The Economist INTELLIGENCE UNIT

Calculating city-level GDP in India

An assessment of methodologies and an evaluation of feasibility

July 2018

A report by The Economist Intelligence Unit

The Economist INTELLIGENCE UNIT

The world leader in global business intelligence

The Economist Intelligence Unit (The EIU) is the research and analysis division of The Economist Group, the sister company to The Economist newspaper. Created in 1946, we have over 70 years' experience in helping businesses, financial firms and governments to understand how the world is changing and how that creates opportunities to be seized and risks to be managed.

Given that many of the issues facing the world have an international (if not global) dimension, The EIU is ideally positioned to be commentator, interpreter and forecaster on the phenomenon of globalisation as it gathers pace and impact.

EIU subscription services

The world's leading organisations rely on our subscription services for data, analysis and forecasts to keep them informed about what is happening around the world. We specialise in:

- **Country Analysis:** Access to regular, detailed country-specific economic and political forecasts, as well as assessments of the business and regulatory environments in different markets.
- **Risk Analysis:** Our risk services identify actual and potential threats around the world and help our clients understand the implications for their organisations.
- **Industry Analysis:** Five year forecasts, analysis of key themes and news analysis for six key industries in 60 major economies. These forecasts are based on the latest data and in-depth analysis of industry trends.

EIU Consulting

EIU Consulting is a bespoke service designed to provide solutions specific to our customers' needs. We specialise in these key sectors:

- **EIU Consumer:** We help consumer-facing companies to enter new markets as well as deliver greater success in current markets. We work globally, supporting senior management with strategic initiatives, M&A due diligence, demand forecasting and other issues of fundamental importance to their corporations. **Find out more at eiu.com/consumer**
- **Healthcare:** Together with our two specialised consultancies, Bazian and Clearstate, The EIU helps healthcare organisations build and maintain successful and sustainable businesses across the healthcare ecosystem. **Find out more at: eiu.com/healthcare**
- **Public Policy:** Trusted by the sector's most influential stakeholders, our global public policy practice provides evidence-based research for policy-makers and stakeholders seeking clear and measurable outcomes. **Find out more at: eiu.com/publicpolicy**

The Economist Corporate Network

The Economist Corporate Network (ECN) is The Economist Group's advisory service for organisational leaders seeking to better understand the economic and business environments of global markets. Delivering independent, thought-provoking content, ECN provides clients with the knowledge, insight, and interaction that support better-informed strategies and decisions.

The Network is part of The Economist Intelligence Unit and is led by experts with in-depth understanding of the geographies and markets they oversee. The Network's membership-based operations cover Asia-Pacific, the Middle East, and Africa. Through a distinctive blend of interactive conferences, specially designed events, C-suite discussions, member briefings, and high-calibre research, The Economist Corporate Network delivers a range of macro (global, regional, national, and territorial) as well as industry-focused analysis on prevailing conditions and forecast trends.

Table of Contents

1 Executive summary

1.1 About city-level Gross Domestic Product (GDP)

Up-to-date, country-level GDP data is available for most countries, supported by broad international guidelines (the System of National Accounts [SNA], the latest version of which is SNA 2008). However, calculating city-level GDP is a more complex exercise. Few statistical agencies produce regular city or metropolitan GDP updates due to the complexity of calculating GDP, which is already a resource-intensive exercise at the national level but becomes increasingly data-intensive at the city level. Much of the required data is also not tracked at the city level, such as inter-city trade.

1.2 Proposed approach: top-down approach based on sectoral income

The Economist Intelligence Unit (EIU) conducted a high-level review of some of the estimation methodologies used for calculating either city-level GDP or regional GDP. In many cases, the statistical authorities emphasise that even these approaches remain experimental.

Based on our review and stakeholder engagement, The EIU identified a top-down approach based on sectoral income data for consideration. We believe that the **top-down approach using sectoral income data is the preferable approach in the longer term because it balances detail and resource-effectiveness.** The United States Bureau of Economic Analysis (BEA) has already established a precedent for using this approach. Implementation of this approach in India would also continue to evolve thanks to improvements to the Annual Survey of Industries (ASI) and efforts to survey the services sector. The EIU also reviewed top-down approaches that use population and employment as weights; however, an approach using sectoral income data is more likely to capture urban wealth and city-specific economic structures. In the short term, a top-down approach using household expenditure data (discussed in more detail in section 4.3) could be a more feasible alternative as it has fewer data requirements, but it is less robust and does not capture specific city characteristics.

Bottom-up approaches are extremely data-intensive and were not put forward for consideration due to concerns about feasibility and the diversion of resources from existing initiatives designed to improve the overall statistical framework in India. Bottom-up approaches are also rarely used by statistical agencies globally for city-level GDP estimates. A full review of these approaches (including those not covered in the following section) can be found in section 4, and their data requirements can be found in the appendix (section 5.1).

1.2.1 Top-down approach based on sectoral income data

Implementing a top-down approach based on sectoral income data recognises that income generated per employee varies by area. This is particularly true for urban areas, where the productivity gains from urban agglomeration are a key economic benefit. This top-down approach follows the BEA's approach[3] and is built on the premise that income by industry across districts in a city is an accurate proxy for

output by industry across all districts (that is, factors of production for each industry are similar between districts, the city and the state). This approach captures information about the economic structure in cities, as industries that are not prominent in a specific city (for example, agriculture) are unlikely to generate much income. As a result, such industries will automatically "drop out" of the equation and the estimates.

$$city\ GDP_y = \sum_{\substack{s=1 \\ c=1}}^{n,\,m} \left(\frac{state\ GDP_{s,y}}{state\ Income_{s,y}}\right) x\ Income_{s,c,y}\right)$$

s=economic sectors[i], with a total of m economic sectors
c=districts or wards (that form the city), with a total of n such districts or wards
y=year
state=state where city is located
Income=wage and salary disbursements,supplements to wages and salaries,and proprietors' income

Based on a review of the datasets, much of the income data required for this approach can be found in the Annual Survey of Industries (ASI). At present, however, the ASI only covers manufacturing. Implementing this approach would therefore require **the ASI (or similar industry surveys) to be extended to include services and agriculture.** In some cases, resource constraints may mean that estimates are driven by indicators other than income (for example, workforce growth, crop yields), although the goal ultimately should be to capture income data for most economic sectors.

This approach would also support the wider government effort to improve official statistics, as well as providing richer, industry-level data for policy-making for both national and city governments. It could therefore be considered part of a longer-term effort to assist in building robust estimates, forming part of the larger drive to improve official datasets.

1.3 Conclusion

The sectoral income approach is a robust, top-down approach. Implementing this approach would also support a broader goal to improve and increase the datasets available for national statistics. A short-term (and less robust) alternative could be a top-down approach using household expenditure data, which is less resource-intensive than the sectoral income approach. Regardless of which approach is implemented, additional work will need to be completed before city-level GDP estimates can be calculated as sampling and data collection are currently focused on state-level estimates. Should the government wish to operationalise a city-level data collection exercise, stakeholder engagement with relevant government agencies and city governments would be required.

[i] This would typically follow the national industrial classifications for consistency.

2 Overview

The Economist Intelligence Unit (EIU) was commissioned by India's Ministry of Housing and Urban Affairs (MoHUA) to evaluate methodologies for calculating city-level gross domestic product (GDP), and to assess their applicability to India. The EIU conducted a review of several approaches undertaken by other government organisations and think-tanks to estimate city-level GDP. What follows is a feasibility assessment that outlines the main opportunities and challenges for the calculation of city-level GDP. The research contained in this document is based on publicly available best-practice literature and select interviews and is a technical assessment of methodologies. The research for this project was independently conducted by The EIU between April and June 2018.

2.1 About the report

The report is structured as follows:

- Executive summary
- Overview of the project
- Overall feasibility assessment, based on data requirements and engagement with the Ministry of Statistics and Programme Implementation (MoSPI)
- Contextual information about the calculation of city-level GDP
- Review of existing methodologies
- Technical discussion of the proposed approach, based on existing methodologies or extensions of such techniques
- Appendix (which lists the data requirements for the proposed approaches and maps those requirements to publicly available government data)
- References

2.2 Calculating city-level GDP

Country-level GDP is a key macro-economic variable. Data on country-level GDP is readily available and is codified based on the System of National Accounts (SNA), the latest version of which is SNA 2008—an internationally agreed-upon set of recommendations for compiling measures of economic activity, widely adopted by national statistical offices around the world. In India, the MoSPI calculates national GDP and sets the methodology for estimating GDP at the state level. While there are three main methods for calculating national GDP (based on production, expenditure or income), countries differ in terms of both their procedures and data collection processes, which can generate comparability issues as well as accuracy and reliability challenges.

Calculating city-level GDP is a more complex exercise than calculating national-level GDP. In part, this is due to the complexity of calculating GDP, which is a resource-intensive exercise at the national level but becomes increasingly data-intensive at the city level. Furthermore, some of the data is simply

not tracked at the city level, such as inter-city trade. City-level GDP also requires clear definitions of city boundaries because GDP is defined as the output generated in a specific area within a specific time.

2.3 Review of existing methodologies

The difficulties associated with estimating city-level GDP mean that:

1. Methodologies are not standardised, resulting in the use of various proxies and estimation techniques.
2. The majority of estimates come from "top-down" approaches, which essentially use city-to-state/region[ii] or city-to-country ratios to estimate city-level GDP, using existing national or state-level GDPs. Some of these estimates are for specific regions or metropolitan areas, although the concept remains valid for the smaller city unit.
3. "Bottom-up" approaches are rarely adopted due to the extremely high data requirements. These approaches mirror the SNA 2008 but are implemented at the city level, necessitating the use of city-level geographic markers during the data collection phase (for example, tagging census data or enterprise returns at the city level).

A high-level review of some of the estimation methodologies used for calculating either city-level GDP or regional GDP is provided in the following table (including details of the methodology, where possible). It is important to note that, even in these cases, the statistical authorities emphasise that these approaches remain experimental, due to the difficulties discussed above. Regional estimation techniques have been included because they could theoretically be extended to the city level.

Technical details of the calculations are discussed in the following section.

[ii] In this particular exercise, we distinguish between city-level GDP and regional GDP. Regional is often defined as a state or larger sub-national level.

Methodology	Example of organisation using the approach	Details
Top-down **Uses population data to estimate output generated in a region**	Organisation for Economic Co-operation and Development (OECD)[iii,1]	● The OECD discussed the use of population weights to calculate metropolitan-level GDP. ● These are derived from TL3 data (where TL3 represents a territory typology used by the OECD, typically corresponding to administrative regions). This data is adjusted for specific municipal/metropolitan areas using GIS techniques to intersect the municipal boundaries with the TL3 boundaries. ● Data is somewhat more abundant at the TL3 level because OECD countries generally collect more detailed data than other countries. ● Using population data is a simple approach. Weighting by employment or income is more robust, but many countries do not collect this data in detail at the regional level.
Top-down **Uses earnings data to estimate output generated by each industry in a region**	U.S. Bureau of Economic Analysis (BEA)[2,3]	● The BEA uses personal income data to scale state-level GDP to county-level estimates. The release schedule shows a target of annual releases for metropolitan area GDP.[4] ● The county-level data is then summed to provide an estimate of metropolitan GDP, assigning activity that is measured at the state level to particular metropolitan areas. ● As the BEA notes, this relatively simple approach allows for regular, updated estimates. ● A key assumption is that earnings by industry across all counties in a state are a robust proxy for output by industry across all counties—that is, the approach assumes that the factors of production for all industries are similar between counties and the state in which they are located. ● For particular industries, such as banking and air transportation, the BEA adjusts data to accommodate additional information. For example, the shares of bank branch deposits in specific areas are taken into account, and growth levels for the air transportation industry are adjusted based on financial information from transport providers in those areas.
Bottom-up approach **Income approach, following Systems of National Accounts 2008 (SNA 2008)**	Statistics Canada[5]	● Statistics Canada released GDP data for 33 metropolitan areas, as well as nine other regions, in 2017. It plans to update these figures annually. ● Its approach is based on the income approach to calculating GDP, which essentially adds up income generated through the production of goods and services. ● While this approach sounds straightforward, it is a taxing exercise as most data used to calculate GDP is obtained from surveys that do not provide regional-level detail. Statistics Canada relies heavily on regional supply-use tables.

[iii] Discussed in a methodology note by the OECD.

Methodology	Example of organisation using the approach	Details
Bottom-up approach **Production approach, with value-add based on geographic units or enterprise data**	Statistics New Zealand[6] (As this is the preferred approach for regional GDP compilation, other national statistical offices are likely to use a similar approach)	● Statistics New Zealand details out its estimation of regional GDP. While it uses a blended approach (discussed later), one technique is a bottom-up approach that adds up value-add based on geographic units or enterprise data. ● As Statistics New Zealand notes, information on GDP is typically collected at the kind-of-activity level. Using this approach for regional or city-level GDP would therefore require the allocation of GDP activity to the correct region or geographic unit. ● This approach is more feasible for regional estimates and is the internationally preferred approach for regional GDP compilation.
Blended approach **Mainly production approach**	Statistics New Zealand[7] (As this is the preferred approach for regional GDP compilation, other national statistical offices are likely to use a similar approach)	● Where detailed survey data is not available with geographic units, Statistics New Zealand uses a top-down approach to apportion national GDP to regional GDP. ● Statistics New Zealand reviews the data and overall approach for consistency with national GDP.
Regression or correlation-based approaches **Proxied using datasets related to economic activity, such as satellite data on lights at night**	*Academics – Night light luminosity* See Henderson et al. (2012),[8] Dai et al. (2017),[9] Bhandari and Roychowdhury (2011),[10] Bickenbach et al. (2013)[11] and Chakravarty and Dehejia (2017)[12]	● These approaches use existing datasets as a proxy for economic activity, particularly when GDP data may not exist or is of poor quality. Night luminosity has been shown to be a good proxy for economic activity. ● While there are more than a few papers covering different versions of this approach, not all of them estimate GDP directly. A significant number focus on trend analysis to study correlation and growth patterns. ● Estimating GDP from this data can be problematic. For example, night luminosity data is subject to sensor saturation (that is, luminosity may not increase beyond a certain value), which means that large cities face a "cap" on their data. ● However, some papers (such as Basihos, 2016)[13] use night luminosity data to build out a provincial GDP dataset, extrapolating night luminosity growth from a base GDP year.[iv] ● Caution is required when using this data to approximate GDP. Papers such as Bickenbach et al. (2013) note that the relationship may not be stable[v] when looking at regional economic activity.

[iv] The full paper outlines the precise methodology, which uses neural networks to estimate the non-linear relationship between GDP and night luminosity.

[v] Specifically, the stability of the elasticity of observable GDP with regard to lights.

3 Overall feasibility assessment

Efforts to calculate city-level GDP are affected by data limitations. However, the least data-intensive approach (using population data) may be overly simplistic, given that the government has access to detailed datasets. Taking into consideration ongoing work and new approaches being studied by the Central Statistics Office in the MoSPI, a top-down approach using sectoral income data (covered in more detail in section 4.2) balances detail and resource-effectiveness.

The BEA in the United States has already established a precedent for using this approach. This approach would also continue to evolve thanks to improvements in the Annual Survey of Industries (ASI) and efforts to survey the services sector. A short-term alternative could be the top-down approach using household expenditure data, which is less robust but would be more feasible in the shorter term as it has fewer data requirements. Detailed feasibility studies on implementing either approach would need to be undertaken in consultation with relevant stakeholders.

Regardless of which approach is implemented, additional work will need to be completed before city-level GDP estimates can be calculated as sampling and data collection are currently focused on state-level estimates. A summary of the feasibility assessment for each approach is highlighted in the following table. For easy reference, feasibility (based on initial research on data availability; see section 5) is colour-coded as follows:

	Approach	Data limitations
Data largely available	Top-down approach, using population data	N/A
Data largely available	Top-down approach, using employment data	Available, although detailed data is only available every five years (last update was 2011–12).
Data partially available, or with reservations on coverage	Top-down approach, using sectoral employment data	Available at a high level (urban/rural stratum); potential to increase granularity of detailed data collected every five years.
Data not available	Top-down approach, using sectoral income data	Data is available for manufacturing but lacking for agriculture and services. Discussion with MoSPI officials also highlighted concerns about the granularity of the approach, which is resource-intensive.
Data partially available, or with reservations on coverage	Top-down approach, using household expenditure data	Available at a high level (urban/rural stratum and state-level data); potential to increase granularity of detailed data collected every five years.
Data not available	Bottom-up approach, based on income	Data is available for manufacturing but lacking for agriculture and services. Discussion with MoSPI officials also highlighted concerns about the granularity of the approach, which is data-intensive. Officials do not think this approach is feasible.
Data not available	Bottom-up approach, based on production (value-add based on geographic units or establishment data)	Data is available for manufacturing but lacking for agriculture and services. Discussion with MoSPI officials also highlighted concerns about the granularity of the approach, which is data-intensive. Officials do not think this approach is feasible.
	Bottom-up approach, based on production (value-add based on geographic units or establishment data)	As this approach focuses on the relationship between GDP and night luminosity, it would not be adequate to produce point estimates without existing baseline data. However, it could serve as a check for growth, or for convergence/ divergence in GDP for cities over time.

4 Approaches reviewed

In this section, we cover the more technical aspects of the various approaches considered.

4.1 Top-down approach, using population or employment data

Conceptually, this is a straightforward approach that apportions data based on population or employment weights. The accuracy of this data will depend on the level of detail or geographic areas included in the "top-level" GDP figures. The OECD, for example, uses TL3 data (where TL3 represents a territory typology used by the OECD, typically corresponding to administrative regions that are smaller than states). The robustness of the approach also relies on the accuracy of detailed population and employment data. By construct, this approach will produce data that is consistent and additive with national and state GDP data.

4.1.1 Using population as weights

$$city\ GDP_y = \frac{population_{city,y}}{population_{state,y}} \times state\ GDP_y$$

$$where\ state\ GDP_y = \sum_{c=1}^{n} p_c(state\ GDP_y)$$

where p_c = population weight of area c, as proportion of state population, y = year

$$such\ that \sum_{c=1}^{n} p_c = 1 \text{ and the state is formed from n areas}$$

An important caveat is that using population data potentially underestimates GDP, as cities are hubs of economic activity and are typically wealthier than surrounding rural areas. Population weights ignore wealth and productivity issues, as well as the economic structure of specific cities.

4.1.2 Using overall employment as weights

This approach improves on the population weights by reweighting using employment instead. This reduces the underestimation bias, as hubs of economic activity should also generate jobs. However, this approach does not account for the difference in GDP and employment generated in various economic sectors, as well as the difference in economic structure between urban and rural areas, or between different cities.

$$city\ GDP_y = \frac{population_{city,y}}{population_{state,y}} \times state\ GDP_y$$

$$where\ state\ GDP_y = \sum_{c=1}^{n} p_c(state\ GDP_y)$$

where p_c = population weight of area c, as proportion of state population, y = year

$$such\ that \sum_{c=1}^{n} p_c = 1 \text{ and the state is formed from n areas}$$

4.1.3 Using sectoral employment as weights

Provided state GDP is broken down into key economic sectors that are roughly consistent with employment numbers, this approach improves on the overall employment approach by using sectoral employment data as weights. This approach partially captures information on the economic structure and relative wealth of cities, although to a lesser extent than sectoral income data (which captures information that is closer to actual output and productivity of specific sectors).

$$city\ GDP_y = \frac{employment_{city,s,y}}{employment_{state,s,y}} \times state\ GDP_y$$

$$state\ GDP = \sum_{\substack{c=1 \\ s=1}}^{n,\ m} e_{c,s,y}(state\ GDP\ by\ sector\ s)_s$$

where $e_{c,s}$
= employment weight of city of district c for sector s, as proportion of state employment f for sector s

$$such\ that \sum_{c=1}^{n} e_c = 1, \sum_{s=1}^{m} state\ GDP\ by\ sector_s = state\ GDP$$

4.2 Top-down approach, using sectoral income data

Improving further on the sectoral employment approach, the use of sectoral income data recognises that the income generated per employee varies by area. This is particularly true for urban areas, where the productivity gains from urban agglomeration are a key economic benefit.

This top-down approach follows the BEA's approach[3] and is built on the premise that income by industry across districts in a city is an accurate proxy for output by industry across all districts (that is, factors of production for each industry are similar between districts, the city and the state). It implies that the state-level relationship between industry earnings and industry output holds for each metropolitan area.

This approach captures information about the economic structure in cities, as industries that are not prominent in a specific city (for example, agriculture) are unlikely to generate much income. These industries will automatically "drop out" of the equation and the estimates, allowing a single survey questionnaire to be developed and implemented. The bulk of the data on income (wage and salary disbursements, supplements to wages and salaries, and proprietors' income) can be derived from data on emoluments and operating surplus in the Annual Survey of Industries (ASI).[vi]

$$city\ GDP_y = \sum_{\substack{s=1 \\ c=1}}^{n,m} \frac{state\ GDP_{s,y}}{state\ income_{s,y}} \times Income_{s,c,y}$$

s=economic sectors [vii], with a total of m economic sectors
c=districts or wards (that form the city), with a total of n such districts or wards
y=year
state=state where city is located
Income=wage and salary disbursements,supplements to wages and salaries,and proprietors' income

Inclusion of the unorganised sector

The ASI focuses on the organised sector. Data on income generated within the unorganised sector can be found in other surveys, such as the Survey on Unincorporated Non-Agricultural (Excluding Construction) Enterprises.[14] This survey contains data on emoluments but not operating surplus or profit, which would need to be included in the survey if data is to be used for computing GDP in the unorganised sector. This would also require state-level GDP data to be broken down into organised and unorganised sectoral GDP estimates.

This would improve the robustness of the estimates but would also increase the resource-intensity of the approach. Surveys on the unorganised sector would also need city-specific samples as the samples are typically designed for state-level estimates.

A potential extension of the formula to incorporate more information on the unorganised sector could be:

$$city\ GDP_y = \sum_{\substack{s=1 \\ c=1}}^{n,m} \left(\frac{state\ GDP_{s,y,\ organised}}{state\ Income_{s,y\ organised}} \times Income_{s,c,y,\ organised)}\right)$$

$$+ \sum_{\substack{s=1 \\ c=1}}^{n,m} \left(\frac{state\ GDP_{s,y,\ unorganised}}{state\ Income_{s,y\ unorganised}} \times Income_{s,c,y,\ unorganised)}\right)$$

s=economic sectors[viii]
c=districts or wards (that form the city)
y=year
state=state where city is located
Income=wage and salary disbursements,supplements to wages and salaries,and proprietors' income
where $state\ GDP_{(s,y,organised)} + state\ GDP_{(s,y,unorganised)} = state\ GDP_{(s,y)}$

vi Note that the income used here refers to personal income (that is, corporate income is excluded). We refer to the ASI (which contains industrial data on emoluments and operating surplus) because the quinquennial National Sample Survey Office (NSSO) surveys focus on household expenditure and not income. A review of public datasets suggests that there are no regular household or personal income surveys or censuses in India, so an alternative would be to use wage and surplus data within the ASI.

4.3 Top-down approach, using household expenditure data

The approaches in sections 4.1.2 and 4.2 focus on data from the economic sectors. An alternative approach is to use household expenditure data to refine the population approach discussed in section 4.1.1. A key limitation of the population approach in section 4.1.1 is the absence of information regarding greater wealth or greater economic activity in cities compared to rural areas. One possible approach therefore is:

$$city\ GDP_y = \frac{state\ GDP_y}{population_{state,y}} \times population_{city,y} \times \frac{household\ expenditure_{city,y}}{household\ expenditure_{state,y}}$$

$$where\ state\ GDP_y = \sum_{c=1}^{n} P_c(state\ GDP_y)$$

where p_c = population weight of area c, as proportion of state population, y = year

$$such\ that \sum_{c=1}^{n} p_c = 1$$ *and the state is formed from n areas*

The term $\frac{household\ expenditure_{city,y}}{household\ expenditure_{state,y}}$ acts as a weight for state GDP per capita to attempt to correct for the greater wealth in cities. While household expenditure is imperfectly correlated with GDP per capita, both indicators show economic wealth. A key assumption here is that household expenditure forms a significant and major proportion of expenditure at the city level. Wealthier urban households are likely to be related to higher-than-average GDP per capita (compared to overall state GDP per capita), making them less likely to underestimate city-level GDP per capita, compared to an approach using pure population weights. This ratio would then be multiplied by the GDP per capita of the state and the city-level population. However, while this approach captures information on urban wealth (relative to the state), it does not capture information on the economic structure of cities.

4.4 Bottom-up approach, based on income

The income approach calculates the sum of income generated by the domestic production of goods and services. At the city level, it must therefore be the income generated *within* the city, as opposed to within national borders. The income approach is one of the three different but equivalent ways of measuring GDP, as listed in the SNA 2008. (The other approaches are based on production and expenditure.)

This approach[15,16] sums up the following income streams:

[vii] This would typically follow the national industrial classifications for consistency.

Income	Indicator	Description
Workers	Wages and salaries	Includes all compensation, both monetary and non-monetary benefits, given to employees in the city
	Supplementary labour income	Contributions that employers make to social insurance plans (such as pension plans) on behalf of their employees
Capital	Mixed income	Production of goods and services by **unincorporated** businesses, such as self-employed people, including rental income
	Operating surplus *(Remainder of corporate revenue after worker compensation and net taxes on production and on imports have been paid from value added)*	Gross profits of **corporations and government business enterprises,** including investment income, dividends, interests and inventories
Government	Indirect taxes, minus subsidies	Taxes that governments collect from, and the subsidies that governments give to, businesses involved in producing goods and services

The data requirements here are extremely high, as this kind of data is typically collected through surveys and administrative records for both the public and private sectors. This information is collected at the national level; extending data collection down to the city level would require additional geographical tagging, as well as a wider survey base (increasing the cost).

4.5 Bottom-up approach, based on production (value-add based on geographic units or establishment data)

This approach is based on enterprise or establishment data, which is usually collected through a large-scale survey (or census) of establishments and businesses, carried out by the national statistical office. An example of such a survey is the Annual Survey of Industries (ASI) in India, which collects data on assets and liabilities, employment and labour cost, receipts, expenses, input items (indigenous and imported), products and by-products manufactured, and distributive expenses.[17] The 2016–17 version of the ASI collected data on 235,912 economic units.[18]

Using this approach, value-add per geographical unit (for example, ports or establishments) are allocated to the city. For instance, data on the fisheries sector could be gathered from specific ports. This is an extremely data-intensive exercise, typically carried out by the national statistical office.

4.6 Regression or correlation-based approaches, proxied using datasets related to economic activity (such as satellite data on lights at night)

As discussed in section 2.3, many of the papers covering this approach focus mostly on trends analysis—for example, to analyse correlation of activity such as electricity consumption or night luminosity with economic growth. Some of these papers—such as Bhandari and Roychowdhury, 2011,[19] and Chakravarty and Dehejia, 2017[20]—use this approach in the context of India, specifically focusing on night luminosity. However, both these papers do not estimate city or district-level GDP with the night

luminosity dataset. Instead, they use it to analyse trends such as economic convergence or growth in economic activity through regression or correlation analysis.

Research using night luminosity datasets generally relates GDP growth to light growth, analysing the closeness of the relationship through various techniques such as regression and neural networks. The relationship tends to be non-linear, as noted for India by Bhandari and Roychowdhury (2011) and for Turkey by Basihos (2016). Papers that do estimate provincial or city-level GDP, such as Basihos (2016), use night luminosity data to build out a provincial GDP dataset, extrapolating night luminosity growth from a base GDP year.[ix]

Estimating subnational GDP from this data can be problematic. Bickenbach et al. (2013) note that the relationship between the elasticity of observable GDP with regard to lights may not be stable[x] when looking at sub-national economic activity. As such, this approach may be better placed to test the point estimates generated by the approaches discussed above, as it can analyse trends such as changes in economic activity.

4.7 Conclusion: implementing top-down approaches to city-level GDP estimation

The sectoral income approach is a robust, top-down approach. Implementing this approach would also support a broader goal to improve and increase the datasets available for national statistics. A short-term (and less robust) alternative could be a top-down approach using household expenditure data, which is less resource-intensive than the sectoral income approach. Regardless of which approach is implemented, additional work will need to be completed before city-level GDP estimates can be calculated as sampling and data collection are currently focused on state-level estimates. Should the government wish to operationalise a city-level data collection exercise, stakeholder engagement with relevant government agencies and city governments would be required.

While think-tanks and other research agencies have developed estimates using different methods, it is important for the government to consider collecting detailed information on economic activity to support the long-term development of India's statistical framework. This information can feed into GDP calculations (both national and sub-national), as well as estimates of industrial production and city-level development. Government estimates of city GDP will depend on robust economic data collection, which in turn will strengthen the data available for policy-making.

[ix] The full paper outlines the precise methodology, which uses neural networks to estimate the non-linear relationship between GDP and night luminosity.

[x] Specifically, the stability of the elasticity of observable GDP with regard to lights.

5 Appendix: List of data requirements and sources

As part of the feasibility assessment, The EIU looked into available data sources on the Ministry of Statistics and Programme Implementation's website, Open Government Data (OGD) Platform India. We did not look at inputs for the regression or correlation-based approaches (for example, proxied using datasets related to economic activity, such as satellite data on lights at night) because these approaches tend to look at growth rather than estimating GDP directly.

Based on discussions with MoSPI officials, these sites should form the basis of our assessment of data availability. The ministry also noted that city-level GDP data was not available from the central office, although various city governments may estimate city-level GDP independently.

The assessment below summarises a review of data availability, based on publicly available information. It is possible that more detailed data can be collated from responsible ministries and the Central Statistics Office, or from local governments. For easy reference, data availability has been colour-coded as follows:

Data largely available	Data partially available, or with reservations on coverage	Data not available

5.1 Top-down approaches

As the data requirements for the top-down approaches "stack", we discuss data availability for them in the same table.

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data largely available	Top-down approach, using population or employment data Top-down approach, using household expenditure data	Population at the city level	Population and area	2011 (census), but the 2018 liveability assessment should provide updated figures	http://www.censusindia.gov.in/2011census/population_enumeration.html	● Available at the Central Statistics Office through the census ● Local governments and the Central Statistics Office may also have projections
Data largely available	Top-down approach, using population or employment data Top-down approach, using household expenditure data	Population at the state level	Area and Population – Statistical Year Book India 2017	2011 (census), although projections are available for 2016	http://www.mospi.gov.in/statistical-year-book-india/2017/171	● N/A

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data largely available	Top-down approach, using population or employment data Top-down approach, using overall employment as weights Top-down approach, using household expenditure data	State-level GDP	Gross state domestic product, both current prices and constant price	2016–17	https://data.gov.in/catalog/gross-state-domestic-product-current-prices https://data.gov.in/catalog/gross-state-domestic-product-constant-prices	● Available at the Central Statistics Office ● Published regularly
Data partially available, or with reservations on coverage	Top-down approach, using overall employment as weights	Employment and unemployment situation in cities and towns in India	Per 1,000 distribution of persons of age 15 years and above, by broad current weekly activity status for different size class of cities/ towns for each state/union territory (UT)	2011–12 (round 68)	http://www.mospi.gov.in/sites/default/files/publication_reports/report_564_20may15.pdf	● Data is available for cities and towns in all states and UTs, but the survey is conducted every five years and the information is somewhat outdated
Data largely available	Top-down approach, using overall employment as weights	Employment at the state level	Worker population ratio (per 1,000) for persons aged 15 years and above, according to the usual principal status (PS) approach for each state/UT	Mixed sources, ranging from 2011 to 2015–16	http://labourbureaunew.gov.in/UserContent/EUS_5th_1.pdf https://data.gov.in/resources/state-wise-employment-public-sector-and-private-sector-till-march-2011 https://data.gov.in/resources/state-wise-employment-organised-sector-2010-11-2011-12from-ministry-labour-and-employment https://data.gov.in/search/site?query=state+employment http://www.mospi.gov.in/statistical-year-book-india/2017/201	● The most frequently updated source seems to be the Annual Employment – Unemployment Survey (2015–16), which has the worker population per 1,000 in each state.[22] ● Specific data for ***formal*** state-level employment is available up to 2015–16, although fragmented data for specific sectors is available on the Open Government Data Platform. ● Caution is required because it has been noted that employment data has significant gaps, which weakens the case for using it as a ratio for the top-down approach. ● Informal employment data, as the Economic Survey 2015–16 noted, is only collected every five years.[23] ● However, this approach may be more feasible in the future, now that the government has started conducted a Quarterly Employment Survey (QES).[24] ● As the data is fragmented, the Central Statistics Office would need to conduct regular collation and reconciliation. Potentially, the QES could be expanded to cover more sectors of interest.

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data partially available, or with reservations on coverage	Top-down approach, using sectoral employment as weights	Sectoral employment in cities	N/A	N/A	http://labourbureaunew.gov.in/UserContent/EUS_5th_1.pdf https://data.gov.in/search/site?query=city+employment+sector	● The most frequently updated source seems to be the Annual Employment – Unemployment Survey (2015–16), which lists employment ratios by city types (categorised by population) using National Industrial Classification (NIC) 2008 codes (A–U) per 1,000 in each state.[25] ● However, data for specific cities was not found in the document. It is possible that detailed data may be available from the compiling agency. ● No other statistics were found in an online search for sectoral, city-level employment data. ● It is possible that local governments, or even city or state governments, may have data collected by their own government agencies. ● Some city-level authorities have previously published work with a sectoral breakdown of employment, such as Mumbai in 2003.[26]
Data largely available	Top-down approach, using sectoral employment as weights	Sectoral employment in states	Per 1,000 distribution of workers aged 15 years and above, by industry section based on NIC 2008, according to the usual Principal Status (PS) approach for each state/UT	2015–16	http://labourbureaunew.gov.in/UserContent/EUS_5th_1.pdf	● The most frequently updated source seems to be the Annual Employment – Unemployment Survey (2015–16), which has worker distribution by NIC 2008 codes (A – U) per 1,000 in each state. [27] ● The codes cover all codes in NIC 2008.

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data partially available, or with reservations on coverage	Top-down approach, using sectoral income data	State-level income data by sector, such as wage and salary disbursements, supplements to wages and salaries, and proprietors' income	Factory sector available through the Annual Survey of Industries (ASI), via industrial classification	2015	http://164.100.34.62/index.php/catalog/170/related_materials	● The ASI covers most of the income data required, as it includes data on rent, interest, wages and salaries, profits and employers' contribution for specific NIC 2008 codes. ● The government conducted a pilot in 2017 to deliver a similar survey on the services sector. However, it has reviewed this programme due to challenges collecting the required data.[28] ● The ASI only covers manufacturing (focusing on factories) and does not cover services (except for repair services and utilities) or agriculture. This was confirmed in the meeting with the MoSPI. ● By design, the ASI covers all factories employing 10 or more workers and using power, as well as those employing 20 or more persons but not using power. Units with 100 employees or more are covered by a census. ● A survey conducted every five years (Key Indicators of Unincorporated Non-Agricultural Enterprises, Excluding Construction) is available for the informal economy, although this excludes the construction sector and the agricultural sector. This survey captures emoluments paid by sector.
			Informal or unorganised economy through the Key Indicators of Unincorporated Non-Agricultural Enterprises (Excluding Construction) in India	2015–16	http://164.100.34.62/index.php/catalog/169/related_materials	
Data not available	Top-down approach, using sectoral income data	City-level income data by sector, mirroring the data requirements for states	N/A	N/A	N/A	● The ASI covers the income data required for manufacturing but is published at the state level and not the city level. ● It is possible that more detailed data is available, but it may require a significant extension of the sample size. This would require a significant investment in time and resources for the survey, as city-specific survey samples would be needed. ● Overall sector estimates could be derived from the city-level survey data, after accounting for non-response.

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data largely available	Top-down approach, using household expenditure	Household expenditure at the state level	Key indicators of household consumption expenditure in India	2011–12 (Round 68)		● State-level household consumption data is available (urban and rural breakdown available by state, for food, non-food and total)
Data partially available, or with reservations on coverage	Top-down approach, using household expenditure	Household expenditure at the city level	N/A	N/A	N/A	● In *Key Indicators of Household Consumption Expenditure in India*,[29] data is only available at an urban/rural stratum level for states, although the largest cities (towns with population of 10 lakhs or more, as per the 2001 population census) may form a separate urban stratum. ● As the survey for household consumption expenditure data is crafted to capture a state sample, using the same survey to obtain district or city-level data may not yield robust results. ● As a result, there may be a need to re-run the survey in specific cities with a good sample size to obtain robust results, which would be resource-intensive. ● Based on discussions with the MoSPI, states[xi] that have pooled their matching samples are most likely to have good granular data, and are better placed to help pilot a city GDP approach.

[xi] Based on discussion, the states with matching samples are Andhra Pradesh, Pradesh, Maharashtra, Telangana, Uttarakhand and Himachal Pradesh.

5.2 Bottom-up approaches

MoSPI officials noted that bottom-up approaches were extremely granular with respect to the data requirements and, as such, were unlikely to be feasible. Government statistics agencies in other countries (such as the US, the UK and other OECD countries) typically use a top-down approach. The exception is Canada, which is working on using the income approach to estimate metropolitan/city GDP. We discuss the data requirements for the income approach, as ultimately the approach will be a more granular (and extensive) version of existing national and state-level GDP data collection. The approach for geographic units or establishment data involves collecting value-add by geographic units—again, a more granular (and extensive) version of existing national and state-level GDP data collection.

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data not available	Bottom-up approach, based on income	Wages and salaries at the city level	N/A	N/A	N/A	● No statistics were found in an online search for city-level wage or payroll data. ● An International Labour Organisation (ILO) report[30] noted that the annual Employment and Unemployment Survey conducted by the Labour Bureau does not publish tables on wage rates in the reports. ● The quinquennial Employment and Unemployment Survey (EUS), published by the National Sample Survey Office (NSSO), collects data on components of wage and salary earnings received in cash and kind, but this data does not seem to be published. ● The data can potentially be disaggregated but this may entail additional investment in the survey.
Data not available	Bottom-up approach, based on income	Supplementary labour income at the city level	N/A	N/A	N/A	● No statistics were found in an online search for city-level wage or payroll data. ● An ILO report[31] noted that the annual Employment and Unemployment Survey (EUS) conducted by the Labour Bureau does not publish tables on wage rates in the reports. ● The quinquennial Employment and Unemployment Survey (EUS), published by the NSSO, collects data on components of wage and salary earnings received in cash and kind, but this data does not seem to be published. ● The data can potentially be disaggregated but this may entail additional investment in the survey.

	Approach	Data	Series found	Latest year found	Source link	Comments (if available)
Data not available	Bottom-up approach, based on income	Mixed income at the city level	N/A	N/A	N/A	● An ILO report[32] noted that the annual Employment and Unemployment Survey (EUS) conducted by the NSSO had a loophole in earnings data from the self-employed segment of the employed labour force.
Data not available	Bottom-up approach, based on income	Operating surplus at the city level (Remainder of corporate revenue after worker compensation and net taxes on production and imports have been paid from value added)	N/A	N/A	N/A	● The ASI covers the income data required for manufacturing (but not agriculture and most services), but this is published at the state level and not the city level. ● It is possible that more detailed data is available, but it may require a significant extension of the sample size. This would require a significant investment in time and resources for the survey.
Data partially available, or with reservations on coverage	Bottom-up approach, based on income	Indirect taxes minus subsidies, at the city level	Own tax revenue of states and UTs from 2007–08 to 2013–14	2013–14	https://data.gov.in/resources/own-tax-revenue-states-uts-2007-08-2013-14	● Based on discussions with MoSPI officials, city and state-level taxes will likely require detailed data provision from tax authorities and city-level governments.

Important notice

The report may be used by the Indian Ministry of Housing and Urban Affairs, only on an internal basis, solely for its own ordinary business purposes (which, for the avoidance of doubt, will exclude the distribution and/or disclosure of the report to any third party). The EIU shall not be held liable for the Indian Ministry of Housing and Urban Affairs or any third party's reliance on the report. While every effort has been made to verify the accuracy of this information, The EIU cannot accept any responsibility or liability for reliance by any person on this report or any information, opinions or conclusions set out herein.

References

1 OECD. (2016). *The OECD Metropolitan Areas Database visualised through the Metropolitan Xplorer: A user's guide*. Available at: https://measuringurban.oecd.org/content/The%20OECD%20Metropolitan%20eXplorer.pdf

2 Panek, S. D., Baumgardner, F. T., & McCormick, M. J. (2001). Introducing new measures of the metropolitan economy: Prototype GPD-by-metropolitan-area estimates for 2005. *Survey of Current Business, 87*(11), 79–87. Available at: https://www.bea.gov/scb/pdf/2007/11%20November/1107_gdpmetro.pdf

3 U.S. Bureau of Economic Activity. (2015). *GDP by metropolitan area methodology*. Available at: https://www.bea.gov/regional/pdf/GDPMetro2015.pdf

4 BEA. (2018). *2018 news release schedule: Regional economic accounts*. Available at: https://www.bea.gov/newsreleases/news_release_sort_regional.htm

5 Statistics Canada. (2017). *Measuring the economy, region by region*. Available at: https://www.statcan.gc.ca/eng/blog/cs/economy

6 Statistics New Zealand. (2007). *Regional GDP concepts, sources and methods*. Available at: https://unstats.un.org/unsd/economic_stat/China/Regional-gdp-concepts-sources-methods-New%20Zealand.pdf

7 Statistics New Zealand. (2007). *Regional GDP concepts, sources and methods*. Available at: https://unstats.un.org/unsd/economic_stat/China/Regional-gdp-concepts-sources-methods-New%20Zealand.pdf

8 Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review*, *102*(2), 994–1028. Available at: http://doi.org/10.1257/aer.102.2.994

9 Dai, Z., Hu, Y., & Zhao, G. (2017). The suitability of different nighttime light data for GDP estimation at different spatial scales and regional levels. *Sustainability*, *9*(2), 305.

10 Bhandari, L., & Roychowdhury, K. (2011). Night lights and economic activity in India: A study using DMSP-OLS night time images. *Proceedings of the Asia-Pacific Advanced Network*, *32*, 218. doi:10.7125/apan.32.24

11 Bickenbach, F., Bode, E., Nunnenkamp, P., & Söder, M. (2013). *Night lights and regional GDP*. Kiel Working Paper No. 1888. Available at: https://www.ifw-members.ifw-kiel.de/publications/night-lights-and-regional-gdp/KWP_1888.pdf

12 Chakravarty, P., & Dehejia, V. (2017). *India's income divergence: Governance or development model?* IDFC Institute Briefing Paper #5. Available at: http://www.idfcinstitute.org/site/assets/files/11015/briefing_paper_april_11_for_web.pdf

13 Basihos, S. (2016). *Nightlights as a development indicator: The estimation of gross provincial product (GPP) in Turkey*. Available at: https://mpra.ub.uni-muenchen.de/75553/1/MPRA_paper_75553.pdf

14 Ministry of Statistics and Programme Implementation. (2017). *Key Indicators of Unincorporated Non-Agricultural Enterprises (Excluding Construction) in India: NSS 37th report*. Available at: http://mospi.gov.in/sites/default/files/publication_reports/NSS_KI_73_2.34.pdf

15 Statistics Canada. (2017). *Measuring the economy, region by region*. Available at: https://www.statcan.gc.ca/eng/blog/cs/economy

16 Singapore Department of Statistics. (1998). *The income approach to gross domestic product*. Available at: https://www.singstat.gov.sg/-/media/files/publications/economy/gdi.pdf

17 Ministry of Statistics and Programme Implementation. (2016). *Annual Survey of Industries: Instructions to field officials*. Available at: http://www.csoisw.gov.in/cms/cms/Files/243.pdf

18 Ministry of Statistics and Programme Implementation. (2016). *Frame of Annual Survey of Industries (ASI)*. Available at: http://www.csoisw.gov.in/cms/cms/Files/690.pdf

19 Bhandari, L., & Roychowdhury, K. (2011). Night lights and economic activity in India: A study using DMSP-OLS night time images. *Proceedings of the Asia-Pacific Advanced Network*, *32*, 218. doi:10.7125/apan.32.24

20 Chakravarty, P., & Dehejia, V. (2017). *India's income divergence: Governance or development model?* IDFC Institute Briefing Paper #5. Available at: http://www.idfcinstitute.org/site/assets/files/11015/briefing_paper_april_11_for_web.pdf

21 Basihos, S. (2016). *Nightlights as a development indicator: The estimation of gross provincial product (GPP) in Turkey*. Available at: https://mpra.ub.uni-muenchen.de/75553/1/MPRA_paper_75553.pdf

22 Government of India. (2016). *Report on Fifth Annual Employment – Unemployment Survey (2015–16): Volume I*. Available at: http://labourbureaunew.gov.in/UserContent/EUS_5th_1.pdf

23 Government of India. (2017). *Economic survey 2015–16*. Available at: https://www.indiabudget.gov.in/budget2016-2017/survey.asp

24 Government of India. (2018). *Quarterly report on employment scenario*. Available at: http://labourbureaunew.gov.in/UserContent/QES_Report_Jan_2017.pdf?pr_id=isPYUxqHP1M%3D

25 Government of India. (2016). *Report on Fifth Annual Employment – Unemployment Survey (2015–16): Volume I*. Available at: http://labourbureaunew.gov.in/UserContent/EUS_5th_1.pdf

26 Government of India. (2018). *Economic profile.* Available at: http://www.mcgm.gov.in/irj/go/km/docs/documents/MCGM%20Department%20List/City%20Engineer/Deputy%20City%20Engineer%20(Planning%20and%20Design)/City%20Development%20Plan/Economic%20profile.pdf

27 Government of India. (2016). *Report on Fifth Annual Employment – Unemployment Survey (2015–16): Volume I*. Available at: http://labourbureaunew.gov.in/UserContent/EUS_5th_1.pdf

28 Suneja, K. (2017). Government reconsiders survey of the service industry. *The Economic Times*. Available at: https://economictimes.indiatimes.com/news/politics-and-nation/government-reconsiders-survey-of-the-service-industry/articleshow/58848053.cms

29 Government of India. (2014). *India – household consumer expenditure, NSS 68th round Sch1.0 type 2: July 2011 – June 2012, type – 2*. Available at: http://mail.mospi.gov.in/index.php/catalog/145/related_materials

30 Papola, T. S., & Kannan, K. P. (2017). *Towards an India wage report*. Available at: http://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new_delhi/documents/publication/wcms_597270.pdf

31 Papola, T. S., & Kannan, K. P. (2017). *Towards an India wage report*. Available at: http://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new_delhi/documents/publication/wcms_597270.pdf

32 Papola, T. S., & Kannan, K. P. (2017). *Towards an India wage report*. Available at: http://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new_delhi/documents/publication/wcms_597270.pdf

LONDON
20 Cabot Square
London
E14 4QW
United Kingdom
Tel: +44 (0) 20 7576 8181
Email: london@eiu.com

NEW YORK
750 Third Avenue
5th Floor
New York, NY 10017
United States
Tel: + 1 212 698 9717
Email: americas@eiu.com

HONG KONG
1301 Cityplaza Four
12 Taikoo Wan Road
Taikoo Shing
Hong Kong
Tel: + 852 2802 7288
Email: asia@eiu.com