WP-2017-015

An analysis of revisions in Indian GDP data

Amey Sapre and Rajeswari Sengupta

Indira Gandhi Institute of Development Research, Mumbai
September 2017

An analysis of revisions in Indian GDP data

Amey Sapre,Rajeswari Sengupta

Email(corresponding author): rajeswari@igidr.ac.in

Abstract

In this paper we study revisions in the annual estimates of India's GDP data. The objective of our analysis is to understand the revision policy adopted by the Central Statistical Organisation (CSO) and the issues therein. Using historic data, we study the magnitude and quality of revisions in the aggregate as well as the sectoral GDP series. We analyze the computation of the sectoral revised estimates and compare the extent of revision in growth rates from the first release to the final estimate. To understand the magnitude of revisions, we compute the standard deviation of revisions in growth rates for each sector and use that to build confidence bands around the initial estimates. The confidence bands provide a means to understand the extent of variation in the final growth rate estimate, and at the same time, provide a mechanism to contain revisions. Based on our analysis, we highlight some of the major issues in CSO's revision policy. We outline possible solutions that can be implemented to improve the quality of GDP data revisions. We identify sectors with large variations in growth rates and argue that improving or changing the low quality indicators can help contain growth rate revisions and enhance the credibility of the estimates.

Keywords: GDP, National Accounts, Revisions, India

JEL Code: E00, E01, C18

An analysis of revisions in Indian GDP data*

Amey Sapre†
amey.sapre@nipfp.org.in

Rajeswari Sengupta‡
rajeswari@igidr.ac.in

September 13, 2017

Abstract

In this paper we study revisions in the annual estimates of India's GDP data. The objective of our analysis is to understand the revision policy adopted by the Central Statistical Organisation (CSO) and the issues therein. Using historic data, we study the magnitude and quality of revisions in the aggregate as well as the sectoral GDP series. We analyze the computation of the sectoral revised estimates and compare the extent of revision in growth rates from the first release to the final estimate. To understand the magnitude of revisions, we compute the standard deviation of revisions in growth rates for each sector and use that to build confidence bands around the initial estimates. The confidence bands provide a means to understand the extent of variation in the final growth rate estimate, and at the same time, provide a mechanism to contain revisions. Based on our analysis, we highlight some of the major issues in CSO's revision policy. We outline possible solutions that can be implemented to improve the quality of GDP data revisions. We identify sectors with large variations in growth rates and argue that improving or changing the low quality indicators can help contain growth rate revisions and enhance the credibility of the estimates.

Keywords: GDP, National Accounts, Revisions, India

JEL: E00, E01, C18

*The views expressed in the paper are those of the authors and not of their respective institutes.

†National Institute of Public Finance and Policy

‡Indira Gandhi Institute of Development and Research

Contents

Tables

Everywhere around the world, the future is uncertain; in India, even the past is uncertain
– *Y. V Reddy*

1 Introduction

In India the Central Statistical Organisation (CSO) under the Ministry of Statistics and Programme Implementation has been releasing annual estimates of Gross Domestic Product (GDP), among other macroeconomic aggregates in the National Accounts, since 1956 (see, CSO (1993), Kolli (2007), CSO (2012), among others). Compilation of GDP is a complex exercise and requires a combination of inputs such as appropriate methods of computation and vast amounts of data across multiple sectors. Since collection of sectoral data is time consuming, the GDP numbers for any given year are released in a sequence of revised estimates based on different levels of data availability.

The sequence of revisions is supposed to reveal the true picture of the economy as and when data becomes available. Revisions in GDP data also happen when a new base year series is introduced. The revisions in such a case may be due to a combination of changes in the methodology of computation as well as sources of data. While revisions in GDP estimates due to base year changes have been the main focus in recent discussions in academic and policy circles, the quality of the GDP estimates is more influenced by the periodic rounds of revisions, which in turn, affect macroeconomic forecasting and policy making. In this paper we attempt to analyse the quality of the periodic revisions done by the CSO in the annual GDP growth estimates.

Initial estimates of Indian GDP for a particular financial year are available roughly one quarter from the start of the financial year. Thereafter, five rounds of revisions take place between the time the CSO publishes its initial and final estimates. The revision cycle gives an indication of the direction in which the economy is headed. In the literature on GDP revisions, several scholars have argued that data revisions contain both 'news' and 'noise' about the economy's growth performance. (see for instance Mankiw and Shapiro (1986), McKenzie et al. (2008)). This view is based on the fact that as initial estimates are typically compiled with incomplete data or proxies based on high frequency indicators, there is likely to be more noise in the these estimates. Gradually over the revision cycle, as more data become available, the extent of noise is expected to diminish, and the revised estimates start reflecting 'news' about the state of the economy.

From the stakeholders' perspective, GDP data revisions pose several challenges as growth rates are used to infer about the direction and momentum in the economy. A major challenge in this context is to understand and distinguish routine data revisions from actual changes in the underlying growth momentum and macroeconomic fluctuations. Moreover large and frequent revisions in the official GDP data raise questions of data credibility. An immediate

consequence of major revisions in growth rates is that it can lead to imprecise data inputs for policy making and present an uncertain environment for formulating business decisions.

Revisions in GDP have been studied extensively in several countries, mostly in the developed world. There is a sizeable body of international literature that use econometric methods to analyse the accuracy and reliability of the revised estimates i.e. the extent to which the initial estimates can predict the final estimates. For example, Young (1993), Fixler and Grimm (2008), and Fixler et al. (2014) analyse the reliability of GDP data revisions for the US, while Roodenburg and Reije (2006) analyse the same for Netherlands. These studies are based on the availability of long quarterly data on GDP revisions, and hence are able to apply sound statistical techniques to assess the quality of the revisions.

1.1 The problems with India's GDP revisions

Some of the problems with the GDP data revisions have been highlighted in Shetty (2006),Shetty (2012) and Nagaraj (2017). There are claims of overestimation in the recent episodes of revision in GDP data. Questions have been raised about the inconsistencies between the GDP numbers and other high frequency indicators that are used to compile the initial estimates.

An important aspect of the initial estimates is that most stakeholders tend to make their immediate business or policy decisions based on the first growth projection of the economy. Unless the revision cycle is clearly understood, the information contained in the first projection may not be sufficient and reliable for incorporating in decisions.

Secondly, due to the time lag in the release of the successive rounds of revised estimates, the information about the true picture of the economy is more likely to lose its relevance given that it is not available when needed for policy discussions. At a conceptual level, it maybe worth noting here that compiling the national accounts as per the guidelines of the System of National Accounts (SNA) is not sufficient in itself to avoid the problems of revisions. GDP revisions have their own character and are determined primarily by the sources and methods of computation specific to a country. Thus, a systematic effort is needed to understand the process using which revisions are made and the issues therein.

Finally and most importantly, lack of information disseminated by the CSO makes it harder to understand GDP revisions. CSO's revision cycle only states the time and use of data in each round of revised estimate. In absence of a comprehensive and consistent revision policy published by the CSO as well as analysis undertaken by the agency to inform the stakeholders from time to time about the accuracy, relevance and reliability of the revisions, very little information on GDP revisions is publicly available. Lack of consistent data on revisions over a long period of time also make independent econometric analysis difficult, thereby preventing the use of statistical techniques applied in the international literature to assess the quality of GDP data revisions.

1.2 Our paper

To the best of our knowledge, there is no detailed study in India that systematically analyses GDP revisions. In our paper we attempt to fill this gap by providing a comprehensive analysis of the revisions in the annual GDP growth estimates. We obtain the information on revisions from CSO's time line of the release of the estimates, various press releases over the years and the sources and methods of National Accounts. Unlike the international studies in this field, we are unable to work with quarterly GDP data estimates due to multiple reasons.

First, the methodology for computing quarterly GDP estimates is different from that of the annual estimates, which makes it difficult to compare revisions across different base year series. Second, the quarterly estimates are all based on high frequency indicators and they do not get revised based on the availability of actual sectoral data. Third, and a related point, while internationally, the GDP revision cycle is available at the quarterly frequency, in India the revision cycle of GDP estimates is available only for annual numbers.

For a comprehensive analysis, we study the revisions both at the aggregate level and for the various sectors that are included in the production approach of computing GD.[1]. We document in detail the process followed by the CSO to revise the annual GDP estimates, including the kind of data and indicators used at each stage to compile the estimates. We compare the magnitude of the revisions in the annual estimates across various rounds and use descriptive statistics commonly applied in other countries to analyse the revisions over time.

We focus on the revisions in constant prices annual growth estimates because they are of immediate relevance in policy, especially when in India the entire discussion focuses almost entirely on real growth rates of GDP. It is widely accepted that constant price estimates are used to infer about the real contribution by each sector and it is important to first understand revisions in those figures. Current price estimates capture a different story of the economy, and analysis of the revisions in these estimates would be taken in future work.

The overall motivation is to study historic revisions and understand the problems they pose for various stakeholders. Specifically, we ask two questions: (i) historically what has been the magnitude of revisions at the aggregate and sectoral level and (ii) how do the revised estimates perform with regard to reliability and predictability?

In addition to the aggregate revisions, we also study the sectoral ones. Our objective there is to identify the sectors that are affected by large revisions, and hence draw attention to the high frequency indicators used in these sectors to compile the initial estimates. We argue that such indicators maybe the potential source of problems. Finally, we develop a method to construct a confidence interval around the initial estimates. The idea is that the confidence band will aid in securing a range within which the final estimates are expected to be contained.

[1]We focus on GDP estimates based on the production approach because we want to analyze the estimates for each sector. The expenditure side of GDP has five components of aggregate demand and in principle does not allow us to analyze any particular sector

As a result of our analysis, we are able to identify some key problem areas in the manner in which revisions are done in India's annual GDP estimates. We highlight the main issues and also propose some solutions to improve the quality of the revisions and hence the GDP estimates, by way of a revision policy. However, due to the absence of a long time-series of data on the revised estimates, we are not able to undertake any detailed econometric analysis of the reliability and predictability of the data revisions, unlike the practice in developed countries. While such an empirical exercise would be useful, this requires the CSO to make the necessary data publicly available to researchers.

The problem areas we highlight in the paper predominantly focus on the magnitude of revisions across various rounds, specifically the degree of variability between the initial and the final estimates and the time taken by the CSO to publish the final estimates for any given year.[2] These in turn are related to several underlying issues such as the quality of the indicators used by the CSO to compile the initial estimates, and the dependence on primary surveys to collect actual data, among other limitations in the sources and methods of GDP estimation.

We also find a lack of transparency in the CSO's revision policy on several aspects, compared to international practices. The ideal revision policy of the national statistical agency should contain discussion on the relevance, reliability, and accuracy of the GDP estimates to convey a transparent picture to the various stakeholders. Regular analysis must be undertaken to assess the revised estimates on grounds of these metrics. In India however, the CSO does not have a comprehensive revision policy that includes these elements. This also implies a lack of systematic time-series data on GDP revisions, which makes regular tracking of the magnitude of data revisions difficult.

The paper is structured as follows. In Section 2 we summarize the time line and the process of revising the GDP estimates as followed by the CSO. In Section 3 we analyze the extent of revisions in the annual GDP growth rates from the initial estimate to the final estimate, both at the aggregate level and at the sectoral level. To analyze the extent and quality of revisions, we use techniques adopted by other countries and modify them to suit the Indian context. In Section 3.3, we summarise the issues involved in GDP revisions and the quality of indicators used to compile the initial estimates. We build a case for constructing confidence bands around the initial estimates to gain predictability of the final estimate. Finally, in Section 5 we end with a discussion on some feasible ways to improve the overall quality of GDP data revisions.

2 The process of GDP data revisions

In principle, computation of GDP aggregates follows the basic procedures outlined in the System of National Accounts (SNA). The country specific national accounts incorporate most of the international procedures, but allow for deviations in methodology on account of

[2] In case of India, the final GDP estimate is released about three years after the initial estimate.

limitations in source data. Typically, the source data for GDP compilation is of two types, (i) primary data and (ii) survey based data.

In case of India, some example of primary data sources include collection of statistics from public administration, national census, land records, tax collections and other routine government functions such as maintenance of health records, budgets etc. Data is also collected through sample surveys conducted within a time frame. In India, these are commonly known as rounds of NSS conducted by the National Sample Survey Organization (NSSO). Each round collects a variety of statistics such as on employment, consumption expenditure, housing, health and sanitation, and so on. Apart from official sources of data collection, statistics are also obtained from the private sector for activities that do not form a part of existing surveys or have not been included in any official system of collecting data.

The periodicity and availability of the primary and the survey data do not match. Hence, different computation methods are required to deal with the absence of data at the time of computing a particular estimate. In terms of compilation, the first category of GDP estimates is known as 'direct estimates'. These are based on data available on an annual frequency (for example for sectors such as electricity, gas and water supply and mining and quarrying etc). The second category is known as 'indirect estimates'. These are compiled for sectors or economic activities for which data is not available on a regular basis and for different frequencies. In most cases, the indirect estimates are derived from the results of surveys and are extrapolated for the years in between two consecutive surveys (for example, employment figures from the quinquennial National Sample Surveys are used in the Labor Input and Effective Labour Input Method). The extrapolation process involves constructing benchmarks, based on which estimates of the previous year are moved forward for subsequent years, till the results of a fresh survey become available.

In terms of coverage, direct estimates are limited to the formal or the organized sectors of the economy. These sectors include the activities of the public sector and the registered private corporate sector. The indirect estimates on the other hand cover the unorganized sections of the economy, including households, unincorporated enterprises, non-profit institutes serving households (NPISH) and parts of unorganized manufacturing and services. A typical survey period ranges from three to five years and actual estimates are available only after the completion of one full round of survey.

Since GDP estimates need to be produced on an annual and quarterly basis, several adjustments and approximations have to be made to adjust for the absence of regular data. The usual practice is to first compile initial estimates that are based on extrapolation of the previous year's estimate. Several high frequency indicators that capture the level of economic activity across different sectors are used as a basis for the extrapolation. However, since extrapolated values of the previous year are not a true depiction of the current state of the economy, the GDP estimates have to be revised periodically as and when actual source data becomes available.

The process of revising GDP estimates is long and cumbersome as data availability for each sector varies considerably. To facilitate the process of formulating policies and annual budgets, and to provide an overall picture of the state of the economy to various stakeholders from time to time, several intermediate GDP estimates are computed. These are the Advance Estimates, (AE), Quick Estimates (QE) and Provisional Estimates (PE). Over time, as actual source data is made available for various sectors, intermediate estimates are revised. These are termed as First (1st RE), Second (2nd RE) and Third (3rd RE) Revised Estimates. As we move further on the revision cycle, the revised estimates are assumed to get closer to the actual state of the economy.

While the process of compilation and revision is more or less similar across all countries, the singular factor that makes a difference in the accuracy, relevance and reliability of the estimates is the quality of source data. These include the high frequency indicators for various sectors, as well as the survey based data from the quinquennial rounds. To get to a sense of the compilation process of these estimates, we analyze the details of each estimate in the revision cycle and discuss the issues associated with each successive revision.

2.1 Time line of revisions

Presently, the annual and quarterly GDP estimates are released as per the advance release calender published by the CSO. The calender also provides brief details of each of the estimates and summarizes the time line starting from the Advance Estimates (AE) for a financial year till its final stage of completion, i.e. the 3rd Revised Estimate (3rd RE). The time line is reproduced in Table 1. The final estimates are compiled from a variety of data sources, some of which are direct in nature, while others are survey based. However, while computing the initial estimates, like the AE and the PE, neither direct, nor survey data are available. The only resource to compile these initial estimates is to use high frequency indicators.

Gradually, as actual data become available from direct sources as well as surveys, the estimates are re-compiled. Thus, methods and data sources change as we move from initial estimates (AE & PE) to the final estimates. In short, the process entails (i) moving from limited data to complete data on high frequency indicators, (ii) moving from indicators to direct data sources(for sectors where direct data are available) and (iii) moving from indicators to indirect data sources (for sectors where data come from surveys)[3]. For a complete description of the sector wise data sources used in the revisions, see Table 14 in the appendix.

The practice of compiling AE was introduced in 1993-94. The practice till recently was to follow up the Advance Estimate with the release of the Revised Estimate (RE). The RE was followed by the Quick Estimate (QE) for the year, and the final estimate of the sector was released after two revisions of the QE. In recent times, the nomenclature and type of estimates have been revised to include two categories of Advance Estimates, viz. 1st Advance Estimate (1st AE) and 2nd Advance Estimate (2nd AE). The Revised Estimate (RE) has been renamed as Provisional Estimate (PE), while the Quick Estimate (QE) is now called

[3]See Kolli (2007), CSO (2002), CSO (2013), CSO (2016) for further details

Table 1: Details of advance release of GDP estimates

Estimate	Data period (months)	Method/Indicators	Release Month	Time elapsed from 1st AE
First Advance Estimates (1st AE)	7-8	Benchmarked to PE of the previous financial year. Based on extrapolation using indicators such as; IIP, 1st AE of crop production, expenditure of Central & State Govt., sales tax, deposits & credits, passenger and freight earnings of railways, civil aviation, no. of telephone connections, etc	January	–
Second Advance Estimates (2nd AE)	9	Benchmarked to 1st RE of the previous financial year. Based on extrapolated values of earlier indicators like; IIP, financial performance of listed companies, 2nd AE of crop production	February	–
Provisional Estimate (PE)	12	Based on 12 month data on previously used indicators	May	2 months
First Revised Estimate (1st RE)	12	Based on detailed information budgets of govt., financial statements of public and private corporations, 42 crops, horticulture, animal husbandry and forestry	January	10 months
Second Revised Estimate (2nd RE)	12	Based on actual expenditure figures available from the govt. budgets, accounts of public and private corporations and local bodies, figures from ASI for Manuf. in place of IIP	January	1 year 10 months
Third Revised Estimate (3rd RE)	12	Improved coverage of govt., public and private corporations and accounts of local bodies	January	2 years 10 months

Compiled from CSO (2016)

the 1st Revised Estimate (1st RE). The remaining two revisions after the Quick Estimate or the 1st RE, are now called the 2nd and 3rd RE.[4].

The main purpose of an early release of GDP aggregates is to facilitate the preparation of annual budgets and provide a reasonable projection of the economy for the upcoming fiscal year. As mentioned earlier, the initial estimate is based on extrapolated values of the previous year's PE by the growth rate of the respective, representative indicator for each sub-sector. The AEs of the various sub-sectors are compiled at both constant and current prices. For constant prices, the benchmark PE for each sub-sector are extrapolated by the growth rate of the representative or key indicator. For current prices, first, the implicit price deflators are calculated from the Wholesale and Consumer Price (WPI) Indices for each category. Next, the current price values for each category are obtained by inflating the constant price values by the deflator.

The PE is the first full year estimate of aggregate GDP as well as sub-sector numbers as they are based on 12-month data on indicators. The 1st RE is the revision of the PE based on data available from the various sectors, as shown in Table 1. As we move further down

[4]The details of the changes can be understood from CSO (2012), CSO (2013) and CSO (2016)

the revision cycle, more primary data sources are used and almost all indicators get replaced by indirect estimates obtained from surveys. For example, the 2nd *RE* adds to the 1st *RE* by incorporating actual expenditure figures available from budgets and also by replacing the high frequency Index of Industrial Production (IIP) with data from the Annual Survey of Industries (ASI) for the manufacturing sector. With the finalization of budgets and financial accounts, the 3rd *RE* is considered the final estimate of the year, which is available after a lag of 2 years and 10 months. To delve into the details of each estimate and its subsequent revision, we need to understand the use of indicators and the data used for computation.

2.2 Use of indicators

The AE for the current year typically sets the macroeconomic tone for the GDP numbers of a particular fiscal year and provides the first overall picture of the state of the economy for that year. To visualize its composition in detail, Table 2 tabulates the description of the various sector-wise indicators that are used to compile the *AE*. Table 3 presents the coverage and data frequency of the sector-wise indicators.

The methodology for compiling the *AE* indicates that the reliability of the estimate depends on the strength of the indicators in capturing the level of economic activity in their sector. In cases where a high frequency indicator is not available or used, averaging previous years growth rates remains the only recourse for making a projection for the current year. Since the *AE* depends on the strength and quality of the indicator, it is imperative to analyze the choice and capacity of the indicator in each sector. Using the information available in the CSO's Sources and Methods, we summarize some important issues involving the indicators.

To begin with, the Sources and Methods do not elaborate on a particular choice of the indicator for any sector. The only explanation that can be gleaned from various official documents is that the indicator captures the level of economic activity of that sector and is available at a high frequency. Other possible explanations are that such indicators show a high correlation with the growth of value added in the sub-sector. Unfortunately, the information is inadequate to decide whether such indicators are indeed sufficient in capturing the level of the economic activity in any given sector.

Secondly, high frequency indicators are by construct noisy indicators of the growth performance of the sub-sector they represent. Also, most high frequency indicators suffer from seasonality and are influenced by business cycle conditions. For instance, components of the Index of Industrial Production (IIP) are used as indicators for the registered manufacturing sector. Typically, manufacturing output has a seasonal variation and fluctuates with other business indicators. Similarly, sales of vehicles, tax collections, passenger & freight revenue, among others suffer from seasonality on account of various events over the calender year.Furthermore, the element of seasonality may also change over time. Given that most of the indexes are based on a fixed sample frame of production units, each indicator has a limited ability in capturing the level of economic activity in its sector. In the compilation process, we cannot ascertain whether data are seasonally adjusted before they are used for

Table 2: List of indicators used for Advance Estimates of GDP, NAS: 2012

Sector	Indicator
1. Agriculture Principal Crops	Advance estimates of crop production *Other crops*: Average of past few year's growth rates *Inputs*: Previous year's input-output ratio
2. Livestock Milk, egg and wool	Targets/projections *Other products*: Average of past few year's growth rates
3. Forestry Fuel wood	NSS consumer expenditure surveys *Other items*: Average of past few year's growth rates *Inputs*: fixed ratios of output, as in the case of previous year
4. Fishing inland and marine fish	Quarterly production data *Inputs*: Previous year's input-output ratio
5. Mining and quarrying	*Coal and crude petroleum*: Monthly Production data *Other Items*: Index of Industrial Production (Mining) *Inputs*: fixed ratios of output, as in the case of previous year, separately for fuel minerals and other minerals
6. Manufacturing	Index of Industrial Production (Manufacturing)
7. Electricity, gas and water supply	*Electricity*: Index of Industrial Production (Electricity) *Water Supply*: budget estimates of central government revenue expenditure deflated by CPI(IW) *Gas*: average of past few year's growth rates
8. Construction	*Pucca*: Production of cement, steel, coal and IIP (27) *Kutcha*: average of past few year's growth rate
9. Trade, hotels & restaurants	Gross Trading Index, which is computed using the value of output of commodity producing sectors and imports
10. Railways	Net tonne Kms. and Net passenger Kms. The two indicators are combined using the weights of respective earnings
11. Transport by other means	*Road*: Number of commercial vehicles on road, estimated using the data on production of commercial vehicles *Water*: Cargo handled at major ports *Air*: passenger kilometers flown and freight tonne kilometers flown (both domestic and international) *Services*: average of past few year's growth rates
12. Communication	Total stock of telephones, both fixed line including WLL and Cellular
13. Banking and insurance	*Banking*: Total of aggregate deposits and bank credits deflated by the wholesale price index for the sub-sector, *Insurance*: Net premium received on life and non-life insurance business deflated by the wholesale price index
14. Public administration	Central and state government revenue expenditure deflated by consumer price index (industrial workers)
15. Other services	For the public component, budget estimates of central and state government revenue expenditure deflated by consumer price index (industrial workers) and for the private part average of past few year's growth rates.

Compiled from CSO 1993, CSO (2012)

extrapolation. Other than seasonality, the coverage of the indicator also poses certain limitation. Presently, a single indicator is considered as representative of the entire sector. While this may be appropriate for sub-sectors with specific types of economic activities such as; Electricity, Gas and Water supply, or Railways, the requirement may be different for sectors with diverse economic activities. Presence of seasonality in the indicator data and its limited ability leads to two main difficulties in the quality of AE. First, seasonality and limited data capture brings more noise than news in the quality of the estimate. It may also lead to an

Table 3: Coverage and data frequency of indicators

1.	Indicator	Coverage	Sector
2.	Production of food grains	NA	Agriculture
3.	Production of cement	Monthly	Registered Manuf.
4.	Consumption of finished steel	Monthly	Registered Manuf.
5.	Electricity production (IIP Electricity)	Monthly	Registered Manuf.
6.	Index of Industrial Production (IIP)	Monthly	Registered Manuf.
7.	Sales Tax collections	Monthly	Trade
8.	Gross Trading Index (GTI)	Monthly	Trade
9.	Pvt. Corporate growth in	Monthly	Trade/restaurants
10.	Hotels & restaurants	Monthly	Trade, hotels etc.
11.	Railway passenger & freight earnings	Monthly	Trade and Transport
12.	Civil aviation passenger & freight earnings	Monthly	Trade and Transport
13.	Cargo handled at major ports	Monthly	Trade and Transport
14.	Telephone connections	Monthly	Communication
15.	Govt. expenditure	Monthly	
16.	Deposits and Credit	Monthly	Finance and Banking
17.	Sale of commercial vehicles	Monthly	

Compiled from various CSO press releases

incorrect assessment of the growth potential of the economy as extrapolation based on noisy indicators cannot convey the actual acceleration or deceleration of growth in the economy.

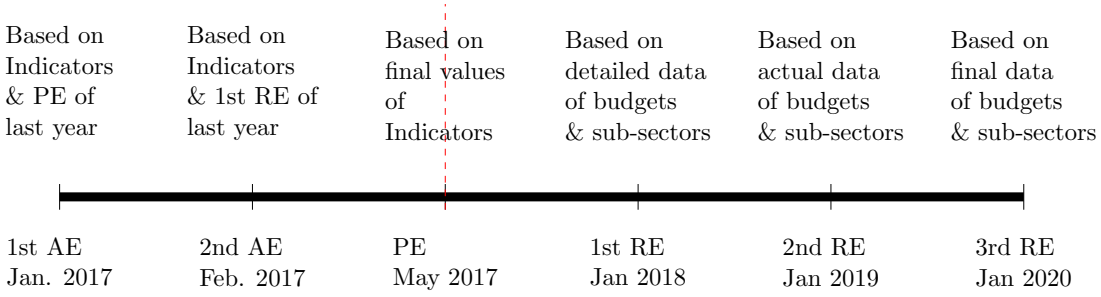
Second, limited coverage of the sector by a single indicator opens the scope for a large revision in levels and growth rates when actual data becomes available over the revision cycle. There are several reasons for the possibility of large revisions. A majority of indicators are volume based, i.e. they capture the growth in physical volume of commodities. The growth rates of volumes (adjusted by index weights) are then applied to previous year's level estimates to obtain the current period projection. As we move down the revision cycle, the level estimates of each sub-sector are re-estimated using a different set of data and methodology. As mentioned earlier, in the manufacturing sector, the volume based IIP index is used to produce the Advance Estimate, whereas the final estimate of the sector is computed using the MCA21 data for the private corporate sector and the Annual Survey of Industries for the unincorporated enterprises. There is a considerable difference in the process of estimating value addition, and since actual annual firm level data may or may not show similar trends as that of the indicator, the final estimates may show large revisions.

Similarly, for a variety of services, the final estimates are produced by methods like the Labor Input (LI) or Effective Labor Input (ELI) method, which in-turn are based on estimates of employment in person-jobs and average productivity of different types of labor. In the initial estimates for the same services, several high frequency indicators like growth in vehicle sales, telecommunication, tax collections, etc. are used. These indicators are most likely to have a different character as compared to actual growth in employment and changes in average productivity of labor. Thus, it is more likely that high frequency indicators that contain noise and pick different signals about the sector's performance may overestimate (or

underestimate) the current period's projection.

The choice of high frequency indicators is guided more by the notion of data availability and coverage, than accuracy. However, a detailed assessment of data on indicators and its choice can help in identifying the sources of divergence between the initial and final estimates. This notion of divergence clearly brings out the difference between the element of noise and news in the aggregates. If we assume that estimates compiled using actual and detailed data reflect the true growth performance of the economy, we can contend that across successive rounds of revisions, the extent of noise diminishes and the proportion of news contained in the revised estimates increases. To visualize this feature, consider the depiction in Figure 1.

Figure 1: Time line of GDP data revisions:
From Noise to News



Given the time line of the estimates, we can assume that part of the actual performance of the economy emerges after the Provisional Estimate (PE) are released. In the above example, some news about the state of the economy till May, 2017 can only be known in January, 2018. In an ideal situation, the first projections ought to be closer to the actual performance of the economy. This also implies that subsequent revisions will corroborate the *news* of the growth performance of the economy and that initial estimates would not get revised substantially. The question to ask is- does that happen at the aggregate and sub-sector level?

In the next section we attempt answer this question by comparing the estimates as per the revision cycle for aggregate GDP and its sub-sectors. We also attempt to decipher, what explains the magnitude of revisions and can the magnitude and direction of revisions be predicted. We argue that comparing the extent of revisions sheds light on the quality of indicators and data sources and allows us to identify sectoral estimates that are subject to large revisions.

3 Magnitude of revisions in GDP data

The magnitude of revisions can be analyzed in both level and growth estimates. The same analysis can be done at constant and current prices. We choose to analyze revisions in growth rates at constant prices instead of level estimates to maintain comparability across different base year series. Ideally, revisions of both levels and growth rates need to be studied for a

detailed analysis of the estimates. However, given different base year series and simultaneous changes in price series (i.e WPI, CPI), the level estimates are more difficult to analyze.

3.1 Revisions in aggregate GDP growth rate

We begin by tabulating the revisions in the annual growth rate of GDP at 2004-05 constant prices We use the 2004-05 base year series because it provides the entire revision cycle for several years. We do not get a complete set of revised estimates from the earlier series or for the latest 2011-12 series. For earlier series, the available data are reproduced in the appendix, both at constant and current prices.

In January 2015, the new 2011-12 base year series of the National Accounts was introduced. The new series led to changes in sources and methods of GDP computation and discontinuation of the 2004-05 series. Similarly, with the introduction of the 2004-05 series, the previous 1999-00 series was discontinued and the practice continues with every new base year revision. As GDP data undergoes four revisions over three years, discontinuation of the previous series before the estimates under the new series are finalized creates a data gap and makes it difficult to conduct a detailed time-series analysis of revisions. Currently, vintage data on annual GDP revisions is available from 1991-92, for the 1980-81 and 1993-94 base year series. The figures are tabulated in Tables 10 and 11 in the Appendix.

In Table 4 we tabulate the AE and the subsequently revised growth figures from 2004-05 onward, using the time line of release of estimates (see Table 1). The figures of AE, PE and 1st RE are taken from various press releases of the CSO, and the final revised growth rates from the annual NAS publication for various years.

Table 4: Revisions in annual growth rate (%) of GDP at Factor Cost
2004-05 & 2011-12 series, constant prices

Fin. Year	AE	PE	1st RE	2nd RE	3rd RE	($AE-$2ndRE)	($AE-$3rdRE)
2004-05	6.9	–	7.5	7.5	7.5	−0.6	−0.6
2005-06	8.1	–	9.0	9.4	9.5	−1.3	−1.4
2006-07	9.2	–	9.6	9.7	9.6	−0.5	−0.4
2007-08	8.7	9.2	9.0	9.3	9.3	−0.6	−0.6
2008-09	7.1	–	6.7	6.7	6.7	0.4	0.4
2009-10	7.2	7.4	8.0	8.4	8.6	−1.2	−1.4
2010-11	8.6	8.5	8.3	9.3	8.9	−0.7	−0.3
2011-12	6.9	–	6.2	6.7	–	0.2	–
2012-13	5.0	5.0	4.5	5.4	5.4	−0.4	−0.4
2013-14	4.9	4.7	6.6	6.3	6.2	−1.4	−1.3
2014-15	7.4	–	7.2	6.9	–	0.5	–
2015-16	7.6	7.2	7.8	7.9	–	−0.3	–
2016-17	7.1	–	–	–	–	–	–

Compiled from various press releases of annual estimates of GDP and various releases of NAS. AE is Advance Estimate, PE is Provisional Estimate
1st, 2nd& 3rd RE are First, Second and Third Revised Estimates, respectively
All figures are in percentages.

To assess the magnitude of revision from the initial to the final estimate of the GDP growth rate and how it has changed over the years, we compute the difference between the *AE* and 2nd *RE*, i.e. (*AE*− 2nd *RE*) and report this metric for each year. Ideally one should consider the 3rd *RE* for computing the difference, but given the data gaps for few years, we use the 2nd *RE* to maintain consistency across time periods.

The direction in revisions (given by *AE*−2nd *RE*) suggests that in most cases *AE*s have been an underestimation of the actual GDP growth rate, as the growth rate was revised upwards in subsequent revisions. There are cases of overestimation as well, as the growth rate was downwardly revised in later revisions. We argue that there are few important learnings from this revision cycle; (i) at the aggregate, the AE's may underestimate the current year's growth and (ii) on average, the absolute magnitude of the revision is close to 0.5%. Over the revision cycle, it is also important to note the difference between the 1st and 2nd *RE*. The 1st *RE* acts an intermediate step that conveys the direction in which the economy is heading. In most cases, the 1st *RE* corrects the underestimation (or overestimation) of the AE and conveys the growth trajectory of the economy. The revision cycle also suggests that in most cases, the 2nd and 3rd revised estimates are closer and are consistently in one direction. Barring a few cases where the 3rd *RE* was downwardly revised, almost all final revisions have been on the higher side.

The pattern from the revision cycle further suggests that at the aggregate level, we cannot ascertain whether AE of overall GDP growth are biased (in either direction) due their design or the quality of indicators. Since the *AE*'s are benchmarked to the *PE*'s of the previous year, the extent of bias in either direction is determined in part by (i) under or over estimation of the Provisional Estimate and (ii) bias in the indicators at the sub-sector levels. The figures also reveal that the absolute magnitude of the revisions between the initial and final estimate has not declined over the years. Given that the 2nd and 3rd *RE* have remained fairly close, the focus of the revisions ought to shift to the Advance Estimates as the discrepancy is much larger in those estimates.

Analyzing revisions at the aggregate level has limitations. Since GDP is an aggregate of sub-sectors, it is much more meaningful to analyze revisions at the sub-sector level. As indicators are used to extrapolate the sub-sector values, the quality of data and the performance of the indicators is revealed more clearly at the sectoral level.

3.2 Revisions in sub-sectors growth rates of GDP

In Table 5, we compile the annual growth rate estimates for all sectors across all successive rounds of revision from 2008-09 to 2015-16. As before, we report the difference between the *AE* and 2nd *RE* for every year to assess both the magnitude and the direction of revisions. In case of the direction, we try to assess whether the AE of the various sectors over or underestimated the actual growth rates i.e. whether (*AE*− 2nd *RE* was > 0 or < 0)). In case of magnitude we see how large the revisions have been between the first and the final round of revisions for every year and for every sector.

Table 5: Growth rate revisions at the sector level,
2004-05 & 2011-12 series, constant prices

Year	Sec	AE	1st RE	2nd RE	(AE-2RE)	Sec	AE	1st RE	2nd RE	(AE-2RE)
2008-09	AGRI	2.60	0.20	0.10	-2.50	CONS	6.50	6.50	5.30	-1.20
2009-10	AGRI	-0.20	1.00	0.80	1.00	CONS	6.50	7.00	6.70	0.20
2010-11	AGRI	5.40	7.90	8.60	3.20	CONS	8.00	10.20	5.70	-2.30
2011-12	AGRI	2.50	3.60	5.00	2.50	CONS	4.80	5.60	10.80	6.00
2012-13	AGRI	1.80	1.40	1.50	-0.30	CONS	5.90	1.10	0.60	-5.30
2013-14	AGRI	4.60	3.70	4.20	-0.40	CONS	1.70	2.50	4.60	2.90
2014-15	AGRI	1.10	-0.20	-0.3	-1.40	CONS	4.50	4.40	3.0	-1.50
2015-16	AGRI	1.10	0.80	–	–	CONS	3.70	2.80	–	–
2008-09	MANF	4.10	10.80	4.30	0.20	TRD	10.30	9.30	5.70	-4.60
2009-10	MANF	8.90	9.70	11.30	2.40	TRD	8.30	7.80	7.90	-0.40
2010-11	MANF	6.20	9.70	8.90	2.70	TRD	11.00	13.80	12.00	1.00
2011-12	MANF	3.90	2.70	7.40	3.50	TRD	11.20	7.00	1.20	-10.00
2012-13	MANF	1.90	1.10	6.00	4.10	TRD	5.20	5.10	11.00	5.80
2013-14	MANF	-0.20	5.30	5.60	5.80	TRD	3.50	11.10	7.20	3.70
2014-15	MANF	6.80	5.50	7.5	0.70	TRD	8.40	10.70	8.5	0.10
2015-16	MANF	9.50	10.60	–	–	TRD	9.50	10.70	–	–
2008-09	M&Q	4.70	10.60	2.10	-2.60	FIN	8.60	9.70	12.00	3.40
2009-10	M&Q	8.70	6.30	5.90	-2.80	FIN	9.90	9.40	9.70	-0.20
2010-11	M&Q	5.80	4.90	6.50	0.70	FIN	10.60	10.10	10.00	-0.60
2011-12	M&Q	-2.20	-0.60	0.10	2.30	FIN	9.10	11.70	11.30	2.20
2012-13	M&Q	0.40	-2.20	-0.50	-0.90	FIN	8.60	10.90	9.60	1.00
2013-14	M&Q	-1.90	5.40	3.00	4.90	FIN	11.20	7.90	4.80	-6.40
2014-15	M&Q	2.30	10.80	14.7	12.40	FIN	13.70	7.90	8.9	-4.80
2015-16	M&Q	6.90	12.30	–	–	FIN	10.30	10.80	–	–
2008-09	EGW	4.30	6.50	4.60	0.30	COMM	9.30	5.60	12.50	3.20
2009-10	EGW	8.20	6.30	6.20	-2.00	COMM	8.20	12.00	11.70	3.50
2010-11	EGW	5.10	5.20	5.30	0.20	COMM	5.70	4.30	4.20	-1.50
2011-12	EGW	8.30	6.50	8.40	0.10	COMM	5.90	6.00	4.90	-1.00
2012-13	EGW	4.90	2.30	2.80	-2.10	COMM	6.80	5.30	6.30	-0.50
2013-14	EGW	6.00	4.80	4.70	-1.30	COMM	7.40	7.90	5.60	-1.80
2014-15	EGW	9.60	8.00	7.2	-2.40	COMM	9.00	11.40	9.3	0.30
2015-16	EGW	5.90	5.10	–	–	COMM	6.90	6.90	–	–

Notes: AGRI denotes Agriculture & Allied activities, MANF is Registered Manufacturing, M&Q is Mining & Quarrying, EGW is Electricity, Gas & Water supply, CONS is Construction, TRD is Trade, Hotels & Restaurants, FIN is Finance, Banking & Insurance, and COMM is Community, Personnel Service and Defense. AE is Advance Estimate, 1st RE and 2nd RE are First and Second Revised Estimates. Source: Various press releases of CSO and National Accounts
The figures are in percentages.

A sector-wise analysis shows that for the agriculture sector in the recent years the *AE*s have been underestimating the actual growth rate. The magnitude of revision between the *AE* and 2nd *RE* seems to have widened over the years in absolute terms. In between there were two years (2010-11 and 2011-12) that witnessed a large positive difference between the first and the final estimates, while it turned negative in the subsequent years. We see an opposite trend for the manufacturing sector where consistently the AE seem to have overestimated the actual growth rate. The extent of overestimation was increasing from 2009-10 to 2013-14.

In other words, during this period, the initial estimates of growth rate of the manufacturing sector showed a different picture about the state of the sector as compared to the actual situation. In 2013-14, the difference between the AE and 2nd RE was as high as 5.8%. In case of the mining sector, recent years have witnessed an increase in the overestimation in AE. In 2014-15, the extent of overestimation was the highest at 12.40%.

For the next two sectors we do not see this consistent trend of overestimation in AE, and we also do not find any consistent trend in the direction or magnitude of revisions. In case of electricity, gas and water supply the recent years show a trend of underestimation in the AE, whereas for the construction sector there is no trend over the years. The sector however witnessed some large swings in the both magnitude and direction on revisions between 2010-11 and 2014-15. For instance, the difference between AE and 2nd RE increased sharply from -2.3 in 2010-11 to 6.0 in 2011-12 and then dropped to -5.3 the next year. The revisions in this sector appear almost random, which would increase uncertainty when the estimates are used for any analysis or policy making.

The sector of trade, hotels and restaurants saw a similar sharp swing in revisions between 2011-12 and 2012-13 when the revisions went increased from -10.0 to 5.8, the largest increase in any year for any sector. Thereafter, the AE have been consistently overestimating the actual growth rate of the sector but the extent of overestimation appears to have reduced over the years. Finance, banking and insurance as well as community, personnel service and defense show a similar picture as electricity, gas and water supply with recent years witnessing an underestimation in the AE.

In summary, while growth rates of some sectors were overestimated in the initial estimates, most sectors do not show any consistent trend over the years. Manufacturing has been the only sector where the AEs were consistently overestimating of the actual growth rate.

One important take-away from this analysis is that the level of revisions at the sub-sector level may be very different as compared to overall GDP growth rate. Second, the extent of revision in both direction and magnitude in any sector remains unpredictable. In particular, a revision may even change the direction of growth of the sub-sector, whereas such directional changes are limited in case of overall GDP. It is expected that revisions at the sub-sector level will cause a change in the overall growth rate of the economy. However, in general, large revisions at the sub-sector level do not lead to a major revision in the overall growth rate.

The analysis points to the fact that considering revisions at the aggregate level alone is insufficient in understanding the state of the economy. It also indicates that sectors that are subject to large revisions in either directions have low quality indicators that produce the initial estimates. Given that revisions at the aggregate and sub-sector level may convey different point of views for different stakeholders, what are the main issues with revisions in GDP data? We summarize the main issues as follows.

3.3 Summary of issues with revisions

Revisions in GDP estimates are a part of constructing the national accounts. The statistical agency conducts the exercise periodically, just like the statistical agencies in other countries, and it operates within the constraints posed by data availability. The issue of data availability on a variety of indicators is the most serious constraint in the Indian case. The actual data on the indicators and the methodology of how the growth shown by the indicator translates into the growth of the level estimate of the PE is presently unknown. This limitation creates an element of uncertainty, especially in cases where high frequency indicators show divergent trends as compared to value addition estimates of sub-sectors and aggregate GDP.

It is equally important to de-link periodic revisions from actual fluctuations in the economy that are inevitable due to the changes in the underlying macroeconomic conditions. Subsequent revisions of initial estimates cannot be concluded as pure economic fluctuations in levels of value addition. Part of this conclusion is also because we are unable to observe consistent patterns in the revisions, and hence it becomes difficult to infer what part of the changes in the estimates is originating from data revisions and what part is triggered by macroeconomic changes.

On the data side, it is known that direct estimates are available for few sectors, while for the rest, the statistical agency has to rely on quinquennial surveys. Unfortunately, the quality and low frequency of surveys limit their use in the initial estimates and the scope of improving the subsequent revisions. To this extent, with more direct estimates and a wider and deeper coverage of the surveys, the magnitude of revisions is expected to reduce. Ironically, this has not happened in the Indian case.

The basic premise of the AE and its subsequent revision is that initially, the state of the economy can be projected for the current year based on its previous position. Subsequently, the projection can be revised, once source data becomes available. However, in this process, revisions often miscommunicate the information about the economy and are taken by surprise, especially when revisions take place contrary to expectations.

Apart from the extent of revisions, for different stakeholders, the number of revisions can also create difficulties in decision making.Ideally, with data availability, the number of revisions ought be as low as possible. In other words, the estimate ought to be finalized within the least possible number of revisions. Unfortunately, over the years, with different nomenclatures, the number of revisions have not reduced. Contrary to best practices, the number of revisions increased when the CSO added the release of the 2nd Advance Estimate in 2016.

Since stakeholders have different uses of aggregate and sub-sector data, it useful to have some directional predictability about the initial estimate, i.e. AE. Presently, there are no readily available metrics to ascertain the magnitudes and predictability of revisions. To build on the concept of revisions and their magnitudes, we first survey the literature on international practices and methods.

In the following section we summarize some of the basic metrics and the limitations in the Indian context in using such metrics. To overcome the limitation, we develop two metrics of our own to provide a measure of short and long term revision, and gain predictability about the revisions.

4 Revision metrics

4.1 International cases

Internationally, particularly in the developed countries, the magnitude of revision is calculated using a variety of techniques. This is possible because they have access to a long time series of consistently comparable estimates of GDP growth rate over successive rounds of revisions. Commonly used metrics are mean revisions (MR), and mean absolute revisions (MAR) along with the basic metric of difference between the first and the final estimates. Such detailed statistical analysis helps to get a comprehensive idea about the overall quality of data revisions, the issues involved and how they have changed over time.

Unfortunately, in India we are unable to generate a long time series of all revisions of GDP and sub-sectors across different base year series. Even for our current study data was put together from the myriad press releases and the NAS publications for various years. The absence of a repository of vintage data on revised GDP estimates makes it difficult to undertake a detailed analysis of revisions.

In the international literature on GDP data revisions, studies use statistical metrics to evaluate the accuracy and reliability of the revised estimates both at annual and quarterly frequency (see for example, Young (1993), Roodenburg and Reije (2006), Fixler and Grimm (2008), and Fixler et al. (2014)) for the respective countries. Young (1993) examines the revisions in the quarterly estimates of the US GDP for 1978-1991, for both current and constant dollar estimates. They use summary measures such as dispersion, bias, relative dispersion, relative bias, upward revisions, and directional misses, to assess the data revisions. The study primarily focuses on a discussion of dispersion and bias in the revised estimates. Dispersion is calculated as:

$$\text{Dispersion} = \frac{\sum |P - L|}{n} \tag{1}$$

where P is the percentage change in current estimates, L is the percentage change in latest available estimates and n denotes the number of quarterly changes. Dispersion gives the average of the absolute values of revisions. Similarly, bias is the average of the revisions calculated which is computed as;

$$\text{Bias} = \frac{\sum (P - L)}{n} \tag{2}$$

Relative dispersion is another metric and is expressed as a percentage of the average of the

absolute values of the latest available estimates, i.e.

$$\text{Relative dispersion} = \frac{\sum |P - L|/n}{\sum |L|/n} \tag{3}$$

Finally, relative bias expresses the bias due to revisions as a percentage of the average of the latest available estimates:

$$\text{Relative bias} = \frac{\sum (P - L)/n}{\sum L/n} \tag{4}$$

In a more recent study on the US, Fixler and Grimm (2008) analyze the reliability of the successive growth estimates using quarterly series for the period 1983-2006. To measure the reliability, they use mean revisions (MR) and mean absolute revisions (MAR) and compute the difference between the earlier estimates and the latest available estimates. The mean revision (MR) is calculated as the average of the revisions in the sample period, i.e.

$$MR = \frac{\sum (L - E)}{n} \tag{5}$$

where E is the percentage change in the earlier estimate, L is the percentage change in the later estimate and n is the number of observations in the sample period. Since the revisions can be positive or negative, the authors also look at the mean absolute revision (MAR):

$$MAR = \frac{\sum |L - E|}{n} \tag{6}$$

The summary measures are similar to those used by Young (1993) except that Fixler and Grimm (2008) look at the difference between the final and initial estimates whereas Young (1993) looks at the reverse in his dispersion measure. Fixler and Grimm (2008) find that the MARs have declined over time and conclude that quarterly estimates are reliable indicators of the growth rate of the economy and where it stands relative to the trend growth. Also improvements in source data and methodologies have contributed to the decline in MRs and MARs over time.

Roodenburg and Reije (2006) analyse the accuracy and reliability of quarterly revisions in the GDP data for Netherlands for the period 1986 to 2002. They look into the aggregate GDP estimates and also the six expenditure and ten production components that constitute GDP. They define reliability as the extent to which initial, provisional estimates are able to predict the final estimates. In particular they test whether data revisions are predictable for the quarterly estimates of GDP growth rates.

The authors categorise the revisions into short-term and long-term. The short-term revision is defined as the revision between the preliminary estimate and the revised estimate after a period of two years. Long term revision is the revision between preliminary estimate and the final estimate. To describe data revisions they look at summary statistics such as the mean, median, standard deviation (SD), root mean square error (RMSE), minimum and

maximum of the short-term and long-term revisions. They find that the mean, SD and RMSE of the long-term revisions are larger. Also the revisions on average are positive which implies that there is a downward bias in the preliminary estimates of their GDP data. They also do a forecast rationality test, which is a test of unbiasedness of the revised data. They apply hypothesis testing methods to test the presence of news versus noise in the preliminary estimates.

4.2 Indian case

In all the above cases, the authors are able to undertake a comprehensive and detailed analysis of GDP revisions because of access to a long time series of quarterly estimates. Given the limited data availability, the only summary statistics we are able to compute to describe the GDP revisions and also to help assess the reliability of the revised estimates, are the mean, minimum, maximum and standard deviation of the short-term and long-term revisions for the period 2008-09 to 2015-16, following Roodenburg and Reije (2006).

We define short-term revision as the difference between the *AE* and the 1st *RE*. The time gap between these two estimates is around 10 months. The difference in the estimates is driven by the transition from using high-frequency to actual data for various sectors. We define long-term revision as the difference between *AE* and 2nd *RE*. The time gap between these two rounds is roughly 2 years. Table 6 presents the basic summary statistics of the short-term and long-term revisions for all the sectors. Mean value denotes the average of the actual revisions over the chosen horizons, for the period from 2008-09 to 2015-16. We argue that if the preliminary estimates contain only news and no noise, then the means of the revisions should be *zero*.

Table 6: Summary statistics of growth rate revisions for various sectors, 2008-09 to 2015-16

	Short term 2008-09 to 2015-16				Long term 2008-09 to 2014-15			
Sector	Mean	Min	Max	SD	Mean	Min	Max	SD
AGRI	0.06	-2.50	2.40	1.58	0.30	-2.50	3.20	2.05
MANUF	-1.79	-6.70	1.30	3.10	2.77	0.20	5.80	1.94
M&Q	-2.85	-8.50	2.60	4.48	2.00	-2.80	12.40	5.34
EGW	0.95	-2.20	2.60	1.51	-1.03	-2.40	0.30	1.20
CONS	0.19	-2.20	4.80	2.07	-0.17	-5.30	6.00	3.68
TRD	-1.01	-7.60	4.20	3.44	-0.63	-10.00	5.80	5.27
FIN	0.45	-2.60	5.80	2.85	-0.77	-6.40	3.40	3.60
COMM	-0.03	-3.80	3.70	2.53	0.32	-1.80	3.50	2.39

Short term denotes AE−1st RE, Long term denotes AE−2nd RE
Revisions include both 2004-05 & 2011-12 series, constant prices

We find that the averages for the short-term revisions of only two sectors, namely agriculture, and community, personnel service and defense, are close to zero. For all the other sectors, there is a relatively stronger indication that the averages are biased in either upward or

downward direction for both short-term and long-term revisions, especially for the manufacturing and the mining sectors. For manufacturing, mining, trade and community services, the short-term mean revisions are negative implying there is an downward bias in the preliminary estimates of these sectors, as the subsequent estimate is higher than the initial one. What is interesting is that the signs get reversed in the long-term revisions for these three sectors. This feature suggests that the AEs contain a downward bias when compared to the 1st RE, but show an upward bias when compared to the 2nd RE. In other words, the preliminary estimates for these three sectors paint a rather subdued picture at first, but it gets reversed in the transition from 1st RE to 2nd RE.

A reverse phenomenon is observed for the electricity, construction, and financial services sectors. Agriculture is found to be the only sector where the direction of mean revision is the same across short-term and long-term horizons. The positive sign in this case implies that the preliminary estimates contain an upward bias.

We are unable perform any test of the reliability of the estimates because of lack of access to adequate data. The publicly available data does not allow us to construct any statistical measure of accuracy over and above the descriptive statistics. However, qualitatively, we can argue that the news about a particular sector's performance begins from the 1st RE, when actual data become part of the estimation.

It is known that all high frequency indicators contain some amount of noise due to leads, lags and seasonality. Thus, when we compare the AE to the 2nd RE, these effects of noise are supposed to decline and the actual picture of the sector is expected to be revealed. We do not find any consistent trend in the standard deviation of the sectoral short-term and long-term revisions. But we do observe that the SD of the revisions is considerably high for most of the sectors, especially for manufacturing, mining, and trade in the short-term and mining, construction and trade in the long-term. In other words, these sectors exhibit a wide variation in the revised estimates.

The question that follows from the lack of suitable metrics is; how do we gain predictability about the revisions? To answer the question, we adopt a simple approach of constructing confidence bands around the initial estimate to determine the range in which the final estimate may arrive. This approach benefits in two ways; (i) it helps us to make an apriori informed judgment about the growth performance of the sub-sectors that would only be revealed after the news containing data is made available, and (ii) it helps us to quantify the magnitude of revisions by constructing a single metric like standard deviation of revision. The method is outlined as follows.

4.3 Confidence bands of final estimates: A possible approach

The approach to constructing confidence bands or intervals is similar to producing error estimates in case of survey results. The conventional approach for capturing the dispersion is to use the mean and standard deviation. We contend that in absence of any other metric,

the same statistics can be used to build a case for constructing confidence bands. We use the geometric mean and geometric standard deviation of the long term actual revisions ($AE-$2nd RE) of each sector to capture the dispersion in revisions. Geometric mean is routinely used in computing indexes (such as WPI and CPI), and is more appropriate when averaging different aggregates. Since the revisions in aggregates are relative to their previous one, the geometric mean provides a suitable metric to capture the average change from the initial to final revised estimate (see Note 1 for more details).

Based on the long term actual revisions for each sector, we use the standard deviation to measure the extent of variation in the growth rate. With one standard deviation around the average, the lower and upper bounds of the interval for growth rate of each sector can be computed conveniently. Table 7 shows the mean, standard deviation and the confidence band for each sector.

Table 7: Descriptive statistics of actual revisions, 2008-09 to 2014-15 constant price series

	Geometric Mean	Geometric SD	Lower Bound	Upper Bound
AGRI	1.2	2.5	AE−2.5	AE+2.5
MANF	1.9	3.3	AE−3.3	AE+3.3
M&Q	2.5	2.7	AE−2.7	AE+2.7
EGW	0.7	3.7	AE−3.7	AE+3.7
CONS	1.9	3.2	AE−3.2	AE+3.2
TRD	1.7	5.3	AE−5.3	AE+5.3
FIN	1.6	3.5	AE−3.5	AE+3.5
COMM	1.2	2.5	AE−2.5	AE+2.5

Actual revision denotes AE − 2nd RE

For instance, the geometric mean of long term revisions (i.e. $AE-$2nd RE) for the period 2008-09 to 2014-15 is 1.2. For comparability, the value 1.2 corresponds to the geometric mean of the figures of actual revisions for the agricultural sector in Table 5 (See note 2 for an example). Values for other sectors are calculated accordingly, and a standard deviation is computed for respective sectors. We construct the lower and upper bounds of interval around the AE by applying one standard on either sides, i.e. $AE - SD$ and $AE + SD$. Next, we apply the interval to each AE of respective sector for the years 2008-09 to 2014-15. The lower (L) and upper (U) bound around each AE is presented in Table 8.

Given one standard deviation, we can obtain a range that is expected to contain the final estimate of the sector. From the standard deviation, it follows that the size of the band will be larger if we consider two or more standard deviation around the AE. However, to gain precision within the narrowest possible range, we argue that one standard deviation provides the necessary and sufficient range to contain the final revised estimate. Since mean and SD are also affected by the number of observations, computing the metric requires a long time series of revisions. Despite limited data availability, we contend that the metrics offers sufficient data points to illustrate its usefulness in constructing confidence bands. Using the lower and upper bounds of the confidence interval, we can estimate the range for the final

Table 8: Confidence bands for Advance Estimates

Sector	Year	AE	2RE	Actual Rev.	L	U	Sector	Year	AE	2RE	Actual Rev.	L	U
AGRI	2008-09	2.6	0.1	-2.5	0.1	5.1	CONS	2008-09	6.5	5.3	-1.2	3.3	9.7
AGRI	2009-10	-0.2	0.8	1.0	-2.7	2.3	CONS	2009-10	6.5	6.7	0.2	3.3	9.7
AGRI	2010-11	5.4	8.6	3.2	2.9	7.9	CONS	2010-11	8.0	5.7	-2.3	4.8	11.2
AGRI	2011-12	2.5	5.0	2.5	-0.04	5.0	CONS	2011-12	4.8	10.8	6.0	1.6	8.0
AGRI	2012-13	1.8	1.5	-0.3	-0.7	4.3	CONS	2012-13	5.9	0.6	-5.3	2.7	9.1
AGRI	2013-14	4.6	4.2	-0.4	2.1	7.1	CONS	2013-14	1.7	4.6	2.9	-1.5	4.9
AGRI	2014-15	1.1	-0.3	-1.4	-1.4	3.6	CONS	2014-15	4.5	3.0	-1.5	1.3	7.7
AGRI	2015-16	1.1					CONS	2015-16	3.7				
MANF	2008-09	4.1	4.3	0.2	0.8	7.4	TRD	2008-09	10.3	5.7	-4.6	5.0	15.6
MANF	2009-10	8.9	11.3	2.4	5.6	12.2	TRD	2009-10	8.3	7.9	-0.4	3.0	13.6
MANF	2010-11	6.2	8.9	2.7	2.9	9.5	TRD	2010-11	11.0	12.0	1.0	5.7	16.3
MANF	2011-12	3.9	7.4	3.5	0.6	7.2	TRD	2011-12	11.2	1.2	-10.0	5.9	16.5
MANF	2012-13	1.9	6.0	4.1	-1.4	5.2	TRD	2012-13	5.2	11.0	5.8	-0.1	10.5
MANF	2013-14	-0.2	5.6	5.8	-3.5	3.1	TRD	2013-14	3.5	7.2	3.7	-1.8	8.8
MANF	2014-15	6.8	7.5	0.7	3.5	10.1	TRD	2014-15	8.4	8.5	0.1	3.1	13.7
MANF	2015-16	9.5					TRD	2015-16	9.5				
M&Q	2008-09	4.7	2.1	-2.6	2.0	7.4	FIN	2008-09	8.6	12.0	3.4	5.1	12.1
M&Q	2009-10	8.7	5.9	-2.8	6.0	11.4	FIN	2009-10	9.9	9.7	-0.2	6.4	13.4
M&Q	2010-11	5.8	6.5	0.7	3.1	8.5	FIN	2010-11	10.6	10.0	-0.6	7.1	14.1
M&Q	2011-12	-2.2	0.1	2.3	-4.9	0.5	FIN	2011-12	9.1	11.3	2.2	5.6	12.6
M&Q	2012-13	0.4	-0.5	-0.9	-2.3	3.1	FIN	2012-13	8.6	9.6	1.0	5.1	12.1
M&Q	2013-14	-1.9	3.0	4.9	-4.6	0.8	FIN	2013-14	11.2	4.8	-6.4	7.7	14.7
M&Q	2014-15	2.3	14.7	12.4	-0.4	5.0	FIN	2014-15	13.7	8.9	-4.8	10.2	17.2
M&Q	2015-16	6.9					FIN	2015-16	10.3				
EGW	2008-09	4.3	4.6	0.3	0.6	8.0	COMM	2008-09	9.3	12.5	3.2	6.8	11.8
EGW	2009-10	8.2	6.2	-2.0	4.5	11.9	COMM	2009-10	8.2	11.7	3.5	5.7	10.7
EGW	2010-11	5.1	5.3	0.2	1.4	8.8	COMM	2010-11	5.7	4.2	-1.5	3.2	8.2
EGW	2011-12	8.3	8.4	0.1	4.6	12.0	COMM	2011-12	5.9	4.9	-1.0	3.4	8.4
EGW	2012-13	4.9	2.8	-2.1	1.2	8.6	COMM	2012-13	6.8	6.3	-0.5	4.3	9.3
EGW	2013-14	6.0	4.7	-1.3	2.3	9.7	COMM	2013-14	7.4	5.6	-1.8	4.9	9.9
EGW	2014-15	9.6	7.2	-2.4	5.9	13.3	COMM	2014-15	9.0	9.3	0.3	6.5	11.5
EGW	2015-16	5.9					COMM	2015-16	6.9				

Notes: AGRI denotes Agriculture & Allied activities, MANF is Registered Manufacturing, M&Q is Mining & Quarrying, EGW is Electricity, Gas & Water supply, CONS is Construction, TRD is Trade, Hotels & Restaurants, FIN is Finance, Banking & Insurance, and COMM is Community, Personnel Service & Defense. AE is Advance Estimate, 2nd RE is Second Revised Estimate, Actual revision is AE-2nd RE L and U are lower and upper bounds of the confidence band

estimate for each year for all sectors. To gain predictability, for each sector we count the number of time the final estimate has fallen within the confidence band. Using the count, we express the accuracy as a percentage of the number of times the final estimate falls within the confidence band. For example, in the case of the agriculture sector, the final estimate is contained in the interval 5 times out of 7 years. Alternatively, using one standard deviation around the AE, the final estimate is within the range 71% of the times. We compute the accuracy for each sector and tabulate the percentages in Table 9.

Comparing the figures of accuracy, we can determine the extent of growth rate revisions in each sector. Except for the sector of Electricity, Gas & Water supply, revisions in all other sectors are less predictable and vary considerably. In particular, revisions in the manufacturing and mining sectors have the least predictability as they vary substantially from their

Table 9: Accuracy of confidence bands for sub sectors

	Lower Bound	Upper Bound	% Accuracy
AGRI	AE−2.5	AE+2.5	71
MANF	AE−3.3	AE+3.3	57
M&Q	AE−2.7	AE+2.7	57
EGW	AE−3.7	AE+3.7	100
CONS	AE−3.2	AE+3.2	71
TRD	AE−5.3	AE+5.3	71
FIN	AE−3.5	AE+3.5	71
COMM	AE−2.5	AE+2.5	71

Accuracy denotes the number of times the 2nd RE falls within the confidence band

initial estimates. Statistically, if we analyze the standard deviation for each sector, larger values suggest that the initial estimates of these sector were imprecise in capturing the state of the economic activity in that sector. Larger standard deviation also corroborates few simple facts about the revisions. In some cases, revisions have altered the direction of growth by a large magnitude, while in other cases, growth rates were upwardly revised by a large magnitude. Such cases leave revisions unexplained and unpredictable in both magnitude and direction.

It is also worth asking whether the range of the confidence band will get narrower with more data availability?. In the present case we contend that the effect of an increase in the number of observation is ambiguous. Since there are no clear patterns of revisions across sectors, changes in mean and standard deviation across sectors remain unpredictable. However, to improve accuracy and gain precision, the confidence band ought to be as small as possible, but large enough to allow for revisions based on accurate and credible data.

The accuracy levels shown by the confidence band also tell us about the quality of indicators used in every sector. Thus, the focus ought to shift to the choice of indicators used in the sectors as their initial estimates are far away from the true picture of the sector.

5 Conclusion

In this paper we study revisions in the constant price growth estimates of annual GDP, both at the aggregate and at the sectoral levels. In particular, we focus on the use of high frequency indicators in preparing the initial estimates, highlight some of the problems with these estimates and draw lessons from international practices in GDP revisions. At the aggregate level the magnitude of revisions in annual GDP growth from the initial to the final estimate has been low, but the indicator based Advance Estimate tends to overstate the growth of the economy in some cases. Revisions at the sectoral level tell a different story.

We find that the extent of revision for almost all sectors is large and at the same time, the magnitude and direction of the revision remain unpredictable.

The extent of revision at the sectoral level suggests that the first projection of growth estimates are insufficient in providing information about the growth performance of the sector. This points to the inadequacy of the underlying high frequency indicators in capturing the level of economic activity in the respective sectors. Since there are methodological differences in compiling the GDP estimate using indicators and actual source data, the lack of quality data is one of the major reasons for large revisions at the sectoral level.

Our analysis has provided some new insights on understanding the process of revisions. First, the extent of revisions at both the aggregate and sectoral level has not reduced over time. With improvements in source data, the magnitude of revisions over the revision cycle is expected to fall but this has not happened in the case of Indian GDP. Second, the direction of revisions remains unpredictable at the sectoral level as there are no clear patters of revisions over the revision cycle for any sector.

We define short term revision as the difference between the Advance and the First Revised Estimate, and long term revision as the difference between Advance and Second Revised Estimate. This analysis provides the insight that on average, the extent of variability in the revisions is much higher in the long term. For sectors such as manufacturing, mining, and trade and community services, the short term revisions are negative, implying that the initial projections of the sector are underestimates, and subsequent revisions are higher than the initial ones. While a reverse phenomenon is observed for sectors such as electricity, construction and financial services, agriculture and trade remain the only sectors where the average of short and long term revisions happen to be in the same direction. Among all sectors, mining and quarrying shows the highest variability in long term revisions.

In order to gain predictability over the final revised estimate, we construct a confidence band around the Advance Estimates. The confidence band allows us to obtain a range around the Advance Estimate which is expected to contain the final revised estimate of the sector. We conduct this exercise for all sectors from 2008-09 onward. We find that the confidence band is accurate for all years for the electricity, gas and water supply sector, while the accuracy is 70% in case of sectors such as agriculture, construction, trade, financial services and community services sector. The predictability that the confidence band will contain the final estimate is lowest at 57% for sectors such as manufacturing and mining and quarrying.

We argue that having a confidence band at the time of the release of the Advance Estimate will provide useful information about the extent of revisions. While the confidence band ought be as narrow as possible, a sufficiently large enough band will provide the appropriate space for genuine and routine revisions. The analysis also serves the purpose of identifying sectors that are consistently prone to over or underestimation and helps us to isolate the high frequency indicators that produce the initial estimates for these sectors.

Our findings hint at important implications for policy in this area. Presently, revisions in GDP data by the CSO are not bound by any revision policy. The revision cycle only states

the time and use of data in each estimate. We can infer a few broad reasons for revisions in some sectors from the official press releases but not for all. Since macroeconomic aggregates are used by various stakeholders, the need of the hour is to develop a comprehensive and consistent *revision policy* that addresses not only the routine revisions, but also disseminates information on several metrics that can enhance the quality of the estimates.

The revision policy must address several key issues such as (i) dissemination of vintage data on revisions, (ii) policy of revision during change of base year or change in methodologies, (iii) information of revisions at both level and growth rates at current and constant prices and (iv) provide explanations to justify the major revisions. Since extent of revision is linked to the reliability and credibility of the underlying data, the revision policy must consider providing information using quality metrics such as confidence bands or error estimates, quality of source data and provide update about revision in past data.

Issues with revisions can have important implications on data quality, credibility and the ultimate usability of the data. Thus, in keeping with international practices, the revisions must be assessed on the basis of their *relevance* (are the estimates still relevant when they become available or is the time lag too long), *reliability* (can the estimates be used to predict the conditions prevailing in the economy), and *accuracy* (how close are the estimates to measuring the outcomes they are designed to measure).

* * * * * * *

Appendix

Table 10: Revisions in growth rates of GDP, 1980-81 series, constant prices

Year	Adv Est.	Quick Est.	Rev. Est.-1	Rev. Est.-2	Rev. Est.-3
1991-92	–	1.2	1.1	0.9	0.8
1992-93	4.2	4.0	4.3	5.1	5.3
1993-94	3.8	4.3	5.0	6.2	6.0
1994-95	5.3	6.3	7.2	7.8	–
1995-96	6.2	7.1	7.2	–	–
	7.0 (Rev.AE)	–	–	–	–
1996-97	6.8	7.5	–	–	–
	6.8 (Rev.AE)	–		–	–
1997-98	5.0	–	–	–	–
	5.1 (Rev.AE)	–	–	–	–

Source: CSO (2001), AE is Advance Estimate, Revised AE was introduced for 95-96, 96-97 & 97-98

Table 11: Revisions in growth rates of GDP, 1993-94 series, constant prices

Year	Latest Est. (base: 1980-81)	Est. released in 1999	Est. released in 2000	Est. released in 2001
1994-95	7.8	7.8	7	7.3
1995-96	7.2	7.6	7.3	7.3
1996-97	7.5 (QE)	7.8	7.5	7.8
1997-98	5.1 (Rev. AE)	5.0 (QE)	5	4.8
1998-99	–	5.8 (AE)	6.8 (QE)	6.6
	–	6.0 (Rev. AE)	–	–
1999-2000	–	–	5.9 (AE)	6.4 (QE)
			6.4 (Rev. AE)	
2000-01	–	–	–	6.0 (AE)

Source: CSO (2001), AE & QE are Advance & Quick Estimates

Table 12: Revisions in growth rates of GDP, 1980-81 series, current prices

Year	Quick Estimate	Revised Estimate 1	Revised Estiamte 2	Revised Estimate 3
1991-92	14.6	16	15.5	15.7
1992-93	13.8	13.7	14	14.1
1993-94	12.7	14.7	16	16.2
1994-95	18.1	17.3	18.4	–
1995-96	14.8	15.9	–	–
1996-97	14.2	–	–	–

Source: CSO (2001)

Table 13: Revisions in growth rates of GDP, 1993-94 series, current prices

Year	Latest Est. (base: 1980-81)	Est. released in 1999	Est released in 2000	Est. released in 2001
1994-95	18.4	18.1	17	17.4
1995-96	15.9	16.9	16.7	17
1996-97	14.2	16.5	15.9	15.9
1997-98	–	11.0 (QE)	11.9	11.8
1998-99	–	15.4	16.5 (QE.)	16.3
1999-2000	–	–	9.6 (AE)	10.5 (QE)
	–	–	9.9 (Rev.AE)	–
2000-01	–	–	–	11.4 (AE)

Source: CSO (2001)

Table 14: Sector wise use of data for revisions

Sector	Advance Estimate	Quick Est.	Revised Est.-1	Revised Est.-2	Revised Est.-3
(a) Agriculture, Forestry and Fishing	Advance Est.s of kharif agriculture production. Limited data on rabi sowings. Quarterly estimates of fish Production	Revised Est.s of agriculture production on principal crops. Provisional data on forestry and fishing. Est.s compiled at All-India level	Final Est.s of agriculture production. Forestry and fishing estimates compiled at state level. State level price data partially available	Land utilization statistics. Data from cost of cultivation studies. Est.s discussed with States	Land utilization statistics and data from cost of cultivation studies, if not available in the previous year
(b) Mining	Production of coal, crude petroleum and IIP (Mining) for 8 months	Provisional data on quantity and output from IBM and inputs/ output from Coal India Ltd./ Tata Iron & Steel Company	Revised data on quantity and output and inputs from IBM and inputs /output from Coal India Limited, (CIL) / TISCO/ Pvt. Sector/ Inputs from Oil & Natural Gas Commission, (ONGC) / Oil India Limited (OIL)	Data on minor minerals, after being finalised with States.	Data on inputs, if not available earlier
(c) Manufacturing	IIP for 8 Months	IIP	Provisional ASI results	Final ASI results	Final ASI results, if not available in the previous year
(d) Electricity, Gas and Water Supply	Index of electricity and budget estimates of expenditure of Centre and States	Annual reports of Central under-takings and some of the State under-takings for electricity; revised estimates of expenditure of Centre and States for water public sector.	Annual reports of Central and state undertakings for electricity; actual estimates of expenditure of Centre and States for water public sector; data from GAIL, M/o Non-Conventional Energy, KVIC and Municipalities	Reports/data if not available in the previous year	Reports/data if not available in the previous years
(e) Construction	Production data for 8 months on cement, steel, coal, IIP (metal and wood products)	Production data for 12 months on cement, steel, coal, IIP (metal and wood products), budget documents (revised estimates)/ annual reports	Budget documents (actual estimates)/ annual reports	Reports not available in the previous year	Reports not available in the previous year
(f) Trade	Gross trading index compiled on the basis of the output of commodity producing sectors	Budget documents (revised estimates)/ annual reports; enterprise survey results, if available	Budget documents (actual estimates)/ annual reports, provisional data on Pvt. Corporate sector; enterprise survey results	Reports not available in the previous year; data on Pvt. Corporate sector; enterprise survey results, if available	Reports not available in the previous years; enterprise survey results, if available
(g) Hotels and Restaurants	-do-	-do-	-do-	-do-	-do-
(h) Railways	Data for 8 months on net tonne kms. and passenger kms.	Annual railway budget (revised estimates)	Annual railway budget (actual estimates)		
(i) Other Transport	Data for 8 months on production of commercial vehicles, cargo handled at major ports and net tonne kms. /passenger kms. in the case of air transport	Budget documents (revised estimates)/ annual reports; enterprise survey results, if available and cargo handled in major ports	Budget documents (actual estimates)/ annual reports; enterprise survey results, if available and number of registered commercial vehicles	Reports not available in the previous year; data on Pvt. Corporate sector; enterprise survey results, if available	Reports not available in the previous years; enterprise survey results, if available
(j) Storage	Projected	Projected	Annual reports of warehousing corporations and provisional results of ASI	Annual reports of warehousing corporations and ASI final results	

Sector	Advance Est.	Quick Est.	Revised Est.-1	Revised Est.-2	Revised Est.-3
(k)Communication	Data for 8 months on postal and telecom. Revenue and on wholesale price index	Budget documents (revised Est.s)	Budget documents (actual estimates) and data on physical indicators		
(l) Banking and Insurance	Data for 8 months on aggregate deposits and aggregate credits	Provisional data on banking and annual reports of non-banking financial institutions and insurance companies	Final data on banking and annual reports of non-banking financial institutions and insurance companies	Annual reports of non-banking financial institutions, if not available earlier	Annual reports of non-banking financial institutions, if not available earlier
(m) Real Estate, Business Services, Ownership of Dwellings	Budget estimates of expenditure of Centre and States - for public part	Projections of workforce on the basis of long-term survey results, budget documents (revised estimates)/annual Reports	Budget documents (actual estimates)/annual reports	Reports not available in the previous year	Reports not available in the previous year
(n) Public Administration	Budget estimates of expenditure of Centre and States	Revised estimates of expenditure of Centre and States	Actual estimates of expenditure of Centre and States	NA	NA
(o) Other Services	Budget estimates of expenditure of Centre and States - for public part	Projections of workforce on the basis of long term survey results, budget documents (revised estimates)/annual reports.	Budget documents (actual estimates)/ annual reports	Reports not available in the previous year	Reports not available in the previous year

Source: CSO (2001)

Notes

1. **Sector details**:

Abbreviation	Sectors
AGRI	Agriculture, Forestry and Fishing
MANF	Registered Manufacturing
M&Q	Mining and Quarrying
CONS	Construction
TRD	Trade and Transport
FIN	Financing, Insurance, Real Estate and Business Services
COMM	Community, Social and Personal Services

2. **Computation of geometric mean and standard deviation**:

a. The geometric mean is calculated using the absolute values of the observations for each sector, i.e. $\mu_g = \sqrt[n]{x_1 \times x_2 \times x_3 \dots x_n}$, where n is the number of observations. For example, the data on revision in case of the agricultural sector has values $[-2.5, 1.0, 3.2, 2.5, -0.3, -0.4, -1.4]$. The geometric mean is; $\sqrt[7]{2.5 \times 1.0 \times 3.2 \times 2.5 \times 0.3 \times 0.4 \times 1.4}$, or $(3.36)^{(1/7)} = 1.189$ which is approximated to 1.2

b. The geometric standard deviation is given by; $\sigma_g = \exp\left[\sqrt{\frac{\sum_{i=1}^{n}\left(\ln\left[\frac{x_i}{\mu_g}\right]\right)^2}{n}}\right]$, where μ_g is the geometric mean.

References

CSO (1993). *National Accounts Statistics.* Central Statistics Office, Ministry of Statistics and Programme Implementation, Govt. of India.

CSO (2001). *Report of Dr. Rangarajan Commission.* National Statistical Commission, Ministry of Statistics and Programme Implementation, Govt. of India.

CSO (2002). Methodology for advanced estimates and quick estimates of national accounts. *Press Release 5th February, Press Information Bureau, Govt. of India.*

CSO (2012). *National Accounts Statistics: Sources and Methods.* Central Statistics Office, Ministry of Statistics and Programme Implementation, Govt. of India.

CSO (2013). Modification in nomenclature of annual gdp estimates released at various points of time for a year and statement on policy of revision of annual and quarterly estimates of gdp and related macro-economic aggregates. *Press Release, Central Statistics Office, Ministry of Statistics and Programme Implementation, Govt. of India.*

CSO (2016). Statement on release calendar of annual and quarterly estimates of gdp and policy of revision of annual & quarterly estimates of gdp and related macro-economic aggregates. *Press Release, Central Statistics Office, Ministry of Statistics and Programme Implementation, Govt. of India.*

Fixler, D. J., Greenaway-McGrevy, R., and Grimm, B. T. (2014). The revisions to gdp, gdi, and their major components. *Survey of Current Business*, August.

Fixler, D. J. and Grimm, B. T. (2008). The reliability of the gdp and gdi estimates. *Survey of Current Business*, February, 2008.

Kolli, R. (2007). Compilation of national accounts statistics of india: A historical perspective. *The Journal of Income and Wealth*, Vol. 29,(No. 1):93–128.

Mankiw, N. and Shapiro, M. D. (1986). News or noise: An analysis of gnp revisions. *Survey of Current Business*, May, 1986(66):20–25.

McKenzie, R., Tosetto, E., and Fixler, D. (2008). Assessing the efficiency of early release estimates of economic statistics. *Working Party on National Accounts, OECD Working Paper,*, STD/CSTAT/WPNA(2008)5.

Nagaraj, R. (2017). Quarterly gdp estimation: Can it pick up demonetisation impact? *Economic and Political Weekly*, Vol. 52,(Issue No. 10):10–11.

Roodenburg, O. and Reije, A. H. (2006). Dutch gdp data revisions:are they predictable and where do they come from? *Applied Economics Quarterly, Duncker and Humblot, Berlin*, 52 (4).

Shetty, S. (2006). Revision of national accounts statistics. *Economic and Political Weekly*, Vol. 41,(Issue No. 23):2317–2326.

Shetty, S. L. (2012). Dealing with a deteriorating statistical base. *Economic and Political Weekly*, Vol. 47(Issue No. 18):41–44.

Young, A. H. (1993). Reliability and accuracy of the quarterly estimates of gdp. *Survey of Current Business*, October, 1993(73):29–43.