UNU WIDER

15
25

Inequality, urbanization, and the Kuznets process

Evidence from India's annual Periodic Labour Force Surveys

S. Chandrasekhar,[1] Karthikeya Naraparaju,[2] and Ajay Sharma[3]

March 2025

WIDER WORKING PAPER

wider.unu.edu

United Nations University
World Institute for Development Economics Research

Abstract: We provide annual estimates of inequality in monthly per capita household earnings in India over the period 2017/18 to 2022/23 based on analysis of India's Periodic Labour Force Surveys. Over the six years, the estimate of inequality as measured by the Gini coefficient is in the range of 0.40 to 0.44 and, as measured by the Mean Log Deviation, between 0.28 and 0.34. We find that a 1 percentage point increase in the level of urbanization may increase the Mean Log Deviation by 0.5 to 0.7 per cent. Our analysis suggests that inequality will start declining only when India's urbanization rate is in the region of 63 to 74 per cent. Further, after accounting for variation in sectoral means and inequalities, we find that the development of the inequality–urbanization relationship at the sub-national level conforms to the Kuznets process.

Key words: household earnings, inequality, urbanization, Kuznets process, decomposition, India

JEL classification: D31, D63, I31

Acknowledgements: We are grateful for their helpful comments to Nancy Birdsall, Ravi Kanbur, Kunal Sen, Andy Sumner, and other participants at the symposium 'Kuznets at 70 – Inequality and Structural Transformation' held in Helsinki. Naraparaju acknowledges the travel support from UNU-WIDER to attend the symposium.

[1] Indira Gandhi Institute of Development Research, Mumbai, India; [2] Indian Institute of Management, Indore, India, corresponding author: **karthikeyan@iimidr.ac.in**; [3] Indian Institute of Management, Indore, India

This study is published within the UNU-WIDER project **Inequality and structural transformation – Kuznets at 70**, which is part of the **Reducing inequalities across and within countries** research area.

ISSN 1798-7237 ISBN 978-92-9267-572-1

https://doi.org/10.35188/UNU-WIDER/2025/572-1

1 Introduction

In the last three decades, income inequality has increased in countries accounting for 71 per cent of the world's population (United Nations 2020), leading to a search for factors that explain differences in inequality within nations (Huang et al. 2019). In this paper we focus on inequality in India, a country with a low level of urbanization, which in recent decades has been characterized by geographical variation in the reduction of poverty, an increase in consumption inequality, a steady increase in the ratio of average urban to rural consumption (Datt et al. 2020), and a lack of convergence of incomes at the sub-national level (Lamba and Subramanian 2020). Our work complements the large emerging cross-country literature on sub-national differences in wellbeing and inequality (Azam and Bhatt 2018; Huang et al. 2019; Kanbur et al. 2020; Wu and Rao 2017).

Given that India is now the world's most populous country and is among the most unequal countries in the world (Chancel et al. 2022), the evolution of India's income inequality will in turn affect the world's income distribution. Despite having grown to become the world's fifth-largest economy in terms of nominal gross domestic product, India has a per capita GDP that continues to lag significantly behind the world average.[1] India has also not undergone a structural transformation, and a large number of workers are engaged in the informal sector. In fact, India is classified as a structurally underdeveloped country, since nearly 50 per cent of its workforce is engaged in the agricultural sector (Sen 2023). Reflecting this, India's urbanization is also relatively low, at around 30 per cent. Furthermore, the urbanization process is driven not by an increase in manufacturing jobs but by an increase in service-sector jobs. Cross-country evidence would suggest that the movement of workers into services is likely to have a positive impact on inequality during the initial stages of structural transformation (Baymul and Sen 2020; Sen 2023). Since regional disparities continue to persist in India, Lamba and Subramanian (2020) suggest that India, unlike other developing countries and China, is an outlier. They posit this persistence as a puzzle. In search of an explanation they hypothesize that the divergence could be due to differences in the level of urbanization across Indian states. Recognizing that urbanization is an important correlate of economic development and that India's urbanization process is ongoing, in this paper we ask what this might imply for the country's income distribution. This leads us to ask two questions of interest. First, how much will inequality change as a result of an increase in the level of urbanization? Second, is there a turning point, i.e. at what level of urbanization will inequality start declining?

While we acknowledge that there are a multitude of factors that affect the evolution of income inequality, we focus on the salience of urbanization. The emphasis on the role of urbanization in the interaction between development and inequality has a long and venerable tradition in development economics, starting with the classic paper by Simon Kuznets in 1955. At the heart of various mechanisms that could underlie Kuznets' hypothesized inverted-U relationship between inequality and development is the impact of an increase in the share of urban population on total inequality.

According to the Kuznets process, the three major assumptions underlying the estimation of urbanization's impact on inequality are: first, that inequality is higher in urban areas than in rural

[1] In 2023, India's per-capita GDP (in nominal US dollars) was 2,485, which is around 19 per cent of the world's per-capita GDP (World Development Indicators).

areas; second, that mean income is higher in urban areas; and third, that the urbanization process is distribution-neutral within each sector (Ravallion and Chen 2022: 752).

The focus of this paper, which is on the impact of urbanization on inequality, is closely related to the work of Kanbur and Zhuang (2013) in the context of China, India, Indonesia, and Philippines and more recently that of Ravallion and Chen (2022) in the context of China.[2] Kanbur and Zhuang (2013) sought to estimate the contribution of the following components to changes in inequality observed at two points in time: level of urbanization, ratio of urban to rural income, and inequality within the rural and urban sectors. The aforementioned studies also sought to understand whether an inverted-U relationship exists and, if it does, whether there is a level of urbanization at which inequality will begin to decline.

In this paper we analyse six rounds of India's annual Periodic Labour Force Survey (PLFS) conducted by India's National Statistical Office. The PLFS, which was first conducted in 2017/18, is the first ever survey in India providing detailed information on the labour market earnings of all household members. In a notable departure from the existing literature on inequality in India, which has primarily focused on monthly per capita consumption expenditure (MPCE) or wage inequality, our focus is on monthly per capita household earnings (MPCHE).

Our key findings are as follows. Over the six years, the estimate of India's inequality in MPCHE as estimated by the Gini coefficient is in the range of 0.40 to 0.44.[3] The rural and urban Gini coefficients vary between 0.35 and 0.37 and between 0.43 and 0.45, respectively. And the difference between rural and urban inequality is consistent with the findings of Datt et al. (2020), who found that inequality in consumption expenditure had risen steadily over time and that the difference between rural and urban inequality was highest in 2010.

India's inequality, as measured by the Mean Log Deviation (MLD), varied between 0.28 and 0.34; when we decompose this by rural and urban, we find that, depending on the year, the between-group component accounts for between 14 and 19 per cent of the inequality.

We estimate that a one percentage point increase in the level of urbanization can increase India's inequality, as measured by the MLD, by between 0.5 and 0.7 per cent, depending on the year. We estimate that inequality will reduce when the share of the urban population is in the range 63–74 per cent across the six years. Our findings are robust across the six years of data.

Given that earlier work in the Indian context has looked at MPCE (e.g. Kanbur and Zhuang 2013), we complement our main results by also analysing India's latest Household Consumption Expenditure Survey (HCES) data—for the year 2022/23. Our estimates from the HCES data suggest that a 1 percentage point increase in urbanization can increase India's inequality in MPCE by 0.6 per cent.

[2] The empirical literature has sought to understand trends and patterns in inequality across countries within the Kuznets framework. Recent contributions include Paredes et al. (2016) and Akita and Miyata (2018) for Chile and Indonesia, respectively.

[3] This is lower than that of Brazil, China, and South Africa, three countries with which India is typically grouped.

Similarly, the HCES data predict that the turning point for inequality will occur when India's share of urbanization is at 63 per cent. Thus, our estimates from both the datasets are broadly consistent.

One hypothesis advanced in the World Development Report 2009 is that convergence in income across regions is at best a slow process (World Bank 2008). Since India's urbanization process is far from complete, the conjecture is that, in the future, richer and urbanized regions will benefit more than poorer regions, and this will exacerbate regional differences (Lamba and Subramanian 2020). This line of reasoning would suggest that India's inequality will be driven by differences across her 36 States and Union Territories. However, on decomposing the MLD by sub-national units, we find that the between-group component accounts only for between 14 and 16 per cent of inequality. This would suggest that India's inequality is largely a story of within-sub-national-unit inequality. When we decompose the inequality in a sub-national unit along the rural–urban axis, we find that the between-group component, which reflects the difference in the average income, i.e. between the average rural and urban MPCHE within a sub-national unit, varies between 2 and 23 per cent across the major units. As we will discuss later, sub-national units where the contribution of the between-group component is higher than the all-India estimates are a heterogeneous group—they vary by level of average income and level of urbanization.

Given the context of an extensive literature on econometric tests for the Kuznets hypothesis, for the sake of completeness we also undertake a regression analysis to understand how inequality at the sub-national level varies with level of urbanization. We follow the framework suggested by Anand and Kanbur (1993), who specified parametric restrictions for testing the relationship between inequality and development. Given that we are proxying level of urbanization for development, we first derive the functional form for the inequality–urbanization relationship according to the Kuznets process. After accounting for heterogeneity in sectoral means and inequalities, we find that the empirical relationship between inequality and urbanization across India's major sub-national units conforms with the Kuznets process.

The rest of the paper is structured as follows. In Section 2, we briefly describe the information collected as part of the PLFS and then present the key summary statistics. In Section 3, we discuss the estimates of inequality at the national and sub-national levels and the findings from the decomposition exercise. Here we also provide estimates of the impact of an increase in urbanization and urban-to-rural mean income differences on inequality. We then provide the functional form for the inequality–urbanization relationship under the Kuznets process and present the empirical results from the estimation of this relationship using sub-national-level data. In addition we provide results from the analysis of consumption expenditure data. Section 4 concludes.

2 Data and summary statistics

The annual PLFS launched by India's National Statistical Office in 2017/18 was the successor to the Employment and Unemployment Survey, which was last conducted in 2011/12. The PLFS is representative at the national and state levels as well as for rural and urban areas. Detailed information is collected on household characteristics, the demographic particulars of household members, and the activity particulars of household members including earnings and hours worked. Table 1 provides year-by-year details of the sample size of individuals and households that were enumerated as part of the survey. Details on sampling methodology and sample size of the first six rounds of PLFS are available in the reports published by Government of India (2019, 2020, 2021,

2022a, 2022b, 2023). Rural households were visited once, while urban households were visited four times. The analysis presented in this paper is based on the first visit for both rural and urban households.

Table 1: Number of observations in the Periodic Labour Force Survey (PLFS)

	Households			Individuals		
Year	**Rural**	**Urban**	**Total**	**Rural**	**Urban**	**Total**
2017/18	56,108	46,005	102,113	246,809	186,530	433,339
2018/19	55,812	45,767	101,579	239,817	180,940	420,757
2019/20	55,291	45,189	100,480	240,231	178,066	418,297
2020/21	55,389	44,955	100,344	236,279	174,539	410,818
2021/22	55,895	45,887	101,782	249,175	179,350	428,525
2022/23	55,844	45,811	101,655	243,971	175,541	419,512

Source: authors' calculations from unit-level data.

For every household, the PLFS provides a complete description of earnings from wages/salary, casual labour, and self-employment.[4] By definition, the earnings of those working as unpaid family workers equals zero. The reference period for the self-employed and regular wage-earners is 30 days preceding the survey, while for casual labourers it is the week preceding the survey. The monthly income from casual labour is arrived at by multiplying the weekly income by the factor (30/7). We calculate the total monthly earnings of the household by adding the income of working members (self-employment, regular wages/salary, and casual labour).[5] For each household we calculate the MPCHE, which is the ratio of total monthly earnings of the household to household size. As a welfare metric, MPCHE is preferred over total household earnings or average earnings per worker in a household. For the purpose of our analysis in the remainder of our paper, we consider MPCHE as the proxy for the income of the household and use these terms interchangeably. In the absence of appropriate spatial price deflators at the sub-national level, we opt not to make ad-hoc price adjustments. Also, it is a fairly standard practice not to adjust for spatial price differences in studies that examine patterns in inequality within countries (Shorrocks and Wan 2005).

In Table 2, we present the summary statistics pertaining to mean and median MPCHE in rural and urban India. We also report the ratio of average urban to rural MPCHE and the 75:25 percentile ratio. In line with the assumption required under the Kuznets process, we find that the mean income in urban areas is higher, by a factor of two, than that in rural areas. The difference in mean earnings between rural and urban households is important when we undertake the decomposition analysis.[6] Based on the estimates for more than 75 countries, a stylized fact is that the rural–urban gap in per capita consumption declines with higher levels of urbanization (World Bank 2008: 65). We find stark variations at the sub-national level in the ratio of urban to rural mean MPCHE (Figure 1). Kerala and Tamil Nadu, two southern states, are urbanized, with relatively high average MPCHE, and their rural–urban mean income differences are much lower than all-India. In contrast, in the equally urbanized states of Karnataka and Maharashtra, the ratio of urban to rural mean is high. This ratio is low in the

[4] The survey did not seek information on transfer incomes such as pensions, remittances, or interest income.

[5] We drop households that report zero earnings from all three sources (around 10 per cent of the total). Imputing a value of 0 to their earnings would be incorrect.

[6] Azam (2019) establishes that the urban–rural differences in consumption steadily increased over the period 1983–2011.

relatively less urbanized state of Bihar and the overall average income in Bihar is also low. The three mineral-rich states, Chhattisgarh, Jharkhand, and Odisha, stand out by exhibiting some of the highest sectoral income differentials across all states while being at the bottom of the earnings rankings.

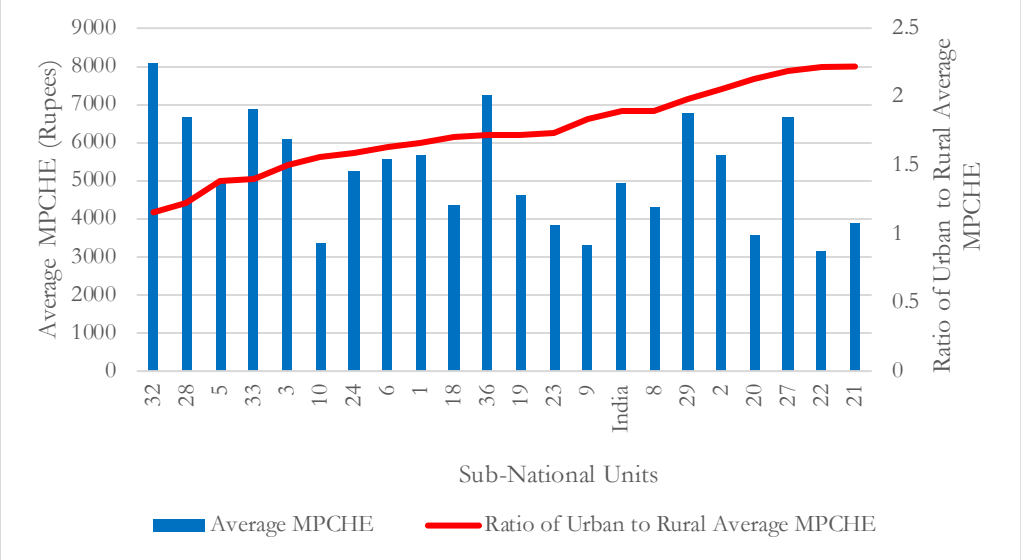
Table 2: All-India monthly per capita household earnings (rupees)

Year	Area	Mean	P-25	P-50	P-75	P-75:P-25 ratio	Ratio of urban to rural mean MPCHE	Urbanization rate (%)
2017/18	Rural	2,599	1,400	2,000	3,000	2.1	2.0	29
	Urban	5,156	2,250	3,500	6,000	2.7		
2018/19	Rural	2,782	1,500	2,250	3,333	2.2	2.1	30
	Urban	5,776	2,400	3,875	6,667	2.8		
2019/20	Rural	2,954	1,500	2,333	3,600	2.4	2.0	29
	Urban	6,039	2,500	4,000	7,143	2.9		
2020/21	Rural	3,229	1,667	2,500	4,000	2.4	1.9	28
	Urban	6,095	2,625	4,200	7,300	2.8		
2021/22	Rural	3,526	1,875	2,850	4,250	2.3	1.9	28
	Urban	6,871	3,000	4,833	8,000	2.7		
2022/23	Rural	4,005	2,200	3,333	5,000	2.3	1.9	26
	Urban	7,613	3,350	5,167	8,750	2.6		

Note: the person weight is calculated as a product of the household weight and household size. India's rural and urban population is arrived at by adding the respective person weights. It is to be noted that confidence intervals of household size can be calculated; hence we can calculate a confidence interval for the urbanization rate. These intervals overlap across the years (not reported here but available on request).

Source: authors' calculations from unit-level data.

Figure 1: Average MPCHE and ratio of urban to rural average MPCHE in 2022/23: major Indian sub-national units



Note: the sub-national units have been arranged from left to right on the x-axis in increasing order of their ratio of urban to rural average MPCHE. Sub-national codes are the following: 1: Jammu and Kashmir, 2: Himachal Pradesh, 3: Punjab, 5: Uttaranchal, 6: Haryana, 8: Rajasthan, 9: Uttar Pradesh, 10: Bihar, 18: Assam, 19: West Bengal, 20: Jharkhand, 21: Odisha, 22: Chhattisgarh, 23: Madhya Pradesh, 24: Gujarat, 27: Maharashtra, 28: Andhra Pradesh, 29: Karnataka, 32: Kerala, 33: Tamil Nadu, 36:Telangana.

Source: authors' calculations from unit-level data.

3 Estimates of inequality

3.1 Inequality measures

The commonly used measures in the literature on income or consumption inequality are the Gini and the Generalized Entropy (GE) class of measures. Within this class, GE(0), which is also known as Mean Log Deviation (MLD), has been shown to have favourable decomposition properties for the purpose of our analysis (see discussion in Anand and Kanbur 1993; Kanbur and Zhuang 2013; and Ravallion and Chen 2022). Unlike the Gini, the MLD is amenable to exact subgroup decomposition and, with the weights being the population shares of the sub-groups, this index has been referred to in the literature as belonging to the set of 'strictly decomposable' inequality indices (Anand and Kanbur 1993: 34).

The formula for GE(0) when there are two groups (g) as in our case, i.e. the two sectors rural and urban, is as follows:

$$GE(0) = \ln(\mu) - \sum_{g=1}^{2} \frac{n_g}{n} \sum_{i=1}^{n_g} \ln(y_{ig}) \quad (1)$$

where μ is the mean MPCHE of the population and Y_{ig} is the MPCHE of the ith household belonging to group 'g', n_g is the size of sector 'g', and n is the total number of rural and urban households.

This can be decomposed into between- and within-group components:

$$GE(0) = GE(0)_{Within} + GE(0)_{Between} \quad (2)$$

$$GE(0) = [\frac{n_R}{n} GE(0)_R + \frac{n_U}{n} GE(0)_U] + [\ln(\mu) - \frac{n_R}{n} ln(\mu_R) - \frac{n_U}{n} ln(\mu_U)] \quad (3)$$

where μ_g is the mean income of location g (rural (R), urban (U)). As is evident, the term in the first square brackets of Equation (3) is the weighted sum of inequality within rural and urban sectors and the second term is inequality due to differences in the mean income of rural and urban sectors. Decomposition of earnings inequality by rural and urban is of interest because it is a central piece in the framework of a Kuznets process.

In addition to its being 'strictly decomposable', there are at least four reasons for the popularity of GE(0). It is possible, first, to estimate how it would change if there were a 1 percentage point increase in the level of urbanization and, second, how it would change on account of a change in the ratio of average urban income to average rural income. Third, one can calculate how inequality would change if both the level of urbanization and ratio of mean urban and rural income changed. Finally, it is possible to estimate the threshold level of urbanization at which inequality will begin to decline.

Following Kanbur and Zhuang (2013), we write Equation (3) as follows:

$$\mathrm{GE}(0) = (1-\mathrm{x})GE(0)_R + \mathrm{x}GE(0)_U + \ln[\mathrm{xk} + (1-\mathrm{x})] - \mathrm{x}\ln(\mathrm{k}) \quad (4)$$

where x and 1-x are the share of urban and rural population, respectively (i.e., $\mathrm{x} = \frac{n_U}{n}$; $and\ 1 - \mathrm{x} = \frac{n_R}{n}$), and k is the ratio of average urban income to average rural income (i.e. $k = \frac{\mu_U}{\mu_R}$).

One can differentiate Equation (4) in order to estimate how inequality will change if the urbanization rate (x) changes by one unit:

$$\mathrm{dGE(0)/dx} = [GE(0)_U - GE(0)_R] + [(\mathrm{k}-1)/(\mathrm{x}(\mathrm{k}-1)+1)] - \ln(\mathrm{k}) \quad (5)$$

Setting dGE(0)/dx equal to zero gives the urbanization rate x^* where inequality will begin to decline.

$$\mathrm{x}^* = 1/[\ln(\mathrm{k}) - (GE(0)_U - GE(0)_R)] - 1/(\mathrm{k}-1) \quad (6)$$

Similarly, from (4) one can estimate the impact of a change in average urban to average rural income, k, leading to the following expression:

$$\mathrm{dGE(0)/dk} = \mathrm{x}/(1-\mathrm{x}+\mathrm{xk}) - \mathrm{x}/\mathrm{k} \quad (7)$$

In addition to GE(0), we provide estimates of inequality using the popular index of inequality, Gini coefficient (G), defined as the following:

$$Gini\ Coefficient = \frac{1}{(2n^2\mu)}\sum_{j=1}^{m}\sum_{k=1}^{m} n_j n_k |y_j - y_k| \quad (8)$$

where y_j, y_k are the MPCHE of households j and k, respectively; n_j is the number of households with MPCHE y_j; m denotes the number of distinct per-capita incomes; n is the total number of households; and μ is the mean of MPCHE across households. Unlike the GE(0) measure, Gini is not exactly decomposable into subgroups; hence we will not be using it for the decomposition analysis.

3.2 All-India estimates of inequality

Table 3 provides the estimates of inequality in MPCHE for the six years using the two inequality indices of Gini coefficient and GE(0), for all-India as well as for the rural and urban sectors separately. We estimate that for rural India the inequality in MPCHE as measured by Gini (GE (0)) is in the range 0.35 (0.21) to 0.37 (0.24) over the six years and for urban India it ranges between 0.43 (0.31) and 0.45 (0.36). Throughout the six years of data, our sectoral inequality estimates are consistent with the assumption made under the Kuznets process, that urban distribution is more unequal than its rural counterpart. Combining the two sectors, inequality at the all-India level as measured by Gini (GE(0)) ranges between 0.40 (0.28) and 0.44 (0.34). While there is a tradition of measuring consumption inequality, we argue that such estimates lead to the misleading conclusion that India is a low-inequality country. In Section 3.4 we provide estimates of inequality in consumption expenditure to reinforce the point about the significant difference between the two measures of inequality.

When we decompose the overall inequality as measured by MLD by rural and urban, we find that the 'between-sector' component, which captures the differences in average MPCHE between the rural and urban sectors, accounts for between 14 and 19 per cent of the total inequality across the six years of data (Table 3). To put this estimate in perspective, Shorrocks and Wan (2005), who review

the cross-country studies on decomposition of inequality by rural–urban, find the average share of the between-group component to be 19.6 per cent.

Table 3: Estimate of inequality in MPCHE – 2017/18 to 2022/23

Year	Gini R	Gini U	Gini	GE(0) R	GE(0) U	GE(0)	Sub-group decomposition of mean log deviation (GE(0)) by rural (R) and urban (U)			
							GE(0) Within	GE(0) Between	Share within GE(0)	Share between GE(0)
2017/18	0.37	0.44	0.43	0.24	0.33	0.32	0.26	0.05	83	17
2018/19	0.36	0.45	0.43	0.22	0.35	0.32	0.26	0.06	81	19
2019/20	0.37	0.45	0.44	0.24	0.36	0.34	0.28	0.06	83	17
2020/21	0.37	0.43	0.42	0.24	0.32	0.31	0.27	0.04	86	14
2021/22	0.37	0.43	0.42	0.24	0.31	0.31	0.26	0.05	84	16
2022/23	0.35	0.43	0.40	0.21	0.31	0.28	0.24	0.04	85	15

Source: authors' calculations from unit-level data.

When we decompose the inequality as measured by MLD by states, sub-national units, or the second tier of government in India, we find that between-state differences in average MPCHE account for between 14 and 16 per cent over the six years of data (Table 4). Thus, about 85 per cent of India's inequality in MPCHE is accounted for by inequality within states.

Table 4: Estimates of sub-group decomposition of mean log deviation (GE(0)) by India's states – 2017/18 to 2022/23

Year	GE(0)	GE(0) within	GE(0) between	Share within GE(0)	Share between GE(0)
2017/18	0.32	0.27	0.04	86	14
2018/19	0.32	0.27	0.05	84	16
2019/20	0.34	0.29	0.05	85	15
2020/21	0.31	0.27	0.05	85	15
2021/22	0.31	0.26	0.05	85	15
2022/23	0.28	0.24	0.05	84	16

Source: authors' calculations from unit-level data.

Plausible impact of increase in urbanization

In order to estimate the extent of the increase in inequality due to a 1 percentage point increase in the level of urbanization, we evaluate Equation (5), which is a partial derivative of inequality with respect to urbanization, keeping within-sector inequality and rural–urban mean differences constant. This is also known as the 'Kuznets Derivative' in the literature (Ravallion and Chen 2022). We find that a 1 percentage point increase in urbanization may increase India's inequality, as measured by the MLD, by between 0.5 and 0.7 per cent. Table 5 provides the estimates for the 'Kuznets Derivative' at the all-India level for the six years.

Table 5: Estimates of the partial derivative of GE(0) with respect to urbanization (Kuznets Derivative) and the turning points for inequality – 2017/18 to 2022/23

Year	GE(0)	Partial derivative of inequality with respect to urbanization (dGE(0)/dX)	% change in GE(0) for a 1 pp change in urbanization	Predicted turning points for inequality
2017/18	0.32	0.168	0.5	0.66
2018/19	0.32	0.214	0.7	0.74
2019/20	0.34	0.198	0.6	0.71
2020/21	0.31	0.153	0.5	0.67
2021/22	0.31	0.159	0.5	0.63
2022/23	0.28	0.188	0.7	0.73

Source: authors' calculations from unit-level data.

Is there a turning point?

While China was 50 per cent urban in 2011, as per World Urbanization Prospects 2018, India is expected to be 50 per cent urban only by 2045. While the Kuznets inverted-U hypothesis assumes there to be a turning point in the trajectory of inequality with an increase in urbanization, such a point might not in fact exist. As shown by Anand and Kanbur (1993) and Ravallion and Chen (2022), the turning point in inequality, as measured by the GE(0) index, exists if and only if the following condition is satisfied:

$$GE(0)_U - GE(0)_R < (1/\mathrm{k}) + \ln(\mathrm{k}) - 1 \quad (9)$$

We find that this condition is satisfied for India using the PLFS data, though the turning point varies slightly for each year: 66 per cent based on the 2017/18 data and, for the successive rounds of data, 74 per cent, 71 per cent, 67 per cent, 63 per cent, and 73 per cent (Table 5).

Inequality and urban–rural mean income differences

While the focus of our paper has been on the impact of urbanization on inequality, Kuznets (1955) also hypothesized that there would be an evolution of differences in per capita income between urban and rural sectors:

> The relative difference in per capita income between the rural and urban populations does not necessarily drift downward in the process of economic growth: indeed, there is some evidence to suggest that it is stable at best, and tends to widen because per capita productivity in urban pursuits increases more rapidly than in agriculture. If this is so, inequality in the total income distribution should increase (Kuznets 1955: 8).

In the Indian context, using data from household consumption expenditure surveys, Datt et al. (2020) document that the ratio of average urban to rural per capita consumption has been rising since 1970. Given this concern, using the expression in Equation (7) above, we quantify the predicted impact on inequality of a rise in the urban-to-rural mean MPCHE differences (defined earlier as 'k') and provide the estimates in Table 6. Given that the values of 'k' vary in smaller magnitudes, we assess the impact of a 0.01 unit change in 'k' on GE(0). Our estimates suggest that a 0.01 increase in 'k' increases GE(0) by 0.2–0.3 per cent across the six years.

Table 6: Estimates of the partial derivative of GE(0) with respect to the urban–rural mean MPCHE ratio ('k') – 2017/18 to 2022/23

Year	GE(0)	Partial derivative of inequality with respect to urban–rural mean income ratio ('k') (dGE(0)/dk)	% change in GE (0) for a 0.01 change in urban–rural mean income ratio ('k')
2017/18	0.32	0.08	0.3
2018/19	0.32	0.083	0.3
2019/20	0.34	0.081	0.2
2020/21	0.31	0.076	0.2
2021/22	0.31	0.077	0.3
2022/23	0.28	0.074	0.3

Source: authors' calculations from unit-level data.

3.3 Sub-national estimates

While we have estimated the impact of urbanization on inequality at the all-India level, as we pointed out earlier, about 85 per cent of inequality in India is due to differences within its states. In this context, it would be useful to establish a few facts about the nature of urbanization in India (based on Chandrasekhar and Sharma 2015 and Chandrasekhar et al. 2017). First, India is among the less urbanized of the top 10 countries in the world by GDP. Also, the distribution of the urban population is uneven both across and within states. As per the Census of India 2011, 42 per cent of the urban population live in cities with over 1 million population and 30 per cent live in towns with less than 100,000 population. The level of urbanization varies across Indian states from around 45 per cent in Kerala to about 10 per cent in Bihar. Second, the contribution of migration to urban population growth has been in the ballpark of 20 per cent.[7] Third, the bulk of the migration is intra-state in nature and not interstate. Fourth, the number of seasonal migrants (seasonal migration also being predominantly intra-state) is greater than the number of individuals who move permanently in any year. Also, the number of individuals commuting between rural and urban areas for work is greater than permanent migration in any year. Taken together, these aspects imply that urbanization in India is largely an intra-state phenomenon.

In Table 7, we provide the sub-national-unit-wise estimates of inequality using the Gini coefficient and GE(0) measures of inequality, along with sector-level estimates of inequality within the sub-national units for the year 2022/23.[8] In almost all the sub-national units, inequality is higher in urban areas. In the relatively urbanized southern states and the states of Maharashtra and Gujarat, there are significant differences in levels of inequality, varying from a Gini of 0.44 in Maharashtra to 0.34 in Kerala. The effect of resource endowment on regional inequality is an empirical question with little by way of guidance from theory (Lessmann and Seidel 2017). In the resource-rich states of Chhattisgarh, Jharkhand, and Odisha, the Gini is above 0.4. In addition to documenting the levels of inequality, we have undertaken the sectoral decomposition of state-level estimates of inequality (using GE(0)) and report the share of the between and within components of total inequality. In states where the ratio of

[7] With the natural increase in urban population and the reclassification of rural areas as urban accounting for the rest of the growth in the urban population.

[8] In the interests of space, we have not reported these numbers for the other five rounds of the survey, but the patterns are broadly similar, and the data are available on request.

urban to rural means is on the lower side,[9] the importance of the between-group component in explaining the state's inequality is low. In Kerala, the average MPCHE is higher than all-India, the ratio of average urban to rural MPCHE is lower (see Figure 1), and it is not surprising that the contribution of the between-group component is the smallest among all the major states, at 1 per cent in 2022/23. It is also not the case that all other states with high urbanization levels, like Kerala, have lower rural–urban inequalities. In urbanized states that also have a large manufacturing base, e.g. Karnataka and Maharashtra, the share of the between-sector component in 2022/23 is 23 per cent and 21 per cent, respectively, significantly higher than all-India.

Table 7: Estimates of inequality in monthly per capita household earnings in 2022/23

Sub-national unit code	Gini R	Gini U	Gini	GE(0) R	GE(0) U	GE(0)	GE WI share	GE BE share	Urbanization rate (%)
1	0.3382	0.4443	0.3725	0.2043	0.3295	0.2443	92	8	16
2	0.478	0.5109	0.4942	0.4171	0.4557	0.4461	94	6	9
3	0.3345	0.4516	0.3936	0.1879	0.3401	0.2611	93	7	35
5	0.3519	0.3388	0.3565	0.2149	0.1858	0.2179	95	5	26
6	0.3008	0.3999	0.363	0.1501	0.261	0.2164	87	13	34
8	0.3501	0.4374	0.4033	0.223	0.3191	0.2864	86	14	23
9	0.3645	0.4298	0.4007	0.2401	0.3083	0.2855	89	11	19
10	0.2755	0.3677	0.2913	0.1258	0.2172	0.1409	94	6	8
18	0.1767	0.3731	0.1915	0.0647	0.228	0.0751	93	7	3
19	0.2928	0.4287	0.3621	0.1478	0.3069	0.2245	86	14	28
20	0.3512	0.4166	0.4007	0.2278	0.2886	0.2849	84	16	17
21	0.3551	0.449	0.4033	0.2221	0.3384	0.2839	84	16	14
22	0.403	0.4413	0.4467	0.2812	0.3262	0.3454	84	16	18
23	0.3284	0.421	0.377	0.1848	0.2909	0.241	87	13	24
24	0.311	0.3534	0.3517	0.1652	0.2061	0.2092	87	13	42
27	0.326	0.4662	0.4439	0.1819	0.3625	0.3323	77	23	41
28	0.3279	0.3885	0.3504	0.18	0.2532	0.2059	98	2	29
29	0.3078	0.431	0.401	0.163	0.3071	0.27	79	21	35
32	0.3462	0.4029	0.375	0.215	0.3116	0.2614	99	1	45
33	0.2937	0.3606	0.3358	0.1488	0.2183	0.1914	93	7	41
36	0.3162	0.3472	0.3578	0.1714	0.1987	0.2167	84	16	37

Note: sub-national unit codes: 1: Jammu and Kashmir, 2: Himachal Pradesh, 3: Punjab, 5: Uttaranchal, 6: Haryana, 8: Rajasthan, 9: Uttar Pradesh, 10: Bihar, 18: Assam, 19: West Bengal, 20: Jharkhand, 21: Odisha, 22: Chhattisgarh, 23: Madhya Pradesh, 24: Gujarat, 27: Maharashtra, 28: Andhra Pradesh, 29: Karnataka, 32: Kerala, 33: Tamil Nadu, 36:Telangana.

Source: authors' calculations from unit-level data.

Recall that for all-India we expect a 1 percentage point increase in urbanization to increase the MLD by between 0.5 and 0.7 per cent. The impact of urbanization on inequality at the sub-national level varies[10] and this can be traced to components of the economic structure of the sub-national units such

[9] In the interests of space, we have not reported the state-wise ratio of urban to rural mean MPCHE here, but the data are available on request.

[10] We find large differences across India's 21 major states and the Union Territory of Jammu and Kashmir; the impact of urbanization varies between -0.02 per cent and 4.29 per cent over the six years.

as sectoral inequalities, rural–urban income differences, and rate of urbanization, all of which influence the estimation of the 'Kuznets Derivative' in Equation (5).[11]

Given the substantial interstate differences in urbanization and inequality levels, it is pertinent to ask whether these differences can be explained using the framework of the Kuznets process.

Anand and Kanbur (1993) provide the appropriate functional form to be used for empirical testing of whether the relationship between an inequality index and the level of per capita income conforms with the Kuznets process. We use this framework to arrive at the functional form for testing the relationship between GE(0) and the *urbanization rate*, x. We undertake this exercise for the following two cases: first, when the sectoral mean incomes and their ratio, k, are constant over time and, second, when these are assumed to vary with time. For both of these cases, we assume that the sectoral inequalities are constant over time. In the Appendix to this paper, we provide the derivations for these functional forms.

In the first case, the functional form of the relationship between GE(0) and x can be given by the following expression:

$$GE(0) = A + B\mathrm{x} + C\ ln\left(\mathrm{x} + \left(\frac{1}{k-1}\right)\right) \tag{10}$$

where

$$A = GE(0)_R + ln(k-1);$$

$$B = GE(0)_U - GE(0)_R - ln(k)$$

$$and\ C = 1$$

In the second case, where sectoral mean incomes are allowed to vary and grow differently over time, such that k is no longer constant, the functional form of the relationship between GE(0) and x can be given by the following expression:

$$GE(0) = A' + B'\mathrm{x} + C'\left(\mathrm{x} * ln\left(\frac{\mu_U}{\mu_R}\right)\right) + D'ln(\mathrm{x} * \mu) + E'ln(\mathrm{x} * \mu_R) \tag{11}$$

where

$$A' = GE(0)_R;$$

$$B' = GE(0)_U - GE(0)_R$$

$$C' = -1,\ D' = 1, and\ E' = -1$$

[11] In the interests of space, we have not reported the state-wise Kuznets Derivative here, but the data are available on request.

For the empirical testing of the functional forms derived for the first case, we estimate the following regression equations as specified in Equations (12) and (13) below, using data for India's 21 major sub-national units across six years, i.e. 126 observations in all. For the inequality–urbanization relationship to be consistent with the Kuznets process, under the assumption that the sectoral mean incomes and their ratio, k, are constant over time, we are required not to reject the hypothesis of parametric restriction of $\gamma = 1$ in Equations (12) and (13).

$$GE(0)_{it} = \alpha + \beta\, \mathrm{x}_{it} + \gamma\, ln\left(\mathrm{x}_{it} + \left(\frac{1}{k-1}\right)\right) + \sigma\, (YearDummy_t) + \varepsilon_{it} \quad (12)$$

$$GE(0)_{it} = \alpha_i + \beta\, \mathrm{x}_{it} + \gamma\, ln\left(\mathrm{x}_{it} + \left(\frac{1}{k-1}\right)\right) + \sigma\, (YearDummy_t) + \varepsilon_{it} \quad (13)$$

where $GE(0)_{it}$ is the inequality in sub-national unit 'i' in year 't' and x_{it} is the urbanization rate in sub-national unit 'i' in year 't'. Equation (12) is a pooled OLS estimation, whereas (13) represents a panel regression framework using sub-national-level fixed effects. The pooled OLS estimates test the hypothesis that all the sub-national units follow the same Kuznets process. The fixed effects regression, on the other hand, allows the Kuznets curve to have different intercepts for the sub-national units, while sharing the same slopes with respect to the level of urbanization. In terms of the functional form derived above, the fixed effect estimates assume that the *levels* of sectoral inequalities (captured in term A in Equation (10)) are different across the sub-national units (but constant over time) whereas their differences in sectoral inequalities (captured in term B in Equation (10)) are constant across states (and time).

The results for these estimations are provided in Table 8a. As can be seen from the 95% confidence interval for the estimate of γ, we can reject the hypothesis that γ=1, i.e. all states are not following the same Kuznets process, in either of the specifications.

Table 8a: Estimates of inequality–urbanization relationship with constant sectoral means

	(Model 1)	**(Model 2)**
	OLS	**Controlling for state-fixed effects**
Urbanization	0.107*	0.841***
	[-0.005,0.218]	[0.515,1.167]
	(0.056)	(0.164)
ln (Urbanization+(1/(K-1))	-0.075***	-0.078***
	[-0.101,-0.048]	[-0.107,-0.049]
	(0.013)	(0.015)
Time dummy	-0.001	0.002
	[-0.008,0.005]	[-0.001,0.005]
	(0.003)	(0.002)
Constant	0.287***	0.077
	[0.250,0.325]	[-0.016,0.170]
	(0.019)	(0.047)
Observations	126	126
R^2	0.220	0.326

Note: ***, **, * refers to statistical significance at the 1%, 5%, and 10% level of significance, respectively. 95% confidence intervals are reported in the squared brackets and standard errors are reported in the curved brackets. 'K' refers to the ratio of mean urban income to mean rural income in a state.

Source: authors' estimates from PLFS data.

Given this, we now relax the assumption that sectoral mean incomes are constant over time and use the functional form derived in Equation (11) above to assess the conformity to the Kuznets process under the second case. For this we estimate Equations (14) and (15) below (the OLS and fixed effects specifications, respectively) using data for the 21 major sub-national units across six years, i.e. 126 observations in all, under the null hypothesis of the following parametric restrictions: $\gamma'' = -1$, $\delta'' = 1$, and $\rho'' = -1$.

$$GE(0)_{it} = \alpha'' + \beta'' \mathrm{x}_{it} + \gamma''\left(\mathrm{x}_{it} * ln\left(\frac{\mu_{U,it}}{\mu_{R,it}}\right)\right) + \delta'' ln(\mathrm{x}_{it} * \mu_{it}) + \rho'' ln\left(\mathrm{x}_{it} * \mu_{R,it}\right) + \sigma''(YearDummy_t) + \varepsilon_{it} \quad (14)$$

$$GE(0)_{it} = \alpha''_i + \beta'' \mathrm{x}_{it} + \gamma''\left(\mathrm{x}_{it} * ln\left(\frac{\mu_{U,it}}{\mu_{R,it}}\right)\right) + \delta'' ln(\mathrm{x}_{it} * \mu_{it}) + \rho'' ln\left(\mathrm{x}_{it} * \mu_{R,it}\right) + \sigma''(YearDummy_t) + \varepsilon_{it} \quad (15)$$

where $GE(0)_{it}$ is the inequality in sub-national unit 'i' in year 't', x_{it} is the urbanization rate in state 'i' in year 't', k = $\frac{\mu_{U,it}}{\mu_{R,it}}$ is the ratio of urban to rural mean incomes in state 'i' in year 't', and μ_{it} is the mean MPCHE in state 'i' in year 't'. We present the results of these estimations in Table 8b.

Table 8b: Estimates of inequality–urbanization relationship with variable sectoral means

	(Model 1)	**(Model 2)**
	OLS	**Controlling for state-fixed effects**
Urbanization	-0.452***	0.263
	[-0.751,-0.153]	[-0.123,0.650]
	(0.151)	(0.195)
Urbanization*ln(K)	-4.873***	-1.493**
	[-6.150,-3.596]	[-2.923,-0.063]
	(0.645)	(0.721)
ln(Urbanization*Total Mean Income)	4.088***	1.560***
	[3.139,5.037]	[0.567,2.553]
	(0.479)	(0.501)
ln(Urbanization*Rural Mean Income)	-3.972***	-1.508***
	[-4.914,-3.031]	[-2.503,-0.514]
	(0.475)	(0.501)
Time dummy	-0.009***	-0.002
	[-0.016,-0.003]	[-0.006,0.002]
	(0.003)	(0.002)
Constant	-0.374***	-0.204*
	[-0.607,-0.142]	[-0.443,0.036]
	(0.117)	(0.121)
Observations	126	126
R^2	0.503	0.480

Note: ***, **, * refers to statistical significance at the 1%, 5%, and 10% level of significance, respectively. 95% confidence intervals are reported in the squared brackets and standard errors are reported in the curved brackets. 'K' refers to the ratio of mean urban income to mean rural income in a state.

Source: authors' estimates from PLFS data.

As can be observed from the 95% confidence intervals of the coefficients γ'', δ'', and ρ'', the OLS estimates reject the hypothesis of the required parametric restrictions (model 1, Table 8b). However,

when we allow the Kuznets curve to have different intercepts across the sub-national units in the fixed effects specification (model 2, Table 8b), from the confidence intervals of the coefficients γ'', δ'', and ρ'' it is evident that we cannot reject the hypothesis of the inequality–urbanization relationship being consistent with the Kuznets process. Thus, once we allow for sub-national-level differences in levels of inequality (but continuing to assume that they are constant over *time*), we observe that the relationship between urbanization and inequality across India's major sub-national units can be reconciled using the framework of the Kuznets process.

From this empirical exercise, we can also estimate the implied turning point. Using Equation (6), the turning point for the case where sectoral means and their ratio are allowed to vary is given by the following equation:

$$\mathrm{x}^{*} = \frac{1}{ln\left(\frac{\mu_U}{\mu_R}\right) - [GE(0)_U - GE(0)_R]} - \frac{\mu_R}{(\mu_U - \mu_R)} \tag{16}$$

Noting from model 2 of Table 8b that the estimate of β'', which captures $GE(0)_U - GE(0)_R$ (see Equation (11)), is statistically indifferent from zero and putting sample mean values for other terms in (16), we estimate the turning point to be around 38 per cent.

3.4 Analysis with household consumption expenditure survey data

In addition to the MPCHE, we analyse household consumption expenditure data using the latest round of the Survey on Household Consumption Expenditure (HCES) conducted between August 2022 and July 2023 by India's National Statistical Office, which evaluated 155,014 households in the rural sector and 106,732 households in the urban sector.[12] The rationale for undertaking an analysis of the consumption expenditure survey is the following: first, the literature on the Kuznets process in the Indian context has traditionally looked at consumption expenditure data (e.g. Kanbur and Zhuang 2013, who analyse consumption expenditure survey data from 1993 and 2008). Second, the literature on inequality in India has traditionally focused on consumption expenditure surveys and has documented, inter alia, the rising importance of urban-to-rural mean income differentials along with sectoral inequalities in driving India's overall inequality (e.g. Datt et al. 2020; Deaton and Dreze 2002). Given this, it is useful to compare our main findings using earnings data with the latest consumption expenditure survey and with earlier findings in the literature.

We consider monthly per capita household consumption expenditure (MPCE) as the indicator of individual welfare. Our main findings from the analysis of MPCE from the HCES are as follows: we estimate that for rural India the inequality in MPCE as per the Gini coefficient (GE(0)) is 0.27 (0.12), whereas the corresponding estimate for urban India is 0.32 (0.16). Thus, as is the case with MPCHE, consumption inequality is higher in urban areas than in rural areas. Combining the two sectors, India's inequality in consumption expenditure is 0.32 (0.16) as per the Gini coefficient (GE (0)). Consistent

[12] For details on the sampling procedure, items covered in the enumeration of household consumption expenditure, and other related aspects, refer to Government of India (2024).

with the literature, we also find that inequality in consumption expenditure is considerably lower than our estimates of earnings inequality.

The mean MPCE in rural areas is estimated to be Rs 3,773 and in urban areas it is Rs 6,459, leading to a ratio of urban to rural mean MPCE of 1.7. This is only marginally lower than the corresponding estimate from the earnings data for the last six years. Thus, while the levels of the sectoral means using the consumption expenditure data are lower than those from the earnings data, the relative gap between the sectoral averages is not significantly different between the two welfare metrics. Further, we note that Kanbur and Zhuang (2013) estimate the ratio of urban to rural mean MPCE in India to be 1.7 in 1993 and 2.0 in 2008.

On decomposing GE(0) by sector, we find that about 19 per cent of the inequality in MPCE is driven by the between-group component, i.e. through differences in average MPCE between the rural and urban sectors. Similarly, on decomposing GE(0) by states, we find that about 81 per cent of the inequality in household MPCE is due to inequality within the states and the remaining 19 per cent is due to mean MPCE differences between states.

In Table 9, we provide estimates of inequality at the sub-national level.

Table 9: Estimates of inequality in monthly per capita consumption expenditure in 2022/23

Sub-national unit code	Gini R	Gini U	Gini	GE(0) R	GE(0) U	GE(0)	GE WI share	GE BE share
1	0.2586	0.2451	0.2706	0.1055	0.1007	0.1165	90	10
2	0.2923	0.3146	0.3029	0.1373	0.1624	0.147	95	5
3	0.2235	0.2708	0.2475	0.0805	0.119	0.0993	95	5
5	0.2445	0.3047	0.2757	0.0956	0.1531	0.1227	88	12
6	0.2385	0.3358	0.3074	0.0949	0.1845	0.1557	82	18
8	0.2868	0.297	0.301	0.1389	0.1445	0.1506	93	7
9	0.2341	0.2981	0.2696	0.09	0.1444	0.1193	85	15
10	0.2158	0.281	0.2281	0.078	0.1274	0.087	94	6
18	0.209	0.2882	0.2425	0.0722	0.1338	0.0973	81	19
19	0.2306	0.3086	0.2836	0.0871	0.1551	0.1313	81	19
20	0.259	0.2986	0.2977	0.1087	0.1454	0.1436	80	20
21	0.2342	0.3357	0.2769	0.0893	0.1825	0.1254	82	18
22	0.2695	0.3157	0.3115	0.1189	0.1619	0.1584	80	20
23	0.2328	0.2939	0.275	0.0883	0.139	0.1235	82	18
24	0.2288	0.2843	0.2969	0.0855	0.1315	0.1418	73	27
27	0.295	0.3178	0.336	0.1429	0.1648	0.1839	83	17
28	0.2462	0.2868	0.2742	0.1007	0.1361	0.124	90	10
29	0.2279	0.3109	0.303	0.0841	0.1574	0.1477	75	25
32	0.2896	0.342	0.3186	0.1377	0.1897	0.1657	98	2
33	0.249	0.2835	0.2842	0.102	0.1303	0.1312	88	13
36	0.211	0.2821	0.2859	0.0728	0.1271	0.1317	74	27

Note: sub-national unit codes: 1: Jammu and Kashmir, 2: Himachal Pradesh, 3: Punjab, 5: Uttaranchal, 6: Haryana, 8: Rajasthan, 9: Uttar Pradesh, 10: Bihar, 18: Assam, 19: West Bengal, 20: Jharkhand, 21: Odisha, 22: Chhattisgarh, 23: Madhya Pradesh, 24: Gujarat, 27: Maharashtra, 28: Andhra Pradesh, 29: Karnataka, 32: Kerala, 33: Tamil Nadu, 36:Telangana.

Source: authors' calculations from unit-level data.

In addition to the sectoral and total level of inequality within each state (reported using both the Gini and the GE(0) measures), we provide estimates of the share of inequality within a sub-national entity that is contributed by inequality within rural/urban sectors vs. the share accounted for by inequality between the sectors. As is well known, the between-group component share is positively correlated with the differences in the average MPCE between the two sectors.[13]

With regard to the Kuznets Derivative, the impact of a 1 percentage point rise in urbanization on inequality in India's MPCE distribution, as measured by GE(0), is 0.6 per cent, which is similar to our estimate using the MPCHE welfare metric. For comparability with the earlier estimates, we note that our estimate from the 2022/23 consumption expenditure survey of the absolute change in GE(0) for a unit change in urbanization (i.e. dGE(0)/dx) is 0.10, slightly lower than Kanbur and Zhuang (2013)'s estimates of 0.15 and 0.18 for the years 1993 and 2008, respectively. In terms of percentage change, we use the inequality estimates provided in Kanbur and Zhuang (2013) to find that in 1993 and 2008, the impact of a 1 percentage point rise in urbanization on inequality in MPCE is at 0.8 per cent for both years. This is slightly higher than our estimates.

Using the consumption expenditure data, we estimate the turning point for inequality in the MPCE distribution to be at a level of urbanization of 63 per cent. This is at the lower end of our estimate of a range of 63 to 74 per cent using the earnings data. Moreover, our estimate of the turning point using consumption expenditure data is consistent with Kanbur and Zhuang (2013)'s estimate of 62 per cent for the year 2008. Thus, the predicted rise in inequality in consumption expenditure with increased urbanization is unlikely to be reversed until India reaches about twice its current level of urbanization.

4 Conclusion

Sustainable Development Goal (SDG) 10 seeks to reduce inequality within and among countries across all the monetary and non-monetary dimensions. In order to measure progress towards SDG 10, having estimates of inequality at national and sub-national levels as well as between rural and urban areas is of the utmost importance. In this paper, we undertake a detailed analysis of household earnings in India using data from the annual PLFS. Our findings are robust across the six rounds of data that we analyse. To the best of our knowledge, this is the first such analysis of household earnings. Thus, we are able to extend the literature on inequality in India, which has primarily focused on inequality in consumption or inequality in the earnings of regular wage workers.

A recurring theme in the literature on India's consumption inequality is the low rural living standards relative to urban areas and how this is an important factor constraining the ability of India's poor to benefit from the growth process. Our primary focus is on household earnings since these reflect livelihood opportunities and hence are a better indicator of standards of living. We highlight the importance of spatial aspects such as uneven urbanization, differential rural–urban incomes, and sectoral inequalities at the sub-national level in explaining the prevailing level of inequality in India. Based on the prevailing patterns of sectoral mean incomes and their distribution, we find that increasing urbanization in India will be accompanied by a significant increase in earnings and

[13] Not reported here but available on request.

consumption inequality. We also confirm that this increase in inequality does have a turning point, although it would come about only when India's urbanization rate doubles from the existing level.

A natural question that arises is whether there is any scope for policy intervention to reduce the marginal effect of urbanization on inequality or to hasten its turning point. Since the period of analysis is a relatively short one, we refrain from delving into the issue of what contributes to changes in inequality over time. Hence, we are unable to empirically pin down whether it is bridging the urban–rural income gap or reducing rural and urban inequalities that will matter more for reducing inequality. Nevertheless, our analysis does allow us to reflect on the underlying structural features of the Indian economy. As mentioned at the outset, India is classified as a structurally underdeveloped country. Like other developing countries it is characterized by a large proportion of workers being engaged in the informal sector, which also has low productivity. What Kuznets probably could not have envisaged 70 years ago was the nature of the evolution of developing country labour markets such as India, namely the growth of the urban informal sector and its persistence over time. This has implications for the evolution of both the ratio of average urban to rural incomes and inequality in urban areas, in turn influencing how inequality changes with urbanization as well as its turning point.

References

Akita Takahiro, S.M. (2018). 'Spatial Inequalities in Indonesia, 1996–2010: a Hierarchical Decomposition Analysis'. *Social Indicators Research* (2018)138: 829–52. https://doi.org/10.1007/s11205-017-1694-1

Anand, S., and R. Kanbur (1993). 'The Kuznets Process and the Inequality–Development Relationship'. *Journal of Development Economics*, 40(1): 25–52. https://doi.org/10.1016/0304-3878(93)90103-T

Azam, M. (2019). 'Accounting for Growing Urban–Rural Welfare Gaps in India'. *World Development*, 122(October): 410–32. https://doi.org/10.1016/j.worlddev.2019.06.004

Azam, M., and V. Bhatt (2018). 'Spatial Income Inequality in India, 1993–2011: a Decomposition Analysis'. *Social Indicators Research*, 138(2): 505–22. https://doi.org/10.1007/s11205-017-1683-4

Baymul, C., and K. Sen (2020). 'Was Kuznets Right? New Evidence on the Relationship between Structural Transformation and Inequality'. *The Journal of Development Studies*, 56(9): 1643–62. https://doi.org/10.1080/00220388.2019.1702161

Chancel, L., T. Piketty, E. Saez, G. Zucman, et al. (2022). 'World Inequality Report 2022'. World Inequality Lab. https://doi.org/10.4159/9780674276598

Chandrasekhar, S., and A. Sharma (2015). 'Urbanization and Spatial Patterns of Internal Migration in India'. *Spatial Demography*, 3: 63–89. https://doi.org/10.1007/s40980-015-0006-0

Chandrasekhar, S., M. Naik, and S.N. Roy (2017). 'On the Importance of Triangulating Data Sets to Examine Indians on the Move'. *Economic and Political Weekly*, 52(47): 60–68.

Datt, G., M. Ravallion, and R. Murgai (2020). 'Poverty and Growth in India over Six Decades'. *American Journal of Agricultural Economics*, 102(1): 4–27. https://doi.org/10.1093/ajae/aaz043

Deaton, A., and J. Drèze (2002). 'Poverty and Inequality in India: a Re-Examination'. *Economic and Political Weekly*, 7: 3729–48.

Government of India (2019). 'Annual Report – Periodic Labour Force Survey 2017–18'. National Sample Survey Office, New Delhi.

Government of India (2020). 'Annual Report – Periodic Labour Force Survey 2018–19'. National Sample Survey Office, New Delhi.

Government of India (2021). 'Annual Report – Periodic Labour Force Survey 2019–20'. National Sample Survey Office, New Delhi.

Government of India (2022a). 'Annual Report – Periodic Labour Force Survey 2021–22'. National Sample Survey Office, New Delhi.

Government of India (2022b). 'Annual Report – Periodic Labour Force Survey 2020–21'. National Sample Survey Office, New Delhi.

Government of India (2023). 'Annual Report – Periodic Labour Force Survey 2022–23'. National Sample Survey Office, New Delhi.

Government of India (2024). 'Survey on Household Consumption Expenditure August 2022–July 2023'. Report 591. National Sample Survey Office, New Delhi.

Huang, B., P.J. Morgan, and N. Yoshino (2019). 'Demystifying Rising Inequality in Asia'. Asian Development Bank Institute.

Kanbur, R., and J. Zhuang (2013). 'Urbanization and Inequality in Asia'. *Asian Development Review*, 30(1): 131–47. https://doi.org/10.1162/ADEV_a_00006

Kanbur, R., Y. Wang, and X. Zhang (2020). 'The Great Chinese Inequality Turnaround'. *Journal of Comparative Economics*, 49(2): 467–82. https://doi.org/10.1016/j.jce.2020.10.001

Kuznets, S. (1955). 'Economic Growth and Income Inequality'. *The American Economic Review*, XLV(March).

Lamba, R., and A. Subramanian (2020). 'Dynamism with Incommensurate Development: the Distinctive Indian Model'. *Journal of Economic Perspectives*, 34(1): 3–30. https://doi.org/10.1257/jep.34.1.3

Lessmann, C., and A. Seidel (2017). 'Regional Inequality, Convergence, and Its Determinants – a View from Outer Space'. *European Economic Review*, 92(February): 110–32. https://doi.org/10.1016/j.euroecorev.2016.11.009

Paredes, D., V. Iturra, and M. Lufin (2016). 'A Spatial Decomposition of Income Inequality in Chile'. *Regional Studies*, 50(5): 771–89. https://doi.org/10.1080/00343404.2014.933798

Ravallion, M., and S. Chen (2022). 'Is That Really a Kuznets Curve? Turning Points for Income Inequality in China'. *Journal of Economic Inequality*, 20: 749–76. https://doi.org/10.1007/s10888-022-09541-x

Sen, K. (2023). *Varieties of Structural Transformation: Patterns, Determinants, and Consequences*. Cambridge University Press. https://doi.org/10.1017/9781009449939

Shorrocks, A., and G. Wan (2005). 'Spatial Decomposition of Inequality'. *Journal of Economic Geography*, 5(1): 59–81. https://doi.org/10.1093/jnlecg/lbh054

United Nations (2020). 'World Social Report 2020: Inequality in a Rapidly Changing World'. Department of Economic and Social Affairs, United Nations.

World Bank (2008). 'World Development Report 2009: Reshaping Economic Geography'. Washington, DC: World Bank Group.

Wu, D., and P. Rao (2017). 'Urbanization and Income Inequality in China: an Empirical Investigation at Provincial Level'. *Social Indicators Research*, 131: 189–214. https://doi.org/10.1007/s11205-016-1229-1

Appendix

Deriving the functional form for the relationship between GE(0) and x.

Case 1: Assuming that the sectoral inequalities, the sectoral mean incomes and their ratio, k, are constant over time

$$GE(0) = x * GE(0)_U + (1-x) * GE(0)_R + ln\,[x * k + (1-x)] - x * ln(k)$$

$$= x * \left(GE(0)_U - GE(0)_R - ln(k)\right) + ln\left(x * k + (1-x)\right) + GE(0)_R$$

$$= GE(0)_R + x * \left(GE(0)_U - GE(0)_R - ln(k)\right) + ln[x * (k-1) + 1]$$

$$= GE(0)_R + x * \left(GE(0)_U - GE(0)_R - ln(k)\right) + ln\left(x * (k-1)\right) + ln\left[1 + \frac{1}{x * (k-1)}\right]$$

$$= GE(0)_R + ln(k-1) + x * \left(GE(0)_U - GE(0)_R - ln(k)\right) + ln(x) + ln\left[1 + \tfrac{1}{x*(k-1)}\right]$$

$$= GE(0)_R + ln(k-1) + x * \left(GE(0)_U - GE(0)_R - ln(k)\right) + ln\left(x + \frac{1}{k-1}\right)$$

Therefore:

$$GE(0) = A + Bx + C\ ln\left(x + \tfrac{1}{k-1}\right) \qquad \text{(A1)}$$

where $A = GE(0)_R + ln(k-1)$;

$$B = \left(GE(0)_U - GE(0)_R - ln(k)\right); and$$

$$C = 1$$

Case 2: Assuming that the sectoral inequalities are constant but the sectoral mean incomes and their ratio, k, vary over time

$$GE(0) = x * GE(0)_U + (1-x) * G.E.(0)_R + ln\,[x * k + (1-x)] - x * ln(k)$$

We note that $k = \frac{\mu_U}{\mu_R}$

Substituting the value of k in the above expression, we get:

$$GE(0) = x * GE(0)_U + (1-x) * GE(0)_R + ln\left[x * \frac{\mu_U}{\mu_R} + (1-x)\right] - x * ln\left(\frac{\mu_U}{\mu_R}\right)$$

$$= x * GE(0)_U + (1-x) * GE(0)_R + ln\left[\frac{x * \mu_U - x * \mu_R + \mu_R}{\mu_R}\right] - x * ln\left(\frac{\mu_U}{\mu_R}\right)$$

$$= x * GE(0)_U + (1-x) * GE(0)_R + ln\left[\frac{x * \mu_U + (1-x) * \mu_R}{\mu_R}\right] - x * ln\left(\frac{\mu_U}{\mu_R}\right)$$

$$= x * GE(0)_U + (1-x) * GE(0)_R + ln\left(\frac{\mu}{\mu_R}\right) - x * ln\left(\frac{\mu_U}{\mu_R}\right)$$

$$= x * \left[[GE(0)_U - GE(0)_R] - ln\left(\frac{\mu_U}{\mu_R}\right)\right] + GE(0)_R + ln\left(\frac{x * \mu}{x * \mu_R}\right)$$

$$= GE(0)_R + x * \left[[GE(0)_U - GE(0)_R] - ln\left(\frac{\mu_U}{\mu_R}\right)\right] + ln(x * \mu) - ln(x * \mu_R)$$

$$= GE(0)_R + x * [GE(0)_U - GE(0)_R] - x * ln\left(\frac{\mu_U}{\mu_R}\right) + ln(x * \mu) - ln(x * \mu_R)$$

Therefore:

$$GE(0) = A' + B'x + C'\left(x * ln\left(\frac{\mu_U}{\mu_R}\right)\right) + D'ln(x * \mu) + E'ln(x * \mu_R) \qquad \text{(A2)}$$

where $A' = GE(0)_R; B' = [GE(0)_U - GE(0)_R]; C' = -1; D' = 1; and\ E' = -1$