Statistical Journal of the IAOS 33 (2017) 547–556
DOI 10.3233/SJI-160281
IOS Press

Data anomaly in mining statistics of India

Molly Chattopadhyay[a,*] and Anupam Lahiri[b]
[a]*Indian Statistical Institute, Economic Analysis Unit, Bangalore, India*
[b]*Central Statistical Organization, Kolkata, India*

Abstract. Comparison of official statistics on mining reveals that there are different agencies that are collecting and disseminating data on the mining sector, but there also exists data anomaly between different agencies. For the purpose of detecting data anomaly in the mining sector, official statistics provided by three government agencies are examined. The variation in the number of mining workers (both organized and unorganized sector) ranges from 2.2 million to 0.54 million to 0.52 million in 2011 between the National Sample Survey Organisation, the Director General of Mines Safety, and the Indian Bureau of Mines which clearly indicates the differences in methods for collecting and collating of official statistics.

Keywords: Mining, National Sample Survey Organisation, Director General of Mines Safety, Indian Bureau of Mines

1. Introduction

Good quality official statistics is indispensable for making informed decisions. This article explores data anomaly in official statistics in the mining sector of India. Considering the pivotal role of the mining sector in the Indian economy, both for research and policy-making, compilation of official statistics in the mining sector assumes importance. Official statistics are defined to be collated numerical facts published by the government agencies or other public bodies such as international organizations [1]. The problem with India's official statistics is that various government agencies collect statistics that are not comparable to each other. Detecting data anomalies is an important step in understanding and improving data quality. When data integrity is compromised, the veracity of the decision-making process is likewise threatened. The problem with poor quality data is that it leads to poor decisions. Data are of high quality if they are fit for their intended uses in operations, decision making and planning [2].

A data anomaly is not the same as data defect. A data anomaly might be a data defect, but it might also be accurate data caused by unusual, but actual behaviour of an attribute in a specific context. Data anomalies have also been referred to as outliers, exceptions, peculiarities, surprises, and novelties [3]. Chandola and colleagues refer to data anomalies as patterns in data that do not conform to a well-defined notion of normal behaviour [4]. This is similar to how Hawkins defines an outlier as observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism [5]. Johnson defines outlier as an observation in a data set which appears to be inconsistent with the remainder of that set of data [6].

There is an existing literature that provides valuable information about discrepancies in the reporting of job characteristics of people [7–10]. Another relevant literature is a set of studies that have compared the effects of labour market programs estimated using survey data collected from individuals to the effects estimated from administrative unemployment Insurance wage records obtained from employers [11–13]. In the Indian context, deteriorating state of official statistics was elaborated upon by Sharma [14], particularly in the field of consumption expenditure from National Accounts Statistics and National Sample Survey. Questions regarding credibility and reliability of India's official statistics were also raised by several researchers [15–17]. Saluja and Yadav [18] drew attention to the reliability of estimates in the case of minor minerals by National Accounts Statistics. A detailed

*Corresponding author: Molly Chattopadhyay, Indian Statistical Institute, Economic Analysis Unit, Bangalore, India. E-mail: molly@isibang.ac.in.

analysis of sources, methods of compilation and identification data gaps of labour statistics – both at the central and state level in India is prepared by Papola [19].

Despite their potentially important implications for a variety of labour market analyses, no existing study has systematically examined data anomaly in the mining sector of India. This article, therefore, attempts to find data anomaly between different sources, thereby analyzing inconsistency in definition and coverage of mines and employment in the mining sector. Section 2 describes sources of data on the mining industry; Section 3 portrays the National Industrial Classification of Industries; Section 4 depicts background on the mining sector of India; Section 5 explains data anomaly in the number of mines and mining employment between government agencies and the reasons behind anomaly; Section 6 describes discrepancy in the number of organized and unorganized sector employment with conclusions in Section 7.

2. Sources of data on mining industry

Mining sector data are provided by the following government agencies: i) Director General of Mines Safety based on official records, ii) Indian Bureau of Mines based on official records, iii) National Sample Survey Organization based on household sample survey and iv) Census of India based on total enumeration.

2.1. Director General of Mines Safety (DGMS)

The Directorate General of Mines Safety, entrusted to enforce the Mines Act, 1952 throughout the country, receives various statutory Returns and Notices under the Coal Mines Regulations, 1957 and in respect of metalliferous and oil mines (non-coal) under Regulation 5 of the Metalliferous Mines Regulations, 1961 and the Oil Mines Regulations, 1984 framed under the Mines Act, 1952.

The publication entitled "Statistics of Mines in India" by DGMS [20] contains information regarding employment, production, productivity, average weekly wages, index of labour earnings in coal and non-coal mines along with other associated aspects such as data on output, value of minerals, gassiness of seams, mechanization, consumption of explosives etc.

In addition, these also contain information regarding accidents in coal mines, metalliferous and oil mines and brief description of findings of enquiry conducted by officers of DGMS in respect of each and every fatal accident that occurred during the reference year. An updated list of all fatal accidents involving four or more deaths and a list of Court of Inquiries held for different accidents in coal mines since 1901 are also included in this publication. Since a large number of metalliferous mines are in the unorganized sector and many of them are seasonal in nature, the number of returns received is less compared to the number of mines worked during the year. The data published is based only on the information furnished by reporting mines.

2.2. Indian Bureau of Mines (IBM)

IBM operates primarily for the development of the mining industry under the statutory provisions of the Mines & Minerals (Regulation & Development) Act, 1957. The Indian Bureau of Mines (IBM) is a subordinate office under the Ministry of Mines. It is engaged in the promotion of scientific development of mineral resources of the country, conservation of minerals and protection of environment in mines, other than coal, petroleum and natural gas, atomic minerals and minor minerals. It performs regulatory functions, with respect to the relevant provisions of the Mines and Minerals (Development and Regulation) Act, 1957 and enforcement of the rules framed there under namely Mineral Conservation and Development Rules, 1988 and Mineral Concession Rules, 1960 and Environmental (Protection) Act, 1986 and Rules made there under. It also, undertakes scientific, techno-economic, and research oriented studies in various aspects of mining, geological studies, ore beneficiation and environmental studies.

IBM brings out technical publications relating to mines and minerals, mineral based industries, trade, beneficiation, R&D activities, etc. The publication Indian Minerals Yearbook [21] contains information on various parameters i.e., mineral reserves and its life index, number of mining leases, lease area, number of mines, production by captive and non-captive mines, public and private sectors, A & B categories, state and grade. Data on employment, stock, principal mines, domestic consumption and exports/imports by countries are also presented in the publication. The total mineral resources as on April 1, 2011 are classified under Reserves and Remaining Resources. Data on production for 2011-12 are based on the monthly production reported by the mine owners through Annual/Monthly returns under the provisions of the Mineral Conservation and Development Rules 1988 while those for 2012-13 are based on the monthly production reported by the mine owners in the Monthly returns.

2.3. National Sample Survey Organisation (NSSO)

National Sample Survey Organisation (NSSO) has been regularly conducting employment-unemployment surveys. The quinquennial survey on employment and unemployment is one of the important surveys conducted regularly by the NSSO. The first such survey was carried out in the 27^{th} round of NSSO during October 1972–September 1973. Since then, seven comprehensive quinquennial surveys on employment and unemployment situation in India have so far been carried out by the NSSO prior to the present survey of the 68th round (July 2011–June 2012).

The main objective of the employment and unemployment surveys conducted by NSSO at periodic interval is to get estimates of various employment and unemployment characteristics at national and State/Union Territories level. The critical issues in the context of labour force enquiries pertain to defining the labour force and measuring participation of labour force in different economic activities. The aspects of the labour force are captured in detail in the employment and unemployment surveys of NSSO and estimates are generated for labour force participation rate, worker population ratio, unemployment rate, wages of employees, extent of underemployment etc. The indicators of the structural aspects of the workforce such as status in employment, industrial distribution and occupational distribution are also derived from these surveys. Besides, from the data collected on the particulars of enterprises and conditions of employment, the aspects of employment in the informal sector and informal employment are reflected through the conceptual framework of the survey. Industries are classified as per the National Industrial Classification and National Classification of Occupations.

2.4. Census of India

The Indian Census is the largest single source of a variety of statistical information on different characteristics of the people of India. With a history of more than 130 years, this reliable, time tested exercise has been bringing out a veritable wealth of statistics every 10 years, beginning from 1872 when the first census was conducted in India non-synchronously in different parts. The census provides information on size, distribution and socio-economic, demographic and other characteristics of the country's population. The Population Census is the total process of collecting, compiling, analyzing or otherwise disseminating demographic, economic and social data pertaining, at a specific time, of all persons in a country or a well-defined part of a country. The workers have been classified by the type of economic activity into nine broad categories as per the National Industrial Classification.

3. Industrial classification

The International Standard Classification of Occupations (ISCO) is an International Labour Organization (ILO) [22] classification structure for organizing information on labour and jobs. It is part of the international family of economic and social classifications of the United Nations. The ILO describes the purpose of the ISCO classification as a tool for organizing jobs into a clearly defined set of groups according to the tasks and duties undertaken in the job. It is intended for use in statistical applications and in a variety of client oriented applications. The ISCO is the basis for many national occupation classifications as well as applications in specific domains such as reporting of teaching, agricultural and healthcare workforce information.

3.1. Classifications used in India

The National Industrial Classification (NIC): In India, industrial classification has been used for conducting the Population Census, Industrial Surveys, National Income Estimates, etc. The National Industrial Classification (NIC) is the standard classification followed for classifying economic activities. The NIC [23] is prepared to suit the Indian conditions and follows the principles and procedures laid down in the United Nations' International Standard Industrial Classification (ISIC) of all Economic Activities as revised from time to time. The Central Statistical Organisation (CSO), as the custodian of national industrial classification, took up the task of evolving a standard industrial classification as early as in 1960. While undertaking a fresh revision of NIC, the suggestions obtained from various stakeholders have been taken care of to the maximum extent possible. The structure of NIC is identical to the structure of ISIC up to 4-digit level 'classes' (except a few shadow classes). Classes were then divided into 5-digit 'sub classes' according to national requirements.

National Classification of Occupations (NCO): In an occupational classification, the grouping of occupations has to be based on the fundamental criterion of 'type of work performed' [24]. All the workers en-

gaged in the same type of work are grouped together irrespective of the industrial classification of establishments where they are engaged. For example, all clerical workers have been classified in one occupational group whether they are engaged in a factory, mine, government office or even a shop. Factors like type of operations involved in the performance of a job, types of qualifications, vocational and professional training, status (e.g., own-account worker, employer), levels of skill, etc., are considered in classifying a person as belonging to a particular occupation. Job definitions or descriptions represent only the average national picture of the various occupations.

In India, the Central Statistical Organization (CSO) in the Ministry of Statistics & Programme Implementation is vested with the responsibility of setting up of a system of classification. It is a constant endeavour of the CSO to develop new classification systems as well as update existing ones to keep pace with the changes in the organization and structure of industries besides accounting for emerging economic activities.

4. Background of Indian mining industry

India is endowed with a significant volume of mineral deposits. It is estimated that India holds abundant reserves of minerals such as non-coking coal, iron ore, bauxite (metallurgical grade), dolomite, gypsum, limestone and mica; and adequate levels of reserves of minerals such as lignite, chromite (metallic), manganese, zinc, graphite; but is deficient in mineral reserves such as coking coal, chromite (refractory), bauxite (chemical grade), copper, lead, apatite, rock phosphate and kyanite.

India produces as many as 87 minerals, which includes 4 fuels, 10 metallic, 47 non-metallic, 3 atomic and 23 minor minerals (including building and other materials) and is the world's largest producer of mica blocks and mica splitting. With the recent spurt in world demand for chromate, India has stepped up its production to reach the second rank among the chromate producers of the world. Besides, India ranks 2nd in Barites, 3rd in production of Coal and Lignite and Steel, 4th in Iron ore, 5th in Manganese ore, 6th in Bauxite and crude steel, 7th in Aluminium and 8th in Copper Ore on the basis of production of minerals during 2009–10 [25].

The industry is characterized by a large number of small operational mines. The number of mines has not changed substantially over the years, because as new

Table 1
Number of working and non-working mines in India

Name of the states	No. of working mines	No. of non-working mines	% of working mines to total working mines	Total
Andhra Pradesh	587	359	15.96	946
Gujarat	522	333	14.20	855
Rajasthan	477	1399	12.97	1876
Tamilnadu	444	345	12.08	789
Madhya Pradesh	430	523	11.69	953
Karnataka	310	179	8.43	489
Orissa	191	373	5.19	564
Jharkhand	151	240	4.11	391
Chhattisgarh	119	161	3.24	280
Maharashtra	119	122	3.24	241
Goa	92	223	2.50	315
Other states*	235	235	6.39	470

Source: IBM report, 2011. *Other states include Uttarakhand, J&K, Himachal Pradesh, Kerala, Uttar Pradesh, West Bengal, Meghalaya, Bihar, Assam, Manipur, Sikkim, and Haryana.

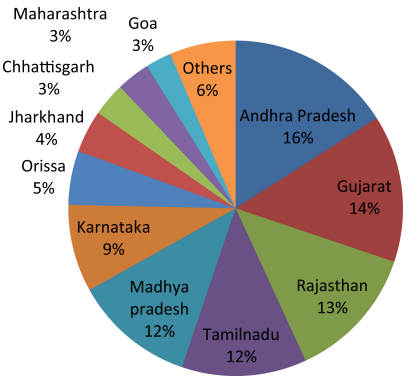

Fig. 1. State wise number of operating mines in India.

mines are explored, empty ones are closed down. The number of reporting mines during the last decade has been around 3,000 to 3,200. However, during 2010–11, it was 2,928, out of which, 573 were fuel mines, 687 were mines for metals, and 1,668 mines for extraction of non-metallic minerals. Of the total number of about 90 minerals, three minerals viz. coal (560 mines – 19% of total number), limestone (553 mines – 19% of total number) & iron ore (316 mines – 11% of total) comprised about half of the total number of reporting mines. Numbers of mines engaged in extraction were also significant in cases of bauxite (189), manganese (141), dolomite (116) & Steatite (113) [21].

The Eastern States – Jharkhand, Chhattisgarh, Orissa and Southern States – Andhra Pradesh, Karnataka as well as Rajasthan in Western India are the most important mineral-rich regions in the country. Six states viz Andhra Pradesh, Gujarat, Rajasthan, Tamilnadu, Madhya Pradesh and Karnataka comprise about 76 per

Table 2
Value of minerals (value in million rupees)

Years	Fuel	Metallic	Non-metallic	Minor minerals	Total
1951	5048	2211	1067	–	8326
1961	12097	3107	2916	1538	19651
1971	34776	6892	8678	4286	54629
1981	302279	29706	41836	23878	452319
1991	1554052	135396	52539	13272	1707959
2001	1352438	418284	48027	187345	2006094
2011	1789220	470320	61258	524904	2845702

Source: IBM report, 2013.

cent of total number of reporting mines and the rest 24% is distributed among seventeen states (Table 1 and Fig. 1).

Of the total of coal mining, maximum contribution in value terms is from Jharkhand (22%) followed by Madhya Pradesh (16%) then Chhattisgarh (15%), Orissa (11%), Maharashtra (10%) and West Bengal (7%). More than 80% of Natural gas is from Bombay high, remaining Natural Gas is contributed from Gujarat, Assam, Andhra Pradesh and Tamil Nadu. Iron Ore is one of the major contributors to the value of minerals and the state which has maximum Iron Ore is Orissa. In 2009–10, Orissa contributed 36% of the total production in value terms. Other contributors are Goa (22%), Karnataka (18%) and Chhattisgarh (16%). Therefore, number of operating mines in a State is not the only indicator of the corresponding States' contribution to the economy in value terms.

The total value of mineral production (excluding atomic minerals) estimated at 2006 billion during 2010–11 has increased by about 11.83% as compared to the previous year due to rise in the production of coal, natural gas (utilised), iron ore, manganese ore, zinc concentrates, ball clay, barites, gypsum, kaolin limestone, pyrophyllite, quartz etc. In the total value of mineral production, the fuel minerals contributed the major share 67.42%, metallic 20.85%, non-metallic minerals 2.4% and minor minerals 9.33%. The mining and quarrying sector (excluding the minerals) declares as prescribed substances under the Atomic Energy Act, 1962) accounted for about 2.6% of the total GDP [21].

In the last two decades, coal mining has witnessed a phenomenal growth of 154% in production from 237 metric tonnes in 1991 to 602 metric tonnes in 2010. The value of output of coal has increased by 1038% from 55.71 million tonnes to 634.43 million tonnes between 1961 and 2013 [20]. Taking into account the coal and non-coal sector, the value of production has increased by 67% from 1708 billion rupees to 2845 billion rupees between 1991 and 2013 (Table 2). Despite increase in production and value of the minerals, average employment in coal mines has decreased between 1991 and 2011 from 0.554 million to 0.369 million. In the non-coal sector also, employment has come down from 0.20 million to 0.17 million workers for the same period. Combining the coal and non-coal industry, employment has decreased from 0.73 million workers to 0.53 million workers; in twenty years between 1991 and 2011, the decrease is by 27% (Table 3). The strategies adopted for rapid expansion of mining activities include increased mechanisation, adoption of new technologies & their adaptation under Indian geo-mining conditions and assimilation of latest scientific innovations in the concerned areas. This may be the prime factor in the decrease of mining workers.

Table 3
Average employment in mines from 1951 to 2011 (in million)

Year	Average employment		Total
	Coal	Non-coal	
1951	0.352	0.197	0.549
1961	0.411	0.260	0.671
1971	0.382	0.235	0.617
1981	0.513	0.222	0.735
1991	0.554	0.200	0.734
2011	0.369	0.166	0.535

Source: DGMS report, 2011.

Table 4
Number of operating mines in India

	2007–08	2008–09	2009–10	2010–2011	2011–12
DGMS					
Coal	569	583	590	590	569
Non-coal	1971	1972	2002	1967	2041
Total	2540	2555	2592	2557	2610
IBM					
Coal	570	574	573	573	573
Non-coal	2453	2576	2483	2546	2663
Total	3023	3150	3056	3119	3236

Source: Annual reports: DGMS, IBM and ministry of mines 2011.

5. Data anomaly in the number of operating mines and miners

The number of mines does not change substantially over the years, because as new mines are explored, empty mines also are closed. It is seen from Table 4 that there is considerable difference in the number of mines between DGMS and IBM. In the case of coal mines, the difference is negligible, but in the case of non-coal mines, total number shown by IBM is greater than 500 than the number shown by DGMS. For non-coal mines for the year 2011–12, DGMS reports it as

Table 5
Employment in mines

	2007–08	2008–09	2009–10	2010–2011
DGMS	549075	562644	575022	538700
IBM	537327	519835	521425	518419

Source: Annual reports: DGMS and IBM, 2011.

Table 6
Differences in the number of mines

	No. of Mines		No. of Workers	
	2007–08	2010–11	2007–08	2010–11
DGMS – all mines	2540	2557	549075	538700
IBM – all mines	3023	3119	537327	518419

Source: Annual reports: DGMS and IBM 2011.

2041 while as per IBM it is 2663. There is discrepancy in the number of miners also between DGMS and IBM (Table 5). In respect of number of mines and miners, the difference between DGMS and IBM is by 21% and 5% respectively.

Let us explore the reasons behind anomalous data on the number of mines and miners. As stated in Section 1, in India, we have two major central government organisations controlling the mining sector, namely the Directorate General of Mines Safety (DGMS), previously known as Mines Inspectorate, and the Indian Bureau of Mines (IBM). The DGMS implements the provisions of the Mines Act (1952) exclusively in the field of safety and labour welfare and the IBM operates primarily for the development of the mining industry under the statutory provisions of the Mines & Minerals (Regulation & Development) Act, 1957.

Thus, although very small mines (tiny SSMs) of these specific minerals (mostly "Minor Minerals") are exempt from the provisions of the Mines Act 1952, the tiny mines of other minerals (non-Minor Minerals) and the other "Minor Mineral" Mines in the upper range of operation are not exempt and are equally liable like Medium and Big Mines.

The exemptions under the Mines Act 1952, restricting the sphere of activities of the DGMS (Mines Inspectors) in this specific category of tiny mines has been made perhaps because chances of fatal and serious accidents in such tiny mines are minimal and as such unnecessary waste of time and money in non-essential inspections and administrative control may be avoided, maintaining authority for intervention under Section 7, 8, 9, 40, 45 and 46 where ever felt necessary.

In the case of IBM such tiny Small Mines are also exempt from their control because these mines are really "Minor Minerals" mines, which are specifically exempt from the jurisdiction of IBM. The control of "Minor Mineral" has been statutorily shifted by the Central Government to the State Governments under Sec 14 and 15 of the MMRD Act'57. Thus, the officers of IBM have no authority to inspect and control the activities of "Minor Mineral" mines and thus have no jurisdiction for collecting statistical figures of production, employment, number of mines etc. from such mines. They have therefore to depend on the figures supplied by 19–20 State governments, which are not always reliable. Except for production value they do not supply the figures about the number of mines and the total employment, although they do not exempt even the tiniest of Small Mines (mostly Minor Minerals) from the payment of government revenue because such mines, working commercially, are liable to pay royalty, cess (tax) etc. on the quantum of minerals raised even though the number of such mines is in the thousands.

Therefore, both the Mines Act'52 and the MMRD Act'57 have no scope of defining Small- Scale Mining although exemptions under Sec 3 of the Mines Act are really for tiny mines and that too of 'Minor Minerals' [26, p. 15]. As the Mines Act'52 does not differentiate between major and "Minor Minerals" and there are many fairly large "Minor Mineral" Small Scale Mines (Stone mines-for example) the statistics of mines maintained by the DGMS, (including MM mines), do not tally with those maintained by IBM (excluding MM mines). Moreover, DGMS statistics are on a calendar year basis and IBM records are on the basis of financial year – April to March. Thus, the statistical records of these two organizations are somewhat different. Similarly, both the Mines Act'52 and the MMRD Act 57 do not define Small-Scale Mining and they do not maintain any statistics of Small-Scale Mines.

In this context, it may be interesting to note from Table 6 the variations in the number of mines and miners in India as reported by the DGMS and the IBM. The incongruity in the number of mines would be clear from the foregoing explanation. But it is not understood how in spite of higher number of mines reported by IBM than DGMS, the employment figures for 2007–08 and 2010–11 do not correspond to the pattern. In both 2007–08 & 2010–11, number of mines by IBM is greater than DGMS, while the number of miners is less in IBM in comparison to DGMS.

There is a difference in the data collected by NSSO, DGMS, and IBM. As per employment-unemployment surveys of NSSO (2011–12), the estimated number of miners (Table 7) is 2,239,828, four times higher than the number of miners provided by the Mining agencies

Table 7
Distribution of Mining workers (in million) by NIC as per Census 2001 & NSSO 1999–00

Sources	Persons	M%	F%
Census	2.2	86.09	13.91
NSSO	2.23	83.81	16.18

Source: Employment-unemployment survey report, NSSO, 1999-2000 and Census of India, 2011.

(DGMS & IBM). Since mines are located mostly in rural areas, 63% of the workers are located in rural areas. The change between NSSO and DGMS is by 125%. Therefore, the question arises, how to account for this huge gap between NSSO and DGMS/IBM?

In this context, Census data on the number of workers is mentioned to show the discrepancy with other agencies. As per Census 2001 [27], the number of total main workers (Table 7) in the mining and quarrying sector is 2.2 million. This is comparable to NSSO the employment and unemployment survey 2000–01; the estimate of workers in mining and quarrying sector is 2.23 million. But census data will not be discussed in this report because till date Census 2011 data by NIC is not available. Discussion will remain confined to DGMS, IBM and NSSO.

One plausible reason that is provided by DGMS is that in coal mines rosters are maintained for three shifts. One person may work for two shifts, even for three shifts; in that case of DGMS records it as one worker, while NSSO considers it as two to three workers.

Another way to match the number of workers as per NSSO with DGMS and IBM may be to distribute workers across organized and unorganized sector assuming that DGMS and IBM covers only organized sector workers. To find employment distribution among organized and unorganized sectors, the following section outlines definition of organized and unorganized sector employment.

6. Organized and unorganized sector of mining industry

There is no universally accepted definition of the unorganized sector. For the purpose of analysis, it is crucial to look into the definitions used by the different data collection agencies.

The unorganized manufacturing sector, for which data are drawn from the NSS 56th (2000–01) and 62nd (2005–06) rounds, is defined as follows: "(i) All manufacturing enterprises except those registered under Sections 2m(i) and 2m(ii) of Factories Act, 1948 and Bidi and Cigar Workers (Conditions of Employment) Act, 1966. (ii) All manufacturing enterprises except those run by the Government (Central Government, State Governments, Local Bodies)/Public Sector Enterprises" [28, p. 8]. By this definition, all those private sectors, manufacturing enterprises which engage less than 10 workers with power or less than 20 workers without power are in the unorganized manufacturing sector. Thus, an analysis of organized and unorganized manufacturing in terms of definitions is complimentary and, therefore, comparable.

The National Commission for Enterprises in the Unorganized Sector (NCEUS, 2007) [24, p. 20] defines 'organized' and 'unorganized' on the basis of various factors including enterprise type, number of workers and social benefits. All enterprises under the domain of the government/public sector, public/private limited.company; co-operatives, trusts, etc. are organized. The enterprise type is unorganized if it is proprietary (male and female); entails a partnership with members from the same household or members from different households; and employer's households (that is, private households employing maid-servants, watchmen, cooks, etc.) coupled with the number of workers, which should be 10 or more. If the enterprise type is not known (missing or other than mentioned above) and employs 10 or more workers, it is considered as organized. When both the organized type and number of workers are not known, then if the enterprise provides social benefits to its workers, it is organized. The residual sectors are considered as unorganized [29].

The National Accounts Statistics, compiled by the Central Statistical Organisation (CSO), categorises sectors as registered and unregistered wherein all public sector units are considered organized while private sector enterprises which are registered under some Act, for example, the Factories Act, Sales Tax Act or the State's Shops and Establishment Act are all organized [29, p. 8]. The rest of the private sector enterprises are considered as unorganized. According to the CSO/NAS, the organized services sector is constituted by public and private corporate sectors, while the household sector is recognised as unorganized [30].

While the criterion for the organized sector is very well defined for manufacturing, this is not the case for the services sector. The problem arises because services enterprises are not subject to an act similar to the Factories Act. Only those service enterprises that engage in some sort of manufacturing activity are required to register under the Factories Act. While

analysing the NSS 63rd Round data for the unorganized services sector, Dehejia and Panagariya state that: most private sector services enterprises, whether small or large, are officially in the unorganized sector. For instance, "large private sector banks such as the ICICI Bank and HDFC Bank and software export giants such as the Infosys, Wipro and Satyam are officially in the unorganized sector" [31, p. 17].

This suggests that there is no consensus over what constitutes the unorganized sector in India.

The estimates of organized and unorganized sector employment in manufacturing and services vary between enterprises as against household surveys. This may be because of two reasons. One is that the subsidiary employment of the worker in other enterprises/ sectors would not be known to the employer, but would be captured during a survey of workers/households [19]. Also, there may be certain sectors, especially in services, which may not be captured in enterprise surveys due to the lack of clarity in defining or difficulty in surveying (for example, in the trade sector), but are captured through surveys of workers/ households. For example, the estimates of employment in unorganized services from NSS enterprise surveys are 28 million, while those from household surveys are 81 million. Thus, while addressing worker-related issues, the policy-maker should keep in mind this distinction and the broader definition of employment and the organized as well as unorganized sectors [28].

For the present report, we have adhered to the definition by NCEUS [24, p. 3].

First, we define the unorganized sector in the following way:

> "The unorganized sector consists of all unincorporated private enterprises owned by individuals or households engaged in the sale and production of goods and services operated on a proprietary or partnership basis and with less than ten total workers".

Second, we define unorganized or informal employment as follows:

> "Unorganized workers consist of those working in the unorganized enterprises or households, excluding regular workers with social security benefits, and the workers in the formal sector without any employment/social security benefits provided by the employers".

The employees with informal jobs generally do not enjoy employment security (no protection against arbitrary dismissal), work security (no protection against accidents and illness at the work place) and social security (maternity and health care benefits, pension, etc.) and therefore any one or more of these characteristics can be used for identifying informal employment.

Table 8
Distribution of mining& quarrying workers by enterprise type, 2011–12

Enterprise type	Number of workers	Percent
Proprietary male	1152743	51.46
Proprietary female	13126	0.58
Partnership same hh	15297	0.68
Partnership diff hh	30719	1.37
Govt/public sector	490228	21.88
Public/private Ltd.	293203	13.09
Co-operative	38379	1.71
Others	206134	9.21

Source: Employment-unemployment survey report, NSSO, 2011–12.

Table 9
Social security benefits as per organized & unorganized sector in 2011–12

Types of benefits	No. of workers	Percent
Eligible for PF + pension	69805	3.11
Only gratuity	12765	0.57
Health care + maternity	8265	0.37
PF + pension + gratuity	16107	0.72
PF + pension + health care + maternity	44129	1.98
Gratuity + health care + maternity	5641	0.25
PF + pension + health care + gratuity + maternity	507485	22.65
Not eligible for any	1575632	70.35
Total	2239829	100.00

Source: Employment-unemployment survey report, NSSO, 2011–12.

For our purpose, we describe unorganized sector workers as those workers who are not entitled to any type of social security benefits irrespective of the workers location in organized or unorganized sector. First, by enterprise type, organized and unorganized sector employment is differentiated. Then, receipt of any kind of social security benefit by sector is delineated. Those, who received any form of benefit is demarcated as formal sector worker; rest are unorganized sector worker. Organized sector covers: a) government/public sector, b) public/private limited company, c) cooperative societies/trust/other non-profit institutions. The rest belongs to unorganized sector.

In view of the aforementioned definition, including government/public sector, public/private Ltd/cooperative as organized sector in the mining sector, total number of organized sector workers is 821,810 (37%). The remaining 1.4 million workers (63.3%) are in unorganized sector (Table 8). On the other hand, in terms

Table 10
Organized sector employment in mining sector, 2011

Definition of organized sector	Total number of workers
Enterprise	821810
Social security benefits	664196
DGMS	452653

Source: Employment-unemployment survey report, NSSO, 2011–12 and DGMS, 2011.

of entitlement to social security benefits, it is found that out of total 2,239,828 workers, total number of organized sector worker is 664,196 (30%); remaining 1,575,632 workers belong to unorganized sector workers (7%) (Table 9). Consequently, it can be said that NSSO data on employment size of miners is congruent with DGMS and IBM if the definition of organized sector employment is accepted as per NCEUS (2007) [24] stipulation of entitlement to social security benefits. Nevertheless, the problem remains. IBM does not provide categorical classification of workers, so it was assumed that the total number of workers belong to organised sector employment. But, DGMS provides the total number of contract workers for both coal and non-coal sector; it is 86,047 [20]. Therefore, as per DGMS, organized sector employment is 452,653 (Table 10). Therefore, the data anomaly between DGMS and NSSO amounts to 0.21 million workers if entitlements to social security benefit are taken whereas if enterprise is the definition accepted, then the status of 0.37 million workers remain unanswered. Perhaps, the answer lies in the statistics of minor minerals which is covered by NSSO during the quinquennial survey but outside the purview of DGMS and within the purview of State governments. Lack of coordination between state and central agencies might have resulted in the huge gap in official statistics.

7. Conclusion

This article compares official statistics on mining data collected by NSSO, DGMS and IBM for the year 2011–12. Total number of mining sector workers varies from 2.2 million (NSSO, 2011–12) to 0.5 million (DGMS & IBM, 2011), a difference by 125%. The difference in the number of workers between NSSO, DGMS and IBM is related to the methods adopted for collecting data. NSSO data is based on a household survey whereas DGMS and IBM data is based on annual returns from the registered mine owners. Differences also are evident in the number of operating mines and miners between DGMS (2610) and IBM (3236), a difference of 21% and 5% respectively. Though both DGMS and IBM data are based on annual returns from the registered mine owners, the discrepancy in the operational definition of Mines Act, 1952 that DGMS follows and MMRD Act 1957 that IBM follows regarding coverage of minor minerals and its workers gives rise to the anomaly in the number of operating mines. Similarly, in case of employment size of organized and unorganized sector workers in the mining sector, variation occurs due to differences in the definition of organized sector employment. In accordance with the NCEUS [28] definition of organized sector employment as entitlement to social security benefits, the gap in the number of miners narrows between NSSO, DGMS and IBM. The question of anomaly in the size of unorganized sector employment persists.

It is quite clear from the data discrepancy that annual returns submitted by enterprises/organizations at central and state levels are not complete. Moreover, labour statistics on minor minerals are not sent to central agencies by the state labour department. Therefore, it can be concluded that effective coverage of mines and mining workers and submission of returns both at the state and central level is necessary for more accurate statistical description.

Acknowledgments

The authors gratefully acknowledge the support provided by Indian Statistical Institute. Authors are also thankful to Prof. Madhura Swaminathan for comments that greatly improved the manuscript.

References

[1] Tendulkar S. Demand for Better Statistics and Use of Data. [Internet]. PARIS21 Secretariat, OECD Development Cooperation Directorate, October 2009 [cited 2015 March 20]. http://paris21.org/sites/default/files/dasause-Tendulkar-Consortium09.pdf.

[2] Juran JM, Gryna Frank M, editors. Quality Control Handbook. 4th ed. New York: McGraw-Hill; 1988.

[3] Lazarevic A, Banerjee A, Chandola V, Kumar V, Srivastava J. Data mining for anomaly detection. In: Tutorial at the European Conference on Principles and Practice of Knowledge Discovery in Databases. 2008 Sep 19.

[4] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. ACM computing surveys (CSUR). 2009 Jul 1; 41(3): 15. DOI. 10.1145/1541880.1541882.

[5] Hawkins DM. Identification of outliers. 5th ed. Netherlands: Springer: 1980; p. 127.

[6] Johnson R, Wichern D. Applied multivariate statistical methods. 6th ed. Prentice Hall, Englewood Cliffs, NJ. 1992.

[7] Mellow W, Sider H. Accuracy of response in labor market surveys: Evidence and implications. Journal of Labor Economics. 1983 Oct 1; 1(4): 331-44. DOI. 10.1086/2534858.

[8] Bound J, Krueger AB. The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? Journal of Labor Economics. 1991 Jan 1; 9(1): 1-24. DOI: 10.3386/w2885.

[9] Bound J, Brown C, Duncan GJ, Rodgers WL. Evidence on the validity of cross-sectional and longitudinal labor market data. Journal of Labor Economics. 1994 Jul 1; 12(3): 345-68. DOI: 10.2307/2535220.

[10] Roemer M. Assessing the Quality of the March Current Population Survey and the Survey of Income and Program Participation Income Estimates. Unpublished paper, US Census Bureau. 2000 Jun 16; [cited 2015 Jan 30]. Available from https://www.census.gov/hhes/www/income/publications/assess1.pdf.

[11] Kornfeld R, Bloom HS. Measuring program impacts on earnings and employment: Do unemployment insurance wage reports from employers agree with surveys of individuals? Journal of Labor Economics. 1999 Jan; 17(1): 168-97. DOI. 10.1086/209917.

[12] Hotz VJ, Scholz JK. Measuring employment and income for low-income populations with administrative and survey data. Studies of welfare populations: Data collection and research issues. In: Ploeg MV, Moffit RA, Citro CF, editors. National Research Council. Studies of welfare populations: Data collection and research issues. Washington DC: National Academies Press; 2002; p. 275-315.

[13] Schochet PZ, Burghardt J, McConnell S. Does job corps work? Impact findings from the national job corps study. The American economic review. 2008 Dec 1; 98(5): 1864-86. DOI: 10.1257/aer.98.5.1864.

[14] Sharma, Savita. Statistical System in India. United Nations Economic and Social Commission for Asia and the Pacific. Paper submitted to the conference WB/UNESCAP/ADB/NSO Thailand-Regional Seminar for Senior Managers on Monitoring and Evaluation of Poverty Reduction Programmes. Bangkok. 2002.

[15] Agrawal A, Kumar V. How Reliable Are India's Official Statistics? East Asia Forum. [Internet] 2012 April 6 [cited 2015 May 10] Available from http://works.bepress.com/vikas_kumar/115/.

[16] Bardhan P. The State of Indian Economic Statistics: Data Quantity and Quality Issues. [Mimeo on the internet]. Berkeley, CA; 2014 [cited 2016 Jan 10]. Available from http://eml.berkeley.edu/∼webfac/bardhan/papers/EconomicStatistics.pdf.

[17] Himanshu A. Credibility Gap In India's Statistics. The Livemint [newspaper on the internet]. 2012 May 10 [Cited 2017 Feb 07]. Available from http://www.livemint.com/Opinion/zgIqH58iolAnAsUo9Ns1DP/A-credibility-gap-in-India8217sstatistics.html.

[18] Saluja MR. Industrial statistics in India: Sources, limitations and data gaps. Economic and Political Weekly. 2004 Nov; 27: 5167-77.

[19] Papola TS. An assessment of labour statistics system in India. International Labour Organization Country Office, India. 2014.

[20] Government of India. Ministry of Labour and Employment. Statistics of Mines in India. Vol I & II. Dhanbad: Director General of Mines Statistics; 2016.

[21] Government of India. Ministry of Mines. Indian Minerals Yearbook. Nagpur: Indian Bureau of Mines; 2011.

[22] International Labour Organization. International Standard Classification of Occupations. Geneva International Labour Office. 2012.

[23] Government of India. Ministry of Statistics and Programme Implementation. Employment and Unemployment Situation in India, 2011-12. Report No. 554. New Delhi: National Sample Survey Organisation; 2013.

[24] National Commission for Enterprises in the Unorganized Sector. Report on Conditions of Work and Promotion of Livelihoods in the Unorganized Sector. New Delhi: Dolphin Printo Graphics; 2007.

[25] Government of India. Annual Report. New Delhi: Ministry of Mines; 2012.

[26] Chakraborty SL. Artisanal and Small-scale Mining in India. Report No. 78. [Internet]. England. International Institute for Environment and Development. 2002. [cited 2014 Dec 20]. Available from http://pubs.iied.org/pdfs/G00724.pdf.

[27] General R. Census Commissioner (2011 A) Census Data 2011-Metadata Government of India, Ministry of Home Affairs, New Delhi, India.

[28] Government of India. Ministry of Statistics and Programme Implementation. Unorganised Manufacturing Sector in India. 2005-06: Employment, Assets and Borrowing. Report No. 525. New Delhi: National Sample Survey Organisation; 2008.

[29] Mehrotra S, Gandhi A, Sahoo B, Saha P. Creating employment in the twelfth five-year plan. Economic and Political Weekly. 2012; 47(19): 63-73.

[30] Shetty SL. Status Paper on Database Issues of the Service Sector. Economic and Political Weekly. 2007; 42(37): 3723-3731.

[31] Dehejia R, Panagariya A. Services Growth in India: A Look Inside the Black Box. School of International and Public Affairs. Columbia Programme on Indian Economic Policies. 2010. Working Paper No. 2010-4. [cited 2015 Jan 7]. Available from http://hdl.handle.net/10022/AC:P:10820.