benefits under some government scheme.

- (8) Although Approach 1 yielded higher profits and GVA compared with Approach 2, there are still possibilities that the enterprises underreported revenues and/or overstated expenses. Moreover, even though the profits(derived) on average were higher compared to profits(direct), they 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 RESEARCH

Considering that in the Indian context the enterprise surveys of NSSO are used to provide benchmark estimates of GVA per worker for estimating the contribution to GDP, and considering also that a single direct question provides lower estimates of GVA, there are severe limitations in resorting to a single direct question approach. However, 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. Nevertheless, it would be interesting for NSSO to undertake pilot surveys 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 NSS 56th round by obtaining the perception of the interviewer through the question “Does the investigator feel that there is any underreporting of net surplus?”—i.e. whether or not the enterprise underreported its direct(profits). A follow-up question for a “Yes” reply required the investigator to report the range of the profits (lower value and higher value) as perceived by her/him. These questions had an inherent bias of assuming that the enterprise would only underreport its net surplus (profits), which may not be true. Unfortunately, the data on these questions has also 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 range of profits of an enterprise in a single-visit interview, as was attempted in the NSS 56th round. Approaches such as those used in the experiments of de Mel *et al.* (2009) in Sri Lanka, which aim at getting such information through indirect questions, could provide a more meaningful understanding of the extent to which enterprises overreport expenses or underreport revenues and profits. de Mel *et al.* (2009) also find fears of income tax in their experiments with firms in Sri Lanka as reasons for misreporting, which may also be the case in India. Further methodological research is needed to test: (i) a short questionnaire versus a long questionnaire with data being collected in a

single visit; (ii) a long questionnaire that collects data, especially on flow variables, in a single visit to the enterprise as against multiple visits to the enterprise; and (iii) introduction of diaries to the enterprises to record daily transactions. Multiple visits and the diary method, while increasing the costs of inquiry, are expected to reduce recall errors and yield more accurate estimates of flow variables.

REFERENCES

- Casley, D. J. and D. A. Lury, *Data Collection in Developing Countries*, Oxford University Press, New York, 1981.
- CSO (Central Statistical Organisation), *National Accounts Statistics 2009*, Ministry of Statistics and Programme Implementation, Government of India, New Delhi, 2009.
- Charmes, J., "The Contribution of Informal Sector to GDP in Developing Countries: Assessment, Estimates, Methods, Orientation for the Future," Paper presented at the 4th Meeting of the Expert Group on Informal Sector Statistics (Delhi Group), Geneva, August 28–30, 2000.
- Deaton, A. and M. Grosh, "Consumption," 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*, World Bank, Washington DC, 2000.
- de Mel, S., D. J. McKenzie, and C. Woodruff, "Measuring Microenterprise Profits: Must We Ask How the Sausage is Made?" *Journal of Development Economics*, 88, 19–31, 2009.
- ECLAC (Economic Commission for Latin America and the Caribbean), *Draft Concept Note for the National Advocacy Workshop on the Informal Sector and Informal Employment*, Castries, St Lucia, 2007.
- ILO (International Labour Office), "Employment and Social Protection in the Informal Sector," Paper presented at the 277th Session of the Committee on Employment and Social Policy, Geneva, 2000.
- Kolli, R., "The Informal Sector in the National Accounts of India," Paper presented at the International Conference on Experience and Challenges in Measuring National Income and Wealth in Transition Economies, organized by the International Association for Research in Income and Wealth and the National Bureau of Statistics of China, Beijing, September 18–21, 2007.
- Kulshreshtha, A. C., "On Measuring Informal Sector: Conceptual and Estimational Issues," Paper presented at the Special IARIW–SAIM Conference on Measuring the Informal Economy in Developing Countries, Kathmandu, September 23–6, 2009.
- Kulshreshtha, A. C. and G. Singh, "Contribution of Unorganized Sector in the Indian Economy," *Manpower Journal*, 34(3), 45–67, 1998.
- Liedholm, C., "Small Scale Industries in Developing Countries: Empirical Evidence and Policy Implications," Paper No. 9, MSU International Development Papers, Michigan State University, Michigan, 1987.
- , "Data Collection Strategies for Small-scale Industry Surveys," GEMINI Working Paper No. 11, Bethesda, MD, 1991.
- OECD, *Measuring the Non-Observed Economy, A Handbook*, OECD, Paris, 2002.
- Scott, C., *Sampling for Monitoring and Evaluation*, World Bank, Washington DC, 1985.
- Suwal, R. and B. Pant, "Measuring Informal Sector Economic Activities in Nepal," Paper presented at the Special IARIW–SAIM Conference on Measuring the Informal Economy in Developing Countries, Kathmandu, September 23–26, 2009.
- SNA, *System of National Accounts 1993*, Commission of the European Communities—Eurostat, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank, Brussels/Luxembourg, New York, Paris, Washington DC, 1993.