India Poverty & Equity Team

WORLD BANK GROUP

India: Trends in Poverty, 2011-12 to 2022-23

Methodology Note

May 2025

A reproducibility package for this report is available in
the Reproducible Research Repository at
https://reproducibility.worldbank.org/index.php/catalog/280

This work is a product of the staff of the World Bank with external contributions. A core World Bank India Poverty and Equity team worked under the guidance of Ximena del Carpio (Practice Manager, Poverty and Equity, South Asia Region). The team included Nancy Puranbhai Devpura (Research Analyst), Maria Fernanda Gonzalez Icaza (Economist), Diego Lestani (Extended Term Consultant), Julia Liniado (Consultant), Nishtha Kochhar (Economist), Nandini Krishnan (Lead Economist), Naresh Kumar (Economist), and Ayago Wambile (Senior Economist). Throughout, the team benefited from technical review and guidance from an advisory group comprising both external and World Bank members, including Laura Moreno Herrera (Senior Data Scientist, World Bank), Himanshu (Professor, Jawaharlal Nehru University), Peter Lanjouw (Professor, Vrije Universiteit Amsterdam), Juan Munoz (Sistemas Integrales), Sergio Olivieri (Senior Economist, World Bank), Jaime Fernandez Romero (Data Scientist, World Bank), and Nobuo Yoshida (Lead Economist, World Bank). The team benefited from technical consultations with the World Bank's Global Poverty Working Group (GPWG) and the Poverty Global Department's Advisory Group.

Contents

Tables

Figures

Boxes

Abbreviations

CAPI	Computer-assisted personal interviewing
CES	Consumer Expenditure Survey
CPHS	Consumer Pyramids Household Survey
CPI	Consumer Price Index
CPI-IW	Consumer Price Index – Industrial Workers
CPI-AL	Consumer Price Index – Agricultural Laborers
CSQ	Consumables and Services Questionnaire, HCES 2022-23
DGQ	Durable Goods Questionnaire, HCES 2022-23
FDQ	Food Questionnaire, HCES 2022-23
FSU	First Stage Unit
GOI	Government of India
HCES	Household Consumption Expenditure Survey (HCES)
HCQ	Questionnaire with household characteristics, HCES 2022-23
ICP	International Comparison Program
IFS	International Financial Statistics
LMIC	Lower-Middle Income Countries
MMRP	Modified Mixed Reference Period
MoH	Ministry of Health and Family Welfare
MoSPI	Ministry of Statistics and Programme Implementation
MPCE	Monthly Per Capita Expenditure
MRP	Mixed Reference Period
NCA	National Consumption Aggregate
NSSO	National Survey Sample Office
PAPI	Paper-and-Pencil Interviewing
PDS	Public Distribution System
PPP	Purchasing power parity
SAS	Situation Assessment Survey of Agricultural Households
SSS	Second Stage Strata
UFS	Urban Frame Survey
UMIC	Upper-Middle Income Countries
URP	Uniform Reference Period
UT	Union Territory
WDI	World Development Indicators

Executive Summary

- This technical note provides updated poverty trends for India using international poverty lines and documents the methodology for generating the corresponding welfare aggregates, poverty, and inequality estimates.
- The new series is based on household microdata from the 2022-23 Household Consumption Expenditure Survey (CES) as well as data from the 2011-12 CES, released by the National Sample Survey (NSS) Office. The series uses the Modified Mixed Reference Period (MMRP).
- Poverty estimates are calculated relative to the international poverty lines (IPLs) of $2.15 (extreme poverty line), $3.65 (poverty line for Lower-Middle Income Countries, LMIC), and $6.85 (poverty line for Upper-Middle Income Countries, UMIC), all expressed in 2017 purchasing power parity (PPP). A range of plausible estimates are also presented for an eventual transition to 2021 PPP and revised IPLs.
- The new estimates suggest a significant decline in poverty between 2011-12 and 2022-23 across all IPL thresholds. Extreme poverty rates fell from 16.2 percent to 2.3 percent, based on the $2.15 per person per day international poverty line (in 2017 PPP). Poverty calculated under the LMIC poverty line of $3.65 (2017 PPP) declined from 61.8 percent to 28.1 percent, between 2011-12 and 2022-23. With the adoption of 2021 PPPs, an extreme poverty line of $3.00 would result in a 5.3 percent poverty rate in 2022-23. A new LMIC line of $4.20 would imply a poverty rate of 23.9 percent.
- The updated poverty rates are lower than those from previous World Bank estimates published in the Poverty and Inequality Platform (PIP)—which were based on the uniform recall period, the imputation methodology from Roy and Van der Weide (2022) and data from the Consumer Pyramids Household Survey (CPHS).

Main results: Poverty rates declined significantly between 2011-12 and 2022-23, a downward revision of the trend compared to previous estimates.

- Poverty rates declined significantly between 2011-12 and 2022-23 across all IPL thresholds.
- Extreme poverty rates fell from 16.2 percent to 2.3 percent, based on the $2.15 per person per day international poverty line (in 2017 PPP) and the MMRP-based welfare aggregate. The previous estimate published in the PIP—based on imputations and URP data—was 12.92 percent for 2021.
- Poverty calculated under the LMIC poverty line of $3.65 (2017 PPP), considered more appropriate for India's level of development, shows a decline in poverty rates from 61.8 percent to 28.1 percent, between 2011-12 and 2022-23. These estimates replace a previous (URP and imputation-based) estimate published in PIP of 44.5 percent for 2021.
- The extreme poverty rate in rural areas fell from 18.4 percent to 2.8 percent, while urban poverty declined from 10.7 percent to 1.1 percent. In rural areas, the LMIC poverty rate fell from 69.0 percent to 32.5 percent, and from 43.5 percent to 17.2 percent in urban areas.

- The overall trends remain robust across various methodological choices, including alternative approaches to constructing inter-temporal deflators, sensitivity checks for potential outliers in consumption data, and different household size variables derived from the survey questionnaires.
- However, the poverty trends are sensitive to changes in the real value of the IPLs. This is evident in interim estimates based on plausible changes to the values of IPLs, that could accompany a shift to 2021 PPPs.
- Given India's inflation rate between 2017 and 2021, a revised extreme poverty line of \$3.00 would constitute a 15 percent higher threshold than \$2.15 expressed in 2021 prices and result in a 5.3 percent poverty rate in 2022-23. A new LMIC line of \$4.20 would imply a 5 percent lower threshold for poverty than \$3.65 adjusted in 2021 prices and yield a poverty rate of 23.9 percent.

The revised trends reflect a concerted effort to improve the measurement of household welfare and maximize comparability between 2011-12 and 2022-23.

- Overall, changes introduced in the questionnaire design, sampling, and data collection methodology in the 2022-23 CES are a positive step towards alignment with international best practices.
- Key changes in 2022-23 survey design and implementation (relative to the 2011-12 CES) include revisions to the sampling strata, a transition from paper-based interviews (PAPI) to computer-assisted interviews (CAPI), and changes in how consumption data is collected—such as modifications to recall periods, the number of visits to each household, and the number of items surveyed.
- However, these changes imply that the CES data from 2011-12 and 2022-23 are not strictly comparable.
- Although this note presents results based on a concerted effort to create a comparable trend, users should be aware of the caveats regarding direct comparability between the two surveys and the reliability of inequality estimates, which are discussed throughout the technical note.
- Additionally, this methodology note is accompanied by a replicability package that includes the necessary codes to reproduce the steps and results presented.

The revised trends reflect methodological improvements aligned with best practices to measure household welfare.

Transition from uniform recall period to modified mixed reference period

- In the past, international poverty estimates for India used the Uniform Reference Period (URP) to capture consumption data in the CES. The 2022-23 CES fully transitioned to the Modified Mixed Reference Period (MMRP), which aligns more closely with international best practices. The 2011-12 estimates have been recalculated using MMRP to ensure consistency across the series.

From expenditure to welfare aggregates

- In the past, international poverty estimates were calculated based on Monthly Per Capita Consumption Expenditure (MPCE), constructed by aggregating monthly consumption expenditures on food, non-food, and durables, and adjusting for household size. To align with international best practices and to transition towards a welfare aggregate (instead of an expenditure aggregate), the following adjustments were made to the 2011-12 and 2022-23 series:

(a) Welfare contribution of subsidized or free items. Free and subsidized items received via the Public Distribution System (PDS) and free school uniforms were re-valued to reflect better their contribution to household welfare rather than their subsidized price or zero price.
(b) Durable goods purchases. Durable purchase values are excluded due to data limitations that make accounting for their use value difficult. Including the purchase value of new durables alone would bias and distort the welfare distribution.
(c) Hospitalization, taxes, and a select set of non-food items are excluded. Hospitalization and taxes are excluded as they are not welfare-enhancing. A small and select set of non-food items (e.g., jewelry) are excluded because they primarily serve as traditional stores of value rather than consumption. These expenses are lumpy, infrequent, and do not reflect typical consumption.
(d) Housing. House rental values for renters are excluded due to challenges in consistently capturing the value of housing services for both renters and homeowners. Since robust estimates of housing services could not be obtained for all households, excluding rent payments (available for renters only) prevents bias against homeowners in welfare rankings.

From nominal to real welfare aggregates

- Spatial, intra-temporal, and inter-temporal price deflation have been introduced to express the welfare aggregate in real terms. In the past, the expenditure aggregate was not deflated to account for within-survey year or spatial price differences, and separate urban-rural 2017 PPP factors were used to proxy for spatial deflation. With the introduction of spatial and intra-temporal price deflators, a single national PPP factor is applied to the real welfare aggregate.

Methodological adjustments to welfare measurement—including the inclusion of subsidies and exclusion of certain expenditures—resulted in a net decrease in average real per capita welfare and a narrowing of welfare inequality, particularly benefiting lower-income rural households while reducing measured welfare among higher-income and urban populations.

- The inclusion of welfare benefits derived from PDS subsidies and school uniforms led to a 2.8 percent increase in the mean per capita real welfare aggregate in 2022-23. In turn, excluding certain expenditures—like durables, taxes, hospitalization, and rents—mechanically results in a downward revision. Together, all methodological adjustments decrease the average real per capita welfare aggregate by 6.6 percent in 2011-12 and 7.4 percent in 2022-23, relative to the baseline (real expenditures).
- In both rural and urban areas, the adjustments narrow the dispersion of the welfare distribution—as the considered subsidies/transfers represent more significant relative gains among lower-

income households. In contrast, higher-income households are more likely to report larger durables and rental expenses.

- The spatial and intra-temporal price adjustments further affect the welfare distribution. On average, rural households experience higher real welfare (relative to nominal), while urban households observe a reduction via these price effects. The extreme poverty rate in rural areas fell from 18.4 percent to 2.8 percent, while urban poverty declined from 10.7 percent to 1.1 percent. Moreover, in rural areas, the LMIC poverty headcount ratio fell from 69.0 percent to 32.5 percent, decreasing from 43.5 percent to 17.2 percent in urban areas. The urban-rural gap in LMIC poverty also narrowed (from 25 to 15 percentage points), although the reduction rate was lower compared to extreme poverty (from 7.7 to 1.7 percentage points).

Consumption-based inequality estimates likely to be downward-biased.

- The estimated Gini index and percentile-based inequality measures suggest a decline in inequality between 2011-12 and 2022-23 and higher inequality in urban compared to rural areas.
- However, these indicators likely underestimate the true extent of inequality due to three key factors: (a) the adjustments made when transitioning from expenditures to real welfare aggregate, which compressed the distribution; (b) concerns regarding the underrepresentation of top-income households in the CES sample; and (c) a low urban population share as derived from survey weights. As a result, these should be viewed as lower-bound estimates.
- Even if consumption were well measured across the distribution, inequality estimates can vary substantially when derived from consumption rather than income or wealth.

Implementing recommendations to strengthen data collection and address research gaps can further enhance the measurement of welfare and poverty in India.

- Reintroducing pilot studies and bridging surveys is recommended to assess the potential impact of design changes ex-ante. An Expert Group could be established to address unresolved issues and propose methods to maintain consistency over time.
- Updating the CES sampling frame to an up-to-date population census is critical to ensure that estimates accurately represent India's demographic composition and urbanization trends.
- Research is needed to better understand the challenges posed by attrition, shifting household composition, and changing respondents, particularly as aggregating data across multiple interviews was unprecedented prior to the 2022-23 design.
- Generating credible inequality measures will require leveraging expenditure-based indicators and alternative (complementing) incomes and wealth data. Achieving this will depend on improvements to the CES instruments and the broader NSSO survey system. Additionally, targeted research and implementation efforts are essential to enhance the accuracy of welfare measurement for top-income households and those at the lower end of the distribution.
- Collecting more detailed socio-economic data in the CES could help better identify the drivers of poverty reduction.

1. Introduction

Accurately assessing household welfare and its distribution is crucial for understanding and addressing economic disparities within a country. Poverty and inequality measures are key metrics for guiding policy design, resource allocation, and development programs. They also help monitor progress toward development goals and evaluate the impact of policies and initiatives.

India's Household Consumption Expenditure Surveys (CES) have been the basis for assessing economic well-being and estimating poverty in the country. Conducted by the National Sample Survey Office (NSSO) under the Ministry of Statistics and Program Implementation (MoSPI) for over 70 years, these surveys gather data on household consumption expenditures of goods and services. The CES data also contribute to national accounts by supporting the rebasing and revisions of Consumer Price Indices (CPI) and the estimation and rebasing of GDP.

The NSSO has enhanced survey design over the years to improve data quality, including the latest 2022-23 CES data, which was released after an 11-year gap.[1] While design and data collection changes hinder direct comparisons with 2011-12 data, bridging the long-standing evidence gap on India's poverty trends necessitates some comparability compromises. The trends in this methodology note reflect a concerted effort to generate comparable estimates but with these limitations and caveats kept in consideration.

This technical note outlines the methodology for measuring poverty in India using international poverty lines and official household survey data. It details the approach, highlights methodological improvements, and presents key findings. The analysis follows international best practices in welfare aggregation and poverty measurement based on the 2011-12 and 2022-23 CES microdata. Poverty is assessed using the World Bank's international poverty lines (IPLs) of $2.15, $3.65, and $6.85 per capita daily (expressed in 2017 purchasing power parity, PPP). Estimates are also presented using 2021 PPP for a range of plausible IPLs. Inequality is measured using household consumption-based Gini Indices, which may be biased downwards due to limitations in CES data and missing top-tails of the income distribution.

This technical note is structured as follows. Section 2 discusses the data collection and sampling approach adopted by MoSPI for the CES 2022-23 and compares it with the previous CES in 2011-12. It details the sample and survey design and describes the changes between the 2011-12 and 2022-23 rounds, as well as the implications for comparability. Section 3 provides an overview of poverty measurement in India, both national and international. Section 4 explains the step-by-step process in moving from a nominal expenditure aggregate to the construction of a real welfare aggregate. Section 5 presents poverty and inequality estimates for 2011-12 and 2022-23. It also presents a range of plausible estimates using the 2021 PPP. Section 6 summarizes the results of sensitivity

[1] The 2022-23 survey round was renamed the Household Consumption Expenditure Survey (HCES), whereas previous rounds, including the 2011-12 round, were titled Consumption Expenditure Survey (CES). For the sake of consistency, this report will refer to all rounds as CES.

checks on treating outliers and survey-based intertemporal deflators. Section 7 discusses potential areas for improving the measurement of poverty and inequality with the CES.

2. Background on poverty measurement in India

2.1 National poverty lines

Poverty measurement in India has evolved significantly over the years. Debates over poverty in India have largely revolved around defining the poverty lines, contested data, and the challenges of determining who qualifies as "poor". The Tendulkar Committee's 2011-12 estimates, accepted by the Planning Commission, set the poverty ratio at 21.9 percent based on MRP consumption expenditure and associated poverty lines. The committee established separate rural and urban poverty lines for each state, reflecting variations in prices and consumption patterns, resulting in all-India rural and all-India urban poverty lines of INR 800 and INR 1,000, respectively.

Critics of the Tendulkar poverty line argued that it understated the cost of living and overlooked significant deprived populations. In response, the Rangarajan Committee (2014) proposed a higher threshold based on daily minimum calorie requirements, estimating that 29.5 percent of Indians were poor in 2011-12 using MMRP consumption expenditure and a rural and urban poverty line of INR 972 and INR 1,407 respectively. However, these recommendations were never officially adopted. The 2011-12 CES MRP data remain the basis for India's last official national poverty estimates.

Until recently, discussions on poverty have relied primarily on alternate data sources in the absence of an official consumption expenditure survey. These data sources include the Private Final Consumption Expenditure from National Accounts, the Periodic Labor Force Survey (PLFS), and the Consumer Pyramids Household Survey (CPHS) [see Bhalla et al. (2020); Roy and Van der Weide (2022); Ghatak and Kumar (2024); Himanshu et al. (2024)].

2.2 International poverty lines

The World Bank defines different thresholds for international poverty lines, which are typically revised with the adoption of new Purchasing Power Parity (PPP) series. The main results presented in this note use 2017 PPPs with the associated international poverty lines: the extreme poverty line at $2.15 per person per day, the lower-middle-income countries (LMIC) poverty line at $3.65, and the upper-middle-income countries (UMIC) poverty line at $6.85 per person per day. For globally comparable estimates, household consumption expenditure is converted to PPP prices using the Household and Non-Profit Institutions Serving Households (NPISHs) Final Consumption Expenditure series (9100000).[2] Each household's real welfare aggregate is compared to these

[2] Available at https://databank.worldbank.org/source/icp-2017#.

poverty lines to determine poverty status.[3] The real welfare aggregate is calculated by adjusting the nominal welfare aggregate for temporal and spatial price variation, to capture differences in costs of living in different regions and over time.

The World Bank typically employs the national PPP (Purchasing Power Parity) conversion factor to translate the international poverty line into the national currency.[4] In the case of India, the official nominal expenditure aggregate in 2011 did not account for spatial disparities in the cost of living. These adjustments were incorporated while constructing national poverty lines, as the Tendulkar Committee established separate poverty lines for each state and sector (urban/rural). Therefore, price adjustments to consumption expenditures were not needed for national poverty estimates since they were already included in the calculation of poverty lines. However, international poverty lines are fixed in PPP base-year prices and applied nationally. Therefore, as an interim solution, it was decided to construct sectoral (rural-urban) deflators based on the 2017 PPP data to adjust for rural-urban price differences. This approach effectively resulted in different international poverty lines for rural and urban India in the 2017 PPP-based estimates, with the rural poverty line being lower than the urban one. The internationally comparable poverty rates were then derived as a weighted average of the rural and urban poverty rates.

For the revised India series using the 2011-12 and 2022-23 CES, adjusting the real welfare aggregate for intra-temporal and spatial price differences allows using a single national PPP. This ensures household expenditures are expressed in national average annualized prices, consistent with international best practices. First, the welfare aggregate is deflated to account for survey month-to-month price differences (intra-temporal). This adjustment is done for each item at the state-sector level using the rural-urban CPI series released by MoSPI. After the intra-temporal adjustment based on CPIs, the welfare aggregate is expressed in average yearly prices for each household, accounting for within-year price variations over the 12-month survey period. Second, spatial price deflators are constructed to account for differences in the cost of living across states and urban and rural areas. This spatially deflated welfare aggregate expresses household expenditures in average national prices of the survey year.

The Consumer Price Index (CPI) used to adjust for inflation across survey years (inter-temporal) has also been changed. The 2011-12 and 2022-23 real welfare aggregates need to be expressed in PPP base year (2017 or 2021 in the future) prices for international poverty calculations. The transition is made from using the CPIs from the outdated Agricultural Labourers (AL) and Industrial Workers (IW) CPI series (to proxy for rural and urban inflation) to the all-India General CPI, which is retrieved

---

[3] The construction of a welfare aggregate is a country-specific exercise, based on the available survey data. The World Bank follows methodological guidelines and adopts international best practices in the process, aligned with microeconomic theory and consistency across countries —such as the practices outlined in Deaton and Zaidi (2002) and Mancini and Vecchi (2022). Developing countries tend to use consumption-based welfare, while higher-income countries have adopted income-based aggregates (see Section 4 for more details).

[4] PPPs are also used to convert different currencies into a common, comparable unit and account for price differences across countries.

from the International Financial Statistics (IFS), to express the welfare aggregate in PPP year prices.[5] The absence of spatial adjustment in the past necessitated the use of urban and rural PPP factors for India, and hence the use of urban and rural price indices for inter-temporal deflation. This is no longer required with the incorporation of a spatially adjusted welfare aggregate. Thus, the uniform national PPP factor is applied to express the real welfare aggregate in 2017 PPP equivalent US$, with direct comparability to the IPLs.

3. Background on the CES and caveats on comparability

Comparing consumption expenditures data from 2022-23 to 2011-12 and previous CES surveys poses several challenges. Significant survey design and data collection changes between 2011-12 and 2022-23[6] may lead to substantially different consumption expenditure estimates, driven by survey effects rather than changes in welfare. While the methodology proposed throughout this note aims to yield the most comparable trend possible given the available CES data, these limitations warrant caution when interpreting results.

3.1 Coverage and timeline for data collection

The 2011 and 2022 CES are large-scale nationwide surveys implemented over a 12-month period, with some differences between the two rounds.

- CES 2022-23, conducted from August 2022 to July 2023, covered the entire Indian Union, excluding a few inaccessible villages in the Andaman and Nicobar Islands. Data was collected from 8,723 villages and 6,115 urban blocks, covering 261,746 households—155,014 in rural areas and 106,732 in urban areas.
- CES 2011-12 was conducted from July 2011 to June 2012 and covered the whole of the Indian Union except: (i) interior villages of Nagaland situated beyond five kilometers of the bus route, and (ii) villages in Andaman and Nicobar Islands that remained inaccessible throughout the year. Data was collected from 7,469 villages and 5,268 urban blocks, covering 203,313 households – 119,378 in rural and 83,935 in urban areas.

3.2 Transition from PAPI to CAPI for data collection in 2022-23

The 2022-23 CES used the Computer-Assisted Personal Interviewing (CAPI) method, replacing the traditional Paper-Assisted Personal Interviewing (PAPI) used in 2011-12 and previous CES rounds. The switch from PAPI to CAPI improves data quality, efficiency, and transparency and aligns with international best practices. However, this complicates direct comparisons with previous survey rounds. For instance, a study in Tanzania showed that CAPI yielded lower consumption

[5] The AL-IW CPI series released by the Labor Bureau rely on base year 1986-87 for AL CPIs and 2016 for IW CPIs.
[6] The 2022-23 and 2023-24 surveys are similar in design and the consumption expenditures modules. In contrast, the 2023-24 dropped questions from the demographic schedule (e.g. dwelling characteristics).

estimates than PAPI due to added data quality checks, leading to differences in poverty and inequality estimates (Beegle, de Weerdt, Friedman, and Gibson, 2012).

3.3 Sample design in 2022-23 and changes from the 2011-12 CES

The 2022-23 CES used a multistage stratified sampling design. Villages/urban blocks were the First-Stage Units (FSUs), and households were the Ultimate-Stage Units (USUs).

First-stage sampling

A state or Union Territory (UT) is the basic geographical unit for stratum formation in rural and urban sectors. The sampling frame for the urban sector comprises the blocks listed in the most recent Urban Frame Survey (UFS). The rural sector includes villages from the 2011 Census, updated to exclude those that have become urbanized and are now part of the latest UFS, with larger villages or UFS blocks sometimes divided into smaller sub-units based on population or household numbers. A special stratum comprises all the uninhabited villages, as per the 2011 census, at the all-India level.

Each state's rural and urban sectors are then further divided into strata. Two strata are formed in rural areas:

a) **Stratum 1:** Comprising villages within 5 Km from the district headquarters or a city/town with more than 5 lakh population. This stratum is not formed if there are less than 50 such villages in the state/UT.
b) **Stratum 2:** Rest of the villages. This substratum ensures that these villages are adequately represented in the sample.

In the urban sector, UFS blocks within each state are divided into strata based on two criteria, i.e., population as per Census 2011 and 'Affluence' status as indicated in the UFS frame.

In rural areas, the FSUs were further sub-stratified into three groups based on the village population. Samples were distributed across these three groups proportional to the population. If any group was allocated more than 20 samples, two or more sub-strata were formed within such group. Similarly, in urban areas, if the sample allocated to any urban stratum was more than 20, two or more sub-strata were formed within the stratum.

A total of 8,723 FSUs were surveyed in rural areas and 6,115 in urban areas, subject to a minimum allocation of 40 FSUs to each state/UT. The FSUs were selected within each stratum following Simple Random Sampling Without Replacement (SRSWOR). The samples were divided into 10 panels during the survey period of August 2022-July 2023, each consisting of three months (**Table 1**). An equal number of FSUs were allotted for each panel to ensure a uniform sample over the entire survey period.

Table 1. Data collection period of the CES 2022-23

Panel	Time period	Panel	Time period
1	August – October 2022	6	January – March 2023
2	September – November 2022	7	February – April 2023
3	October – December 2022	8	March-May 2023
4	November 2022 – January 2023	9	April – June 2023
5	December 2022 – February 2023	10	May – July 2023

Second Stage Sampling

Each FSU's households are divided into three second-stage strata (SSS). All households of a selected village/urban block were classified into three groups based on land possession in rural areas and car ownership in urban areas. Finally, a total of 18 households were sampled within an FSU, with proportional representation from the three groups.[7]

Key changes in sample design, survey coverage and respondent attrition

1. **The First Stage Unit (FSU) sampling methodology has been modified between the CES surveys of 2022-23 and 2011-12.** While both surveys used villages and urban blocks as FSUs, the 2022-23 survey introduced a change by creating two strata in rural areas. Stratum 1 was created to ensure that villages nearer to district headquarters and larger towns are included in the sample while Stratum 2 comprises of all the remaining villages. At the All-India level, a special stratum consists of all uninhabited villages as per Census 2011.
2. **The second stage stratum (SSS) definition has also changed between the two surveys.** In the 2022-23 survey, the criteria for selecting the Ultimate Sampling Units (USUs), i.e., households, were revised: the SSS in rural areas was based on land ownership, while in urban areas, it was based on car ownership. In contrast, the 2011-12 survey identified households based on durable goods possession, non-agricultural activity in rural areas, and Monthly Per Capita Expenditure (MPCE) in urban areas.
3. **Table A.1.1** in the annex outlines the definitional changes in SSS between the two surveys. Changes in household selection criteria, particularly in urban areas, may increase the likelihood of selecting wealthier households while reducing representation from lower-income groups, as the remaining category is merged into a single stratum. This shift could affect poverty estimates and complicate comparisons with the 2011-12 survey. For subgroup analysis, differing stratification methods may impact estimate precision, potentially leading to misleading conclusions about poverty trends. Additionally, as in many developing countries, wealthy households at the top of the distribution are likely underrepresented due to rising non-response rates.
4. **Missing Information on Attrition and Substitution:** In 2022-23, substitution of households was allowed only for the first visit, and there was no such provision for households missing after the initial survey. However, information on the characteristics of original households substituted in the first visit or lost due to attrition thereafter is not available. The possible

[7] NSS HCES Report 2022-23. Table 1.1: Number of FSUs and number of households surveyed.

non-randomness of substitution or attrition could impact the surveys' representativeness. Attrition was not an issue in 2011-12 and previous years, as the unique questionnaire was fully answered during a single interview.

3.4 Changes in the recall period between 2011-12 and 2022-23

In 2011-12, two different CES schedules were collected: Schedule 1 and Schedule 2. Both followed the same item structure but differed in recall periods. Schedule 1 gathered expenditure data using the Uniform Reference Period (URP) and Modified Reference Period (MRP) and was administered to half of the sample (101,662 households). Schedule 2, on the other hand, used the Modified Mixed Reference Period (MMRP), which was conducted on a separate sample of 101,651 households. The CES 2022-23 questionnaire used only the Modified Mixed Reference Period (MMRP), comparable to the Type 2 schedule in CES 2011-12.

In 2022-23, the NSSO shifted to collecting data exclusively using MMRP and will continue to use it in the future. Until March 2025, the 2011-12 estimates on the World Bank's Poverty and Inequality Platform (PIP) were based on the URP aggregate. The series is revised to an MMRP aggregate to enable a consistent and comparable analysis between 2011-12 and 2022-23. While the MMRP approach is broadly similar to the 2011-12 survey, there are differences in the recall periods and level of aggregation and disaggregation of items within groups. For example, expenditure on milk and milk products was collected based on a 30-day recall period in 2011-12 but has now been shifted to the 7-day recall period category (**Table 2**). Further, several items in 2011-12 were collapsed into a single item, while new items were introduced in 2022-23. Since most items can be matched between surveys, harmonizing recall periods is achievable, unlike the comparability challenges mentioned above, which are harder to reconcile.

Table 2. Classification of items in different recall periods

Item groups	Recall period		
	URP	MRP	MMRP
Education, medical (institutional), clothing, bedding, footwear, personal goods, transport equipment, sport goods, medical equipment, cooking and other household appliances, crockery and utensils, furniture and fixtures, goods for recreation, residential building and other durables, jewelry and ornaments, other taxes and cesses.	Last 30 days	Last 365 days	Last 365 days
Milk and milk products, vegetables, fruits, egg, fish & meat; edible oil, spices, beverages, served and packaged processed food, pan, tobacco and intoxicants.	Last 30 days	Last 30 days	Last 7 days
Cereals and cereal substitutes, pulses, salt and sugar, fuel and light, toilet articles, household consumables, conveyance, consumer services, entertainment, medical (non-institutional), rent.	Last 30 days	Last 30 days	Last 30 days

3.5 Changes in the questionnaires between 2011-12 and 2022-23

Multiple changes were introduced to the survey instrument in 2022-23, which have implications on comparability.

1. **The CES 2022-23 gathered expenditure data using three separate questionnaires:** the Food Questionnaire (FDQ), the Consumables and Services Questionnaire (CSQ), and the Durable Goods Questionnaire (DGQ). These were administered over three consecutive monthly visits to the selected households alongside the Household Characteristics Questionnaire (HCQ), which collected household characteristics and demographic information. In total, the survey covered 405 items across the three categories.
2. **Item list expansion:** Changes in consumption patterns and the expansion of the CES item list also present comparability challenges. The number of consumption items increased from 347 in 2011-12 to 405 in 2022-23, while obsolete items with declining consumption were removed. These updates align with best practices to reflect shifts in consumer behavior. As societies grow wealthier and preferences evolve, household consumption patterns change. However, expanding the item list systematically raises reported consumption expenditures (Beegle, de Weerdt, Friedman, and Gibson, 2012), which can bias comparisons. **Table A.2.1** provides a detailed description of changes in the item list.
3. **Introduction of Multiple Questionnaires and Household Revisits**: The 2022-23 CES introduced multiple questionnaires and household revisits to enhance data collection. Unlike the single questionnaire used previously, this approach improves measurement by enhancing recall accuracy and reducing interviewee fatigue. Moreover, this shift implies moving from a single-visit survey to a staggered three-month visit. Survey experiments (Jeong et al., 2023) suggest that this change may also systematically increase reported consumption.
4. **Multiple Respondents and differing household sizes:** As a consequence of the above change, in 2022-23, different household members could be survey respondents on different visits, unlike in 2011-12 when one person (usually the head of the household) answered the entire survey. This could affect consistency in responses. Changes in household composition over the three survey months could potentially lead to the same household reporting different household sizes in the three questionnaires. If the household size reported in CSQ and DGQ differs from FDQ, the per capita consumption expenditure calculated in these modules is adjusted based on the household size reported in FDQ. While MoSPI uses this approach to calculate per capita expenditure for each household, the rationale for choosing household size reported in the FDQ as the reference point is unclear.

3.6 Implications and threats to comparability

There is broad agreement in the literature that changes in survey methods substantially impact welfare measurements. Randomized field experiments in developing countries have found that different methods for measuring household consumption can significantly impact reported consumption and poverty rates (Beegle, de Weerdt, Friedman, and Gibson, 2012).[8] In particular, the length of the survey questionnaire and the details of the items presented to respondents have significant consequences for welfare measurement. The effects of changing survey design can be "nothing short of staggering," influencing fundamental development indicators and key stylized facts (Beegle, de Weerdt, and Gibson, 2020). In another randomized study, Jolliffe (2001) found that poverty estimates based on "short" questionnaires were nearly 50 percent higher than those derived from longer versions of the survey questionnaire.[9] These studies caution that changes to the survey questionnaire “result in spurious estimates of change in consumption and absolute poverty levels" (Jolliffe, 2001).

In India, a particularly prominent debate on the 55th round of the CES in 1999-2000, with questions based on weekly and monthly recall posted in the same schedule, raised uncertainty regarding which to employ for the consumption aggregate. A broad consensus emerged from this debate that consumption expenditure estimates are very sensitive to both the choice of recall period for eliciting information and the level of aggregation of items/item groups. Subsequently, consumption distributions based on URP, MRP, and MMRP were found to vary in estimated mean consumption and generate markedly different inequality measures. In a given year, different consumption definitions will thus generate markedly different poverty estimates for the same poverty line. Thus, the recommendation was that the choice of poverty line be specific to the distribution used. These considerations prompted the establishment of the Tendulkar Committee, which was charged with deriving new poverty lines in 2009. The resultant Tendulkar Committee lines underpin the poverty estimates officially endorsed by the Government of India in 2011/12. These mark the last officially endorsed estimates that are currently available. These considerations pose important caveats for trends using international poverty lines, which are constant and not sensitive to or tailored to changes in the underlying welfare aggregate.

The changes in the CES's sampling design are likely to bias consumption estimates upward and underestimate the level of inequality. The survey changes will likely affect both the level and the distribution of consumption expenditures and welfare measures. The revised sampling strategy of 2022-23 is likely to sample relatively better-off households in both rural and urban areas while potentially missing poorer households.[10] For example, the changes introduced to the selection criteria and definition of SSS affect the profile of rural households represented in the survey towards a higher probability of selecting peri-urban households.

On the other hand, the relative share of urban households implied in the CES survey weights likely under-represents the urbanization rate in India between 2011-12 and 2022-23. The implied

[8] The authors randomized eight different methods for measuring household consumption in Tanzania. Their findings revealed substantial variations in reported consumption and poverty levels across the different methods, despite the population groups having identical actual consumption levels.

[9] Jolliffe (2001) examines an experiment in El Salvador, where two different consumption questionnaires were randomly administered to sampled households.

[10] On some issues related to sampling and comparability with earlier surveys, see Anand (2024).

urban share of the population remains fixed at ~28 percent in both 2011-12 and 2022-23. Considering India's economic development path over the past decade, a constant urbanization share is highly unlikely. Meanwhile, projections from the Ministry of Health and the World Development Indicators (sourced from UN projections) estimate an urbanization rate of ~35 percent (**Figure A.1.1**). The under-representation of the urban population would bias the national-level welfare indicators towards rural areas in 2022-23. As rural households report a lower level of consumption expenditures than urban peers, this bias could, on the one hand, overestimate poverty and, on the other hand, underestimate the true extent of inequality.

In the 2022–23 CES, adding new items and the disaggregation of grouped categories to reflect evolving consumption patterns would tend to systematically increase reported household consumption expenditure. These newly added or split items accounted for an average of 3.7 percent of nominal expenditure, with 99.5 percent of households reporting consumption of at least one such item. In contrast, items present in 2011–12 but removed in 2022–23 accounted for ~0.07 percent of monthly per capita expenditure in 2011–12, with only 5.5 percent of households reporting spending on at least one of these items. **Table A.2.1** presents the share of households consuming each added item, along with the average share of spending per item in monthly per capita expenditure in 2022–23.

Ultimately, in the absence of a pilot or bridge survey, it is impossible to assess conclusively the impacts of changes in the survey on the comparability of the consumption aggregate and potential biases in poverty and inequality measures. Changes in survey and sampling design, along with changes in recall period, are not uncommon. However, such changes in India in the past were typically preceded by experimental rounds, pilot studies, or bridge surveys to examine the sensitivity of consumption measures to key adjustments.[11,12] Unfortunately, without similar ex-ante data or studies to evaluate the effects of survey and sample changes between 2011-12 and 2022-23, the magnitude and direction of any biases introduced remain unclear. Users are, therefore, advised to consider these limitations when comparing estimates across years.

[11] Four experimental rounds were carried out in the late 1990s (rounds 51st to 54th) to examine the implications of the recall periods. The NSSO published findings from this analysis in report no. 447. These rounds used independent but comparable samples, for 30-day and 7-day recall periods. Similarly, the NSSO report no. 475 published findings from a "pilot survey" on suitability of reference period for measuring household consumption.

[12] The bridge surveys would collect two parallel samples, one using the existing survey design and another with a new survey design. This procedure was followed in 2009-10 (66th round) and 2011-12 (68th round).

4. Methodology to construct a real welfare aggregate for India

4.1 Theoretical considerations on welfare measurement

Poverty measurement consists of two key components: (i) constructing a real (if feasible) welfare aggregate to rank the population based on monetary welfare and (ii) determining a poverty line, the minimum standard of living below which individuals are considered poor.

Monetary welfare is typically measured through income or consumption. While income is the standard metric in developed countries, consumption expenditure is more reliable in developing economies like India due to high-income volatility, seasonal employment, informal economic activity, self-consumption, and practical difficulties in collecting income data through household surveys. As a result, consumption better reflects living standards, especially in agrarian economies (Ravallion, 1998; Deaton & Zaidi, 2002). India has consistently used nationally representative household expenditure surveys to assess welfare over the last 50 years.[13]

Using consumption as a metric for individual welfare requires adjusting household consumption expenditure to account for welfare-improving non-monetary benefits and household composition. For example, expenditures on debt repayments or medical emergencies do not enhance welfare but reduce the disposable incomes for spending on welfare-enhancing consumption. On the other hand, access to public services like free healthcare, subsidized food, or education can reduce private spending, meaning lower expenditures may not always reflect lower welfare. By excluding certain welfare-reducing expenditures and incorporating non-monetary benefits, the welfare aggregate offers a more accurate and policy-relevant measure of poverty, ensuring that interventions effectively target those most in need. The household welfare aggregate is typically divided by household size or adult equivalent scales to measure individual or per capita welfare.

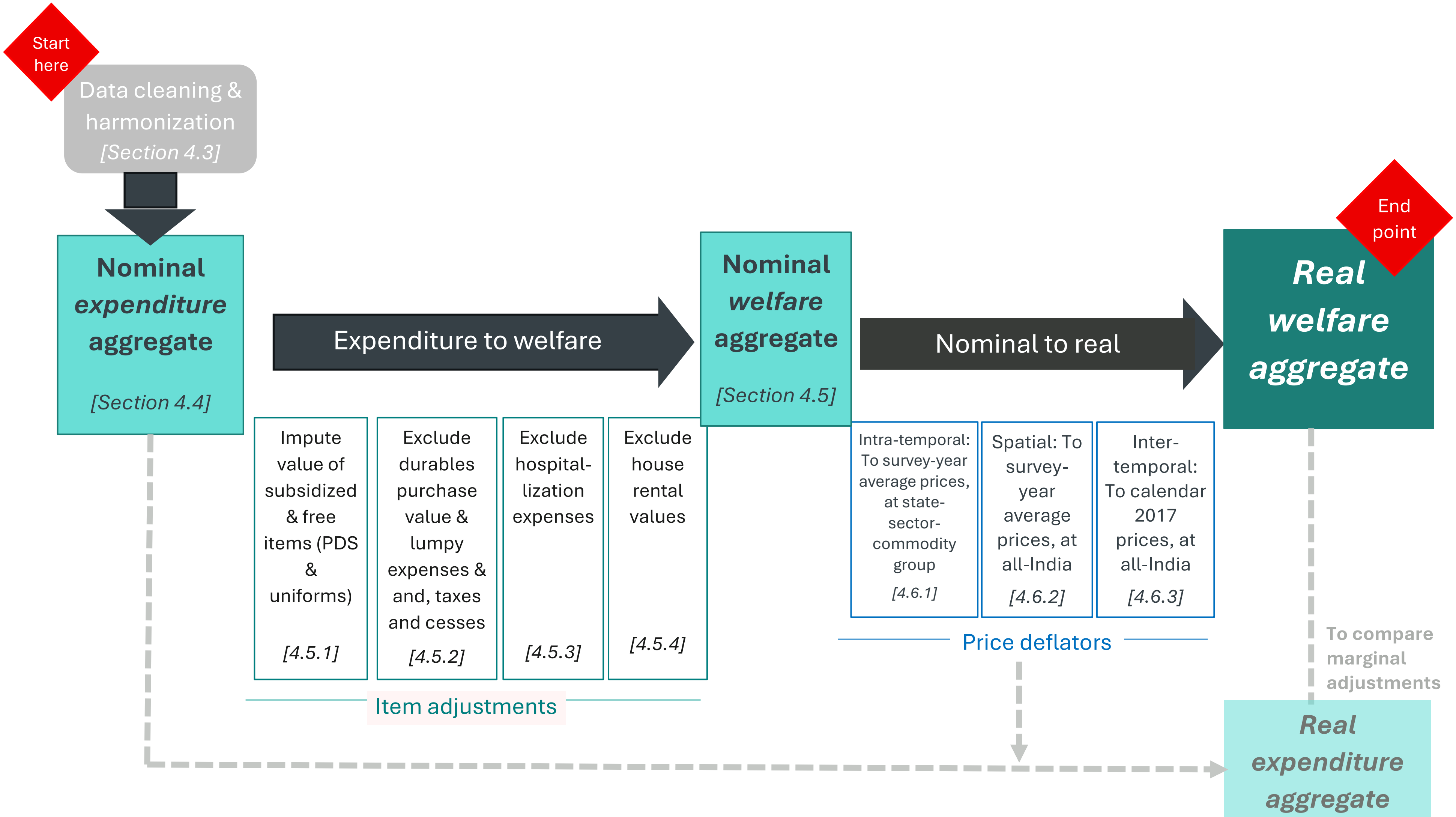
Lastly, for a consistent welfare ranking of households, the welfare aggregate is adjusted for cost-of-living differences across space and time (Deaton and Zaidi, 2002). When using a single cross-sectional household survey to construct measures of welfare, two sources of price variations must be addressed for comparability across households: (a) regional price differences across sampled households and (b) temporal price variations due to the different months in which households are surveyed during the reference period. The nominal aggregates are then deflated using these price indices to correct for spatial and temporal price variations, yielding real aggregates that allow for a consistent ranking of households across space and within the survey period.

[13] In practice, the total value of expenditures during the reference period towards household consumption of goods and services are used as a measure of household consumption used for poverty analysis. The household consumption expenditure is a proxy measure for money metric utility underlying the welfare measure in poverty measurement (Ravallion, 1998; Deaton and Zaidi, 2002).

4.2 Constructing a real welfare aggregate using 2011-12 and 2022-23 CES

This section discusses the methodological improvements and adjustments implemented to construct the most comparable trend in welfare aggregates between 2011 and 2022 based on the available CES data. Figure 1 summarizes the steps to construct a real welfare aggregate in India. First, a nominal expenditure aggregate is constructed using the total expenditures on consumption based on the MMRP schedule in the 2011-12 and 2022-23 CES (details in **Section 4.4**). Second, adjustments are made to transition from the expenditure aggregate into a distributionally consistent welfare aggregate, i.e., one that excludes welfare-reducing expenditures and, wherever possible, values welfare-improving expenditures appropriately for all households (details in **Section 4.5**). Lastly, spatial and temporal price deflators are applied to both the nominal expenditure aggregate and the nominal welfare aggregate to obtain real expenditure and welfare aggregates, respectively (discussed in **Section 4.6**).

Figure 1. Workflow: From nominal expenditures to real welfare aggregate



4.3 Data cleaning and harmonization[14]

The following preparatory adjustments are consistently applied to the 2011-12 and 2022-23 microdata to construct a comparable welfare aggregate.

1. **Alignment of state classification between 2011-12 and 2022-23:** Administrative boundaries in India changed after 2011, with the creation of new states such as Telangana and the creation or merger of union territories such as Daman & Diu and Dadra & Nagar Haveli.[15] Since the original 2011-12 data was collected prior to these changes, the state classifications were adjusted to ensure consistency and comparability with the state codes in 2022-23 CES.
2. **Creation of survey month variable:** The 2022-23 CES data does not include a survey month variable. For 2011-12, the survey month is directly derived from the interview date, while for 2022-23, it is created from the survey type, survey sequence, and panel information described in **Table 1**. This ensures each item has an assigned survey month, enabling intra-temporal adjustments within the survey year using monthly CPI.
3. **Mapping of items to CPI subgroups:** Each item in the two survey rounds is mapped to a CPI subgroup based on the detailed CPI composition and classification available from the MoSPI.[16] For smaller states and UTs where CPI subgroup information is not reported, items were mapped to the respective CPI group.
4. **Dropping aggregate subgroup expenditures:** The survey block-wise aggregated subgroup expenditures in raw data were dropped to avoid double counting.
5. **Adjustments of units:** To make units homogenous, all units are converted into kilograms and liters. While all quantities are in standard units reported in kilograms, grams, liters, and meters are divided by 1000 in 2011-12,[17] only quantities reported in grams are divided by 1000 in 2022-23.
6. **Dropping households with zero food expenditures:** In 2022-23, 50 households did not report any quantities or expenditures on food consumption. These households are entirely excluded from the analysis.[18]

4.4 Nominal expenditure aggregate

Household monthly expenditure for 2022-23 is derived by aggregating monthly consumption expenditure collected through three separate questionnaires: FDQ, CSQ, and DGQ. The household size recorded in the FDQ is considered a base for calculating the MPCE,[19] while weights from the last visit are used to generate the MPCE estimates. Following the methodology used by

---

[14] Despite these cleaning and harmonization efforts, users are advised to refer to **Section 2** to understand methodological and questionnaire changes between the 2011-12 and 2022-23 surveys.
[15] Exact details are available in the accompanying reproducibility package.
[16] Available at https://cpi.mospi.gov.in/Default1.aspx
[17] While the 2011-12 CES questionnaire has multiple units (kilograms, grams, liters or meters), the raw data input on quantities was recorded only in grams and liters.
[18] While these households report positive non-food spending, a part of the food questionnaire that includes quantities and expenditures is not answered by these 50 households.
[19] NSS Household Consumption Expenditure Report 2022-23. Chapter 2. Survey method & estimation of MPCE.

MoSPI to create nominal expenditures, the expenditure aggregate includes imputed value for home-grown consumption, gifts, free collection, and goods received in exchange for goods and services. The aggregate also includes reported actual rental expenditures for renters,[20] lumpy expenditures like hospitalization, durable goods purchases (including cars and motorcycles), and taxes and cesses.[21]

The calculation of the nominal expenditure aggregate involves three steps. First, all itemized expenditures are expressed in monthly terms, considering that expenditure data on different items are recorded using three different recall periods (7, 30, and 365 days). Second, monthly expenditures on all items are summed to obtain the total household monthly consumption expenditure. Third, the aggregate is expressed in PCE *(per capita expenditures)*. Since household sizes are separately recorded in each schedule, these may differ across the three questionnaires. The household size recorded in the FDQ is taken as the base, and expenditures in CSQ and DGQ are adjusted accordingly.

The total household expenditure (TE) is scaled for the base household size reported in FDQ module using the following formula:

$$TE = E_F + (E_C/P_C) * P_F + (E_D/P_D) * P_F$$

where:

TE = Total Household Expenditure
E_F = Household Expenditure on Food
E_C = Household Expenditure on Consumables and Services
E_D = Household Expenditure on Durable Goods
P_F = Number of household members reported in FDQ
P_C = Number of household members reported in CSQ
P_D = Number of household members reported in DGQ

The 2022-23 PCE is estimated TE divided by the P_F , the household size as recorded in the FDQ schedule.

In 2011-12, the nominal expenditure aggregate is derived in three steps. First, spending on all items[22] are expressed as monthly expenditures. Second, these monthly expenditures are aggregated for each household to obtain the total household monthly consumption expenditure. Third, this total is divided by household size to derive the 2011-12 MPCE.

[20] It excludes imputed rental expenditures (item 539 in 2022-23 questionnaire) reported by homeowners.
[21] Tax on consumer services include road cess, water charges, municipal rates, etc. which are levied on the household as a consumer unit. Professional and income tax are not included in the survey.
[22] Following MoSPI's guidelines, imputed rent (item 539 in 2011-12 questionnaire) is not included in the nominal expenditure aggregate.

4.5 From expenditures to welfare: Adjusting the nominal expenditure aggregate

As discussed in Section 4.1, the nominal expenditure aggregate is adjusted to include and exclude certain expenditures, as per the international best practices to measure welfare from household surveys. First**,** an imputed market-equivalent value of items received at a subsidized price or free of cost–food rations, kerosene, and school uniforms and shoes provided to students in government institutions—is included in the aggregate to reflect household welfare better. Second, certain expenditures are excluded from the welfare aggregate. Specifically, large expenditures on purchases of durable goods—such as cars, motorcycles, appliances, and furniture, which contribute to household welfare over multiple years—are not included due to data limitations that do not allow estimating the consumption use value of these assets. Purchases of large durables within the last 365 days are excluded so as not to distort the welfare rankings of households. Precious items like jewelry serve as a store of value and are also excluded. Moreover, while out-of-pocket health expenditures are included, hospitalization expenses are excluded from the survey as they reflect losses in welfare. Finally, although international standards recommend including the value of housing services for all households—both renters and owners of their dwellings—it is not possible to consistently and reliably estimate this component due to data limitations. Without a reliable and consistent imputation or prediction of housing services for house owners, rent expenses are excluded to avoid introducing biases or re-ranking households across the welfare distribution solely due to missing data. **Table 3** provides an overview of all the adjustments and the main rationale.

Table 3. Summary of adjustments to get the final real welfare aggregate

Revaluation of subsidized & free items	Exclusion of durables and lumpy expenditures	Exclusion of lumpy health expenses	Adjustment of housing (Exclusion of rents)
• Imputation of market equivalent value of PDS rations (free or partially subsidized) and free school items (uniforms, shoes). • Imputations were conducted based on median market prices at local levels with a hierarchical approach, ensuring a minimum of 50 observations.	• Durable purchases (e.g., automobiles) are excluded due to data limitations that hinder an accurate estimate of the consumption flow of services from using durable assets over the reference period. Purchase values within the last 365 days are dropped as they would bias welfare rankings. • Exclusion of taxes and cesses, like road and house taxes, that are not welfare-enhancing and may be lumpy. • Valued assets, such as jewelry and wristwatches, which are a store of value, are also excluded.	• Hospitalization expenditures are excluded because of their lumpiness, infrequency, and occurrence in response to potential welfare-reducing health shocks.	• Robust estimates of housing services were not achievable due to data limitations (i.e., demographic and household characteristics and actual rents, which mostly cover urban areas), especially in 2011-12. • Hence, actual rents paid are excluded from welfare aggregates to prevent bias and re-ranking of households due to missing data. • This potential contributes to underestimating inequality.

4.5.1 Re-valuation of subsidized items

(a) Free and subsidized food

India's Public Distribution System (PDS) provides food rations for purchases at subsidized prices or free of cost to consumers.[23] Food subsidies via the PDS constitute the largest fiscal outlay in the government's large set of social schemes.[24] With 84 percent of India's population holding a ration card,[25] the widespread consumption of subsidized rations via the PDS has significant impacts on the measurement of household welfare.

The monetary amount reported in the CES modules—rupees spent on PDS items purchased at either government-fixed subsidized prices or free of cost—does not necessarily reflect the true welfare benefits from consuming those items. Hence, the main challenge is allocating a value to

[23] The PDS-related items collected as part of the 2011-12 and 2022-23 CES included food items and kerosene. The methodology described in this section was equally applied to food rations and kerosene. The term "PDS items" is used to describe any item purchased via the PDS, whether it was consumed at a partially subsidized price or free of cost to consumers (full subsidy).

[24] In 2022-23, the Union government spent 6.5% of its budget on the PM Pradhan Mantri Garib Kalyan Anna Yojana (PMGKAY) to provide free and subsidized food rations.

[25] In 2022-23, 84% of Indians had access to a ration card, including 59% holding a Below Poverty Line (BPL), Antyodaya Anna Yojana (AAY), or Priority Household (PHH) card in their household. In practice, 74% actively consume food rations (or kerosene) via the PDS/PMGKAY, with rice and wheat as the most common. Coverage of ration cards is higher in rural areas (at 89% of the population) compared to urban areas (72%).

those PDS items to accurately reflect their marginal utility of consumption. Although this challenge is common for countries with public provision of rations, the answer is not straightforward and depends on the context of the specific country and the availability of survey microdata.

Mancini and Vecchi (2022) identify several potential approaches to impute a value to subsidized food rations. These include leveraging (a) information from a secondary market for rations, (b) finding a value based on close substitutes to those items, (c) using a self-reported assessment from consumers, and (d) using expert judgment as a last resort. In the case of India, accurate information on secondary markets of food rations or a self-assessment of consumers is unavailable. Option (b) (valuation based on close substitutes) was adopted as the preferred and feasible option.

The methodology re-values the consumption of items procured from PDS ration shops based on the unit value (UV)[26] of close substitutes in the local market. The CES questionnaires record separately the purchase of PDS subsidized and free items (e.g., the same household can report purchases of free rice and purchases of subsidized rice), as well as purchases of similar items from the market without any subsidy (i.e., rice from "other sources"). This disaggregation of the "source" of purchase allows us to leverage information from local market purchases to approximate the value of subsidized and free items had they not been received via the PDS. The list of matching items is presented in **Table 4.**

Table 4. Correspondence of items purchased from the PDS and market

	Source		
Product	Free of Cost	PDS at Subsidized Price	Market purchases
1	rice-free (061)	rice-PDS (101)	rice-other source (102)
2	wheat/atta-free (062)	wheat/atta-PDS (107)	wheat/atta-other source (108)
3*	coarse grains-free (070)	coarse grains-PDS (01)	coarse grains-other sources (02)
3.A*	jowar-free (063)	jowar-PDS (055)	jowar-other source (115)
3.B*	bajra-free (064)	bajra-PDS (056)	bajra-other source (116)
3.C*	maize-free (065)	maize-PDS (057)	maize-other source (117)
3.D*	barley-free (066)	barley-PDS (058)	barley-other source (118)
3.E*	ragi-free (067)	ragi-PDS (059)	ragi-other source (121)
3.F*	small millets-free (068)	small millets-PDS (60)	small millets-other source (120)
4*	pulse-free (071)	pulses-PDS (158)	item 140, 141, 143, 144, 145 & 148
5*	gram-free (072)	N/A	gram whole-other source (142)
6*	salt-free (073)	salt-PDS (178)	salt-other source (170)
7	sugar-free (074)	sugar-PDS (171)	sugar-other source (172)
8*	edible oil-free (075)	N/A	item 181, 182, 183 & 184
9	N/A	kerosene-PDS (334)	kerosene-other source (335)

Source: WB staff based on inputs from the MoSPI (HCES Factsheet, Appendix D). Notes: The item codes from the CES questionnaires are shown in brackets. N/A = Combination of item and source is not available in the survey questionnaire. * Items purchased via the PDS were only available in 2022-23 (not in 2011-12).

[26] Unit value (UV) is calculated as the expenditure on an item divided by the quantity consumed.

The main assumption for the validity of this methodology is that the items hold similar characteristics (such as quality) regardless of their source. In other words, items purchased from the PDS and the local market would provide similar welfare to the consuming household. While second-stage strata are included in the definition of the local market for estimating median prices, differences in quality between PDS and market purchases cannot be *fully* accounted for. Finally, valuation methodology is applied consistently for both surveys, although the PDS item list has shifted between surveys.[27]

The imputed value is calculated as the median UV among unit values from purchases reported by households in the same cluster. The approach is similar to the "hierarchical" approach proposed by Deaton and Zaidi (2002) for pricing in-kind or own-produced food consumption.[28] The hierarchy of clusters is described in Table 5. For all products, the hierarchy first calculates the median UV at the most local or disaggregated level (1). Nonetheless, a sample size of market purchases n≥50 is required. This requirement reduces noise in the estimation of median prices.[29] If the preferred (most disaggregated) cluster does not contain a minimum of 50 market purchases, the next cluster is considered, and the process is repeated sequentially until the minimum sample size requirement is met. In 2011-12, to expand the sample of observations within clusters, microdata from the Type1 questionnaire (URP) was also used in calculating local market medians. In practice, out of 390,587 PDS-related purchases in the CES 2022-23, 85.5 percent receive a median price imputation from hierarchical level (1), which increases to 99.9 percent for hierarchical levels (1)-(4). Survey population weights are applied in the calculation of all median UVs.

Table 5. Hierarchy of clusters to calculate median UVs from market purchases

Level	Cluster for Median calculation	Condition	Preference
(1)	Product, month, state, sector, and SSS	n>50, else move to next level	Most preferred
(2)	Product, month, state, and sector	n>50, else move to next level	↓
(3)	Product, month, and state	n>50, else move to next level	↓
(4)	Product, month, and sector	n>50, else move to next level	↓
(5)	Product and month	n>50, else move to next level	↓
(6)	Product	n>50, else no imputation	Least preferred

Finally, the imputed UVs are tested to ensure that the re-valuation exercise does not result in an implicit welfare loss for PDS-consuming households. In some cases, the methodology described above could yield an estimated median UV lower than the price paid by PDS-subsidized households (implied in their raw data). Although this result is possible, it would mean that the welfare value of PDS items was less than what the household paid for them. This scenario would be equivalent to a PDS "tax" and inconsistent with a non-decreasing welfare measure when adding a

[27] The 2011-12 CES only included four PDS items at subsidized prices: rice, wheat, sugar, and kerosene. Due to the expansion of the PDS policies, the CES 2022-23 includes all items described in Table 5, and the introduction of free items.
[28] The MoSPI proposed a somewhat similar approach to impute the value of food items received for free (Factsheet, Appendix D). In their approach, nonetheless, modal values are used instead of weighted medians. The MoSPI does not impute a value to items purchased at subsidized prices, and kerosene is not included. The MoSPI's imputation exercise is not ultimately included in their nominal consumption aggregate.
[29] Given the sample size of the surveys, the FSU has insufficient sample size to be a valid cluster.

government transfer. Hence, the imputed UV is invalidated for those cases, and the UV is kept as reported by the household in the raw data.

(b) Free non-food items - school uniforms and footwear

Besides free and subsidized food rations, the Government of India (GoI) provides households with other free non-food, non-durable items (such as school uniforms and textbooks), and other subsidized essentials like Liquified Petroleum Gas (LPG) and electricity. Since these items improve household welfare, a market-price-based value of free and subsidized non-food items should ideally be added to the welfare aggregate.

The 2022-23 CES collected information on the number of school uniforms and shoes a household received for free. Among households with any member attending a government educational institution, 17 percent and 8.4 percent report receiving any uniform and shoes free of cost, respectively (**Table 6**). The survey also records information on market purchases of school uniforms and shoes (Section 13 of the DSQ) in the 365 days prior to the interview. This data allows us to impute a market equivalent value of free school uniforms and shoes that a household receives during the reference year.[30,31]

Table 6. Incidence of free clothing and footwear, 2022-23

	National	Rural	Urban
Share of HHs with anyone attending/ attended government institution	32.9%	38.7%	20.2%
Among these HHs, % HHs received clothing (school uniforms) free of cost	17%	18.1%	12.5%
Among these HHs, % HHs received footwear (school shoes) free of cost	8.4%	9.1%	5.9%

The methodology used to impute the values of free school uniforms and shoes is summarized below.[32]

1. Estimate the annual household spending on school uniforms and footwear: The total expenditure on market purchases of school uniforms (items 356 and 357, recording school uniforms for girls and boys, respectively) and footwear (items 390 and 394, recording the purchase of leather

[30] The imputed value for textbooks, LPG cylinders, and electricity is not included. Firstly, assigning a market-based unit value for textbooks is infeasible due to missing information on the number of textbooks purchased and those received for free. Secondly, while the amount of LPG subsidy received by households is available, information on the quantity per LPG cylinder is missing, making it difficult to ascertain the total subsidized quantity received by the household. Finally, the block pricing model for electricity and the variation in the subsidy schemes by state make it challenging to estimate an accurate market value for free electricity.

[31] Since this information is unavailable in 2011-12, the imputed value is only estimated for 2022-23.

[32] A similar approach is followed to the one used by the MoSPI (2022-23 Factsheet) for imputations, with some differences. First, weighted medians are applied, consistent with the approach used to impute other components (food rations) instead of modes. Secondly, like for food ration, a hierarchical approach is used to impute median unit values for school shoes and expenditure per child attending a government institution. Third, these median unit values are only imputed for households that report receiving free uniforms/footwear and with any member reporting attending a government educational institution (and not those attending private institutions). Some households with members attending private educational institutions also report receiving free uniforms and shoes, which could be on account of scholarships, or free items to poorer students, among other reasons. Value of items acquired in private institutions may highly vary due to heterogeneity of items, and therefore, these are excluded from imputations.

shoes or boots, and other footwear, respectively) is added to get the annual expenditure on these items.

2. Estimate market-price unit value for footwear: The survey also records the number of pairs of footwear purchased in the market. The unit value of shoes is the ratio of the market spending to the number of pairs bought in the market.

3. Calculate market-purchase-based unit value on school uniforms as expenses per child attending a government institution: In the absence of information on the number of uniforms purchased, a unit value for uniform is estimated as the ratio of the total spending on school uniforms to the number of children in the household attending a government educational institution.[33] This step assumes that expenditure is evenly distributed between all children in a household attending a government educational institution.

4. Calculate weighted median unit values: Using unit values estimated in steps 2 and 3 above, weighted median unit values are estimated for school uniforms and shoes at different geographical levels (state-sector-second stage sampling unit, state-sector, state, sector, national). These medians are computed over a sample of households with at least one child attending a government institution.

5. Impute unit values faced by a household using a hierarchical approach: Following a hierarchical replacement approach similar to that described in **Table 5,** the weighted median unit value at the most local geographical level with at least 50 non-missing values over which the median is estimated, is assigned as the imputed unit value faced by the household. Specifically, a household is assigned the weighted median unit value at the state-sector-second stage sampling unit if at least 50 households with any member attending a government institution report market purchases of school uniforms and shoes, respectively. If the number of households with market purchases at the most local level falls short, the imputed unit value for a household is determined by the weighted median at a higher geographical level, state-sector, state, sector, or national level, in that specific order.

6. Impute the value of free school uniforms and shoes for households with children attending a government institution: The value for free school uniforms and footwear is estimated as the number of free items received times the imputed unit values using the hierarchical approach described in step 5.

7. Adding the imputed value of free items to the market purchases of respective items: The final step entails adding the imputed value of free school uniforms a household receives for children attending a government institution to the market purchases for school uniforms for girls and boys.

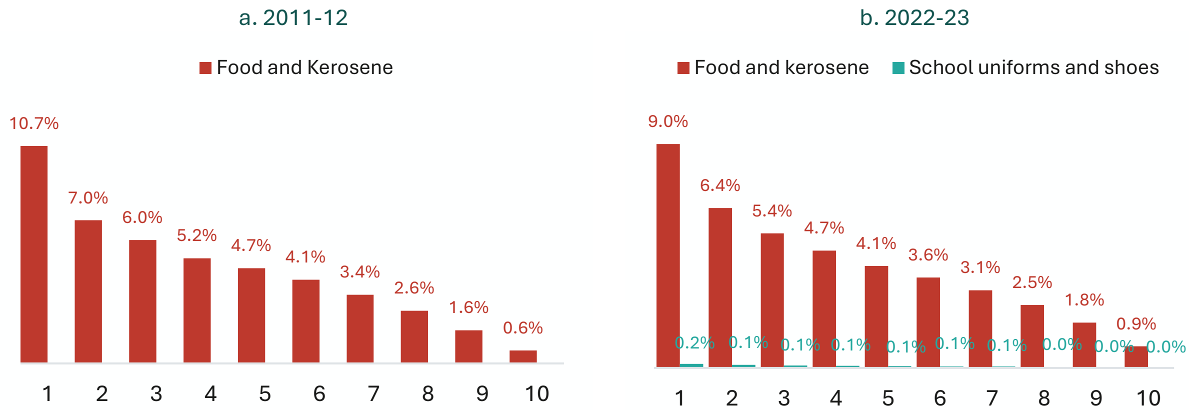
Figure 2 shows the share of the imputed value of food and non-food items as a share of the nominal expenditure aggregate by deciles of consumption expenditure in 2011-12 and 2022-23.

---

[33] Calculating UV in this manner assumes that everyone attending a government institution purchases one uniform from the market in the past 365 days. Moreover, in few households, the reported household size is less than the number of children attending a government educational institute, which implies that these unit values, and consequently, the implied imputed value of free uniforms is a lower-bound on the estimate.

Given that these subsidies are targeted toward poorer households, the average imputed value of free items is larger for those with lower overall expenditures. It gradually decreases, with lower shares among those at the top. Compared to food rations, the imputed value of non-food subsidies is much lower, with only 7 percent of households with a child in a government institution reporting receiving any school uniform or footwear.

Figure 2. Share of imputed free and subsidized items in nominal expenditures, by decile



Note: The figure plots the imputed value of free and subsidized food (as in Section 4.5.1(a)) and imputed value of non-food items (as in Section 4.5.1(b)) as a share of the nominal expenditure aggregate before any adjustments are made. The construction of the nominal expenditure aggregate is described in Section 4.4. The figure presents the share of spending on a category in the total nominal expenditure across deciles, where Decile 1 includes the bottom 10 percent on the nominal expenditure distribution, whereas Decile 10 forms the top 10 percent in the nominal expenditure distribution.

Future rounds of Household CES could be broadened in scope to capture details on households' access to social schemes and the benefits they receive.[34] This would further improve the accuracy of valuing the welfare associated with items received for free or at subsidized prices.

4.5.2 Exclusion of purchase value of durable goods

Durable assets, such as household appliances or electronics, differ from non-durable items due to their ability to add to household utility through repeated use over an extended period, typically spanning more than the survey reference period (past 365 days). Ownership and buying patterns of durable assets vary across households, as these are typically large and somewhat infrequent purchases. Therefore, including the item purchase value rather than its consumption use value over the reference period distorts the welfare ranking of households. Specifically, it tends to overestimate the welfare aggregate of households with more recent purchases while underestimating the welfare aggregate for households that made no purchases during the survey year (but may own several durables) and were using an item purchased before the reference period.

To accurately capture the household's consumption over the reference year, the welfare aggregate ought to include the "value of services" (*consumption flow*) that the household

[34] The World Bank India Poverty and Equity team has drafted a survey questionnaire to identify access to and benefits from social schemes. The questionnaire could complement future rounds of the CES.

receives from owning and using the durable asset over the reference period rather than the purchase price of assets bought during the survey period**.** However, measuring the flow of services from durables is not a trivial exercise as it requires detailed information on the stock of assets, purchase price, purchase year, and estimates of current sale value. Deaton and Zaidi (2002) propose several methods to estimate the consumption flow of durable goods over the reference period, which require a minimum set of information to be included in the survey modules.[35]

The choice of how durables are treated has important implications on the distribution of the welfare aggregate and household rankings. Following international best practices, in the absence of such detailed data in the CES, which includes only the purchase value of items bought during the reference period, the contribution of consumption of durable goods is excluded from the welfare aggregate. Although such exclusion of the flow of services from durables narrows the distribution of the welfare aggregate, it preserves the ranking of households based on the welfare aggregate (Deaton and Zaidi, 2002; Mancini and Vecchi, 2022). **Table 7** lists the durable assets excluded from the welfare aggregate.[36] On average, purchases of durable assets account for 2.4 percent of the nominal expenditure aggregate. Expenditures on durable purchases increase as the aggregate level of consumption expenditure increases: the expenditure share among the top 10 percent was six (three) times that of the bottom 10 percent in the consumption distribution in 2011-12 (2022-23), indicating a significant variation in spending on asset purchases by the level of consumption expenditure (**Table 7**). This variation also implies that excluding durable purchases implies a larger reduction in expenditures at the top of the distribution.

[35] **Annex 4** summarizes these approaches and discusses the set of information that is typically captured in the survey to estimate the consumption flow of durables.

[36] Since the durable purchases are not included, the flow of services households derive from any assets received for free from the government or non-governmental organizations (NGOs) is not imputed.

Table 7. Durable assets excluded from the welfare aggregate

Item No.	Item Description	Item No.	Item Description
43	Electric Air Heaters & Blowers	590	Water Purifier
44	Air cooler	591	Electric Iron, Heater, Toaster, Micro-Oven, Electric Kettle, Vacuum Cleaner, & Other Electric Heating Appliances
45	Air Purifier	592	Other Cooking/ Household Appliances
46	Geysers & Hot Water Systems, Sterilizers	599	Cooking and other household appliances
47	Ventilation System such as Exhaust, Chimney, etc.	600	Bicycle
48	Mixer/Grinder, Juicer, Food Processor	601	Motorcycle, Scooter
560	Television	602	Motor Car, Jeep
561	Camera & Photographic Equipment	603	Tires & Tubes
562	CD, DVD, Pen-Drive, External Hard Disk	604	Other Transport Equipment
563	Musical Instruments	609	Personal transport equipment
569	Goods for recreation	621	Other Machines for Household Work
580	Electric Fan	622	Pc/ Laptop/ Tablet, Other Peripherals Incl. Software
581	Air Conditioner	623	Mobile Handset
582	Inverter, Generators, Stabilizers	624	Telephone Instrument (Landline)
583	Lantern, Lamp, Electric Lampshade, Emergency Light	629	Other personal goods
584	Sewing Machine	587	Pressure Cooker/ Pressure Pan
585	Washing Machine	588	Refrigerator/Freezers
586	Stove, Gas Burner, Induction		

4.5.3 Exclusion of certain non-food non-durable expenditures

The non-food, non-durables expenditure aggregate excludes lumpy and infrequent expenditures that may contribute to household well-being, but a precise value of this contribution is difficult to assess. Additionally, it contains other expenditures that do not necessarily increase household welfare, such as hospitalization costs.

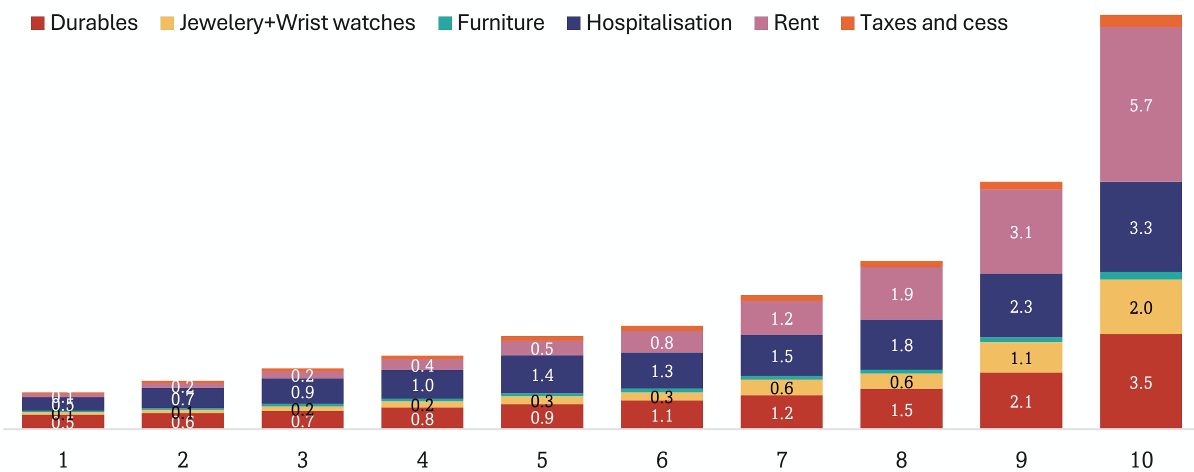
1. **Jewelry and Wristwatches:** Jewelry (DGQ items 640, 641, 642, 649 in the survey) and wristwatches (DGQ item 620) are typically excluded from consumption flow estimates because they do not depreciate and are likely to appreciate over time. Including these items would distort estimates based on the asset's purchase value (Deaton and Zaidi, 2002; Mancini and Vecchi, 2022). These purchases account for an average share of 1 percent of the nominal welfare aggregate in 2022-23, with a significantly higher expenditure by the top 10 percent in the consumption distribution (at 2.4 percent) (**Figure 3**).
2. **Furniture:** Expenditures on items such as bedsteads, couches, and other furniture and fixtures are lumpy, with variability in purchase and sale values on a heterogeneous set of items that make it difficult to assess their worth accurately. Therefore, furniture goods (DGQ items 550 – 559) are not included in the welfare aggregate. These purchases account for an average 0.1 percent of the

nominal expenditure aggregate in 2022-23. In both rounds, the share of expenditures on furniture increases with the aggregate level of spending (**Figure 3**).

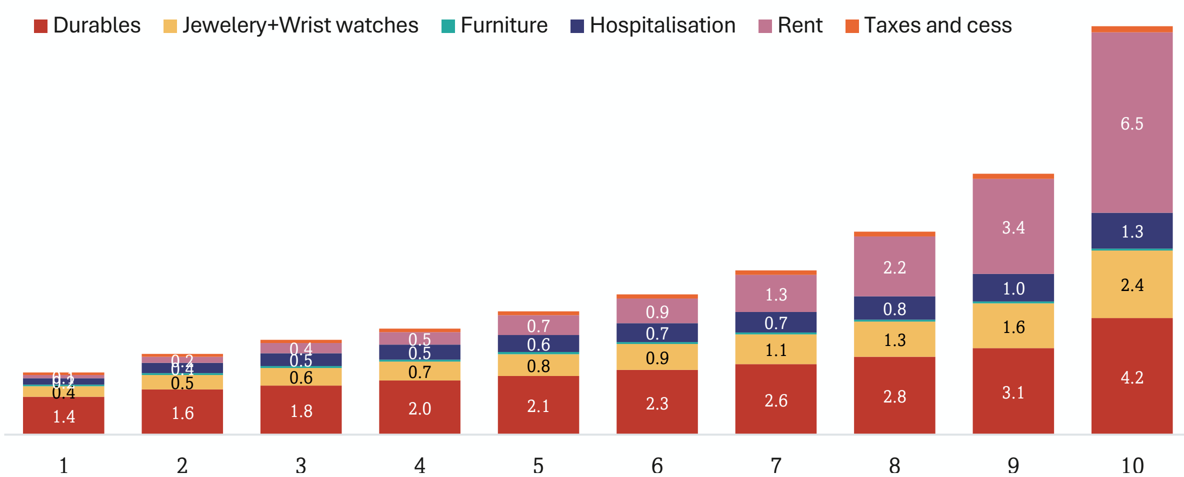
3. **Other taxes and cesses:** Households' spending on other taxes and cesses (items 541 in 2011-12, and 899 in 2022-23) are excluded. These taxes and cesses include road tax, house tax, vehicle tax, and taxes on many other items. While some of these are consumption-related taxes, most are lump-sum taxes paid while purchasing big-ticket items such as cars and are ideally to be distributed across the period over which such assets are owned and used. These expenditures account for a very small share in the nominal welfare aggregate in both years, with a marginally higher share among the top deciles (**Figure 3**).

Figure 3. Share of durables, hospitalization, and select non-food items in nominal expenditures
(by deciles)

a. 2011-12



b. 2022-23



Note: The figure plots the share of each category in the nominal expenditure aggregate, across deciles. Notes to Figure 2 apply. Deciles are created for each sector (rural and urban) separately.

4.5.4 Exclusion of medical hospitalization expenses

Including health costs without accounting for the loss of welfare due to illness would be misleading (for further details, refer to Deaton and Zaidi (2002)). Moreover, some health expenditures, such as hospitalization, are lumpy in nature. According to Mancini and Vecchi (2021), "*If a large hospitalization expenditure is recorded during, say, a three-month reference period, annualizing it and adding it to the nominal consumption aggregate amounts to assuming that whichever household member was recently hospitalized does so on average every year, four times per year—which is patently absurd. The "lumpiness" and infrequency of these health expenditures would suggest exclusion from the [consumption aggregate ...] in accordance with the principle of having to proxy typical consumption*". Therefore, while out-of-pocket preventive healthcare expenses are included in the final aggregate, large lumpy medical expenditures on hospitalization, such as surgeon/doctor fees, medicines, tests, and other hospitalization charges (CSQ items 410 to 419 in the survey), are excluded from the welfare aggregate. In 2022-23, hospitalization spending accounts for an average share of 0.7 percent of the nominal welfare aggregate, with higher expenditures on top of the distribution. In both 2011-12 and 2022-23, the share of expenditures on hospitalization is at least thrice among the top decile compared to the bottom deciles.

4.5.5 Housing services

The housing component of the welfare aggregate captures the value of the services that households derive from their dwellings. Housing services are among the most relevant components of household consumption expenditures (Deaton and Zaidi, 2002). Including the value of housing ensures that the aggregate better captures the relative welfare position of households and the identification of poor and nonpoor households (Ciriani et al., 2023). Conceptually, housing is similar to other durable goods in the welfare aggregate (Mancini and Vecchi, 2022), and the use value of housing over the reference period must be included in the welfare aggregate.

However, survey questionnaires may not have sufficient information to estimate the use-value of housing consistently across all households. Households are typically split between homeowners and renters. While renters usually report rental payments in the survey, the welfare value of housing services that homeowners enjoy cannot be directly measured from rent payments needs to be estimated. The 2011-12 and 2022-23 CES collected rental values paid by households over the last 30 days (actual rents) and self-reported (called *imputed*) rents for dwelling owners but primarily in urban areas.[37]

Not including the housing value of owner-occupied houses may underestimate the welfare of households who own their dwellings and, consequently, skew the rankings in the welfare distribution.[38] In India, only 12.3 percent (13.5 percent) of Indian households were renters in the CES 2022-23 (CES 2011-12). More importantly, the rental market is extremely thin in rural areas—only 3.6

[37] The surveys enquire into the household's unit of accommodation for residential purpose, with options "owned", "rented" or "otherwise occupied without paying any rent" (excluding employer-provided accommodation).
[38] In the past, the expenditures aggregate incorporated the rental values (actual rent) paid by renters but excluded any value of housing services enjoyed by homeowners.

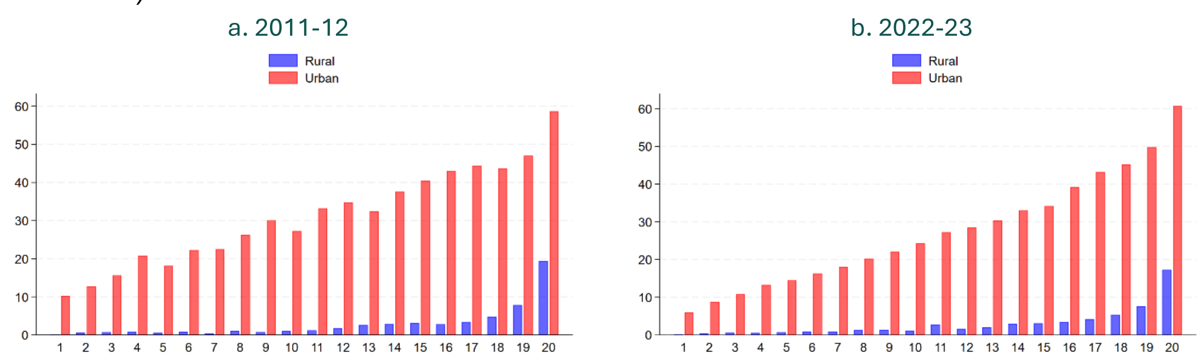
percent of rural households reported "actual" rent payments in rural areas in 2022-23 (**Table 8**). Moreover, relatively well-off urban households are more likely to be renters, with a very low incidence of reported rents in the bottom 40 percent of the consumption distribution in both urban and rural areas (**Figure 4**). Among renters, the average share of reported (actual) rents is also higher among those whose aggregate consumption spending was higher in both survey periods.

Table 8. Share of households reporting actual rental values (in %)

		Dwelling status		Report of rental values		
				Renters	Homeowners	
		Renters	Homeowners	Report actual rent	Report imputed rent	Do not report imputed rent
2011-12	All	13.5	84.0	13.0	19.0	64.9
	Urban	34.7	61.3	34.0	60.9	0.4
	Rural	3.8	94.3	3.4	0	94.3
2022-23	All	12.3	85.5	12.3	32.7	52.8
	Urban	31.1	65.5	31.1	64.9	0.6
	Rural	3.6	94.7	3.6	17.9	76.8

Note: Applying household survey weights. 2 percent of renters also recorded self-assessed rents for owner occupied dwellings in 2022-23, these observations are not shown.

Figure 4. Share of renters, by vigintile and sector
(% of renters)



Note: Share of households living in a "rented" dwelling and reporting a non-zero value of "actual" rent paid over the past 30 days. Household weights are applied. Vigintiles are constructed based on the per capita nominal expenditures aggregate.

A hedonic imputation method is the most common approach to estimate the value of housing services among households that do not report a rental value (Deaton and Zaidi, 2022). Hedonic models seek to predict the implicit rental value of a dwelling based on its observed characteristics—such as location, structural attributes, and neighborhood characteristics for homeowners (Ciriani et al., 2023). An experimental hedonic model was conducted for 2022-23 and presented in **Annex 6.** A similar hedonic rents model was not possible for 2011-12—as dwelling characteristics were not enquired or collected in either the 2011-12 or the 2023-24 CES, and the 2011-12 questionnaire did not collect self-reported (imputed) rents in rural locations, leading to a sample that is missing not-at-random (MNAR).

Despite best efforts to impute rent, the available CES data seems inadequate to support accurate and unbiased predictions via a hedonic rents model. The thin rental market captured in the CES and skewness in the available observations towards wealthier and urban households likely contributes to an overestimation of housing values at the bottom of the distribution and for rural dwellings. This is supported by an in-sample accuracy test in **Annex 6**, which suggests that the hedonic model may overestimate rental values at the lower end of the renters' distribution. In countries where few households pay house rents, rental predictions and equivalents can be inaccurate, and "the benefits of completeness need to be weighed against the costs of error" (Deaton and Zaidi, 2002). Unfortunately, the likelihood of introducing costly prediction errors under the current CES configuration is high.

Considering these arguments, rent expenses (for renters) were excluded from the welfare aggregate to avoid inconsistencies in valuation between renters and homeowners. Given the inability to consistently estimate housing value for homeowners in 2011-12 and 2022-23, the only viable option is to remove the distorting effect of actual rents reported by renters only. As explained above, partially including rents for renters alone distorts the household rankings—artificially making homeowners appear to have lower welfare than renters due to the exclusion of their housing services.

Further research, along with improvements in questionnaire design and implementation, will be fundamental to accurately estimating the welfare contribution of housing. Although the CES 2022-23 collected information on basic dwelling characteristics,[39] other important variables that differentiate dwelling sizes and quality (e.g., number of rooms) are missing. Including more disaggregated geographic variables could similarly improve the precision of rental imputations.[40] Finally, addressing the issues of missing data in rural areas—77 percent of rural homeowners did not report an imputed rental value in 2022-23—will be complex but key to reducing biases in the available data and training an accurate hedonic rents model in the future.

4.6 From nominal to real: Constructing the real welfare aggregate

Next, the nominal aggregates are converted to real aggregates by adjusting for spatial and temporal variation in prices. A three-step sequential adjustment process, summarized in **Table 9** is applied to express expenditure and welfare aggregates in real terms.

[39] The CES 2022-23 questionnaire included dwelling characteristics: construction materials for floors, roofs and walls; primary source of drinking water; time taken by the household for a single trip to reach the source of water; and access to latrine and type of latrine; primary cooking fuel, primary energy source for lighting, and total area owned.

[40] Beyond traditional survey data, other variables that can help predict rent and housing prices include the historical significance or architectural style, local zoning regulations or development potential, flood risk or other exposure to natural hazards, and proximity to negative externalities (e.g., landfills, industrial sites).

Table 9. High-level summary of temporal and spatial price adjustments

Intra-temporal Adjustment	Spatial Adjustment	Inter-temporal Adjustment
To adjust monthly expenditure on each item for within-survey-year price variations. ❑ Using 20 CPI subgroups at the state-sector month level to deflate expenditure. ❑ Result: Expenditure expressed in annual survey-year average prices at the state-sector level.	To make household expenditures comparable across the country. ❑ Fisher Price Indices are constructed following the Tendulkar Committee (2009) to adjust for price variation across states and urban-rural areas. ❑ Fisher Indices for 15 commodity groups where information on quantity is available (food, clothing, and footwear). ❑ For the remaining items, nominal budget shares and deflated expenditure on 15 commodity groups were used to calculate the total household expenditure adjusted for spatial deflation. ❑ Result: Expenditure is expressed in the national average prices for all households.	To express household expenditures in 2017 prices, the base year for PPP conversion. ❑ 2011-12 per capita expenditures are expressed in 2017 prices. ❑ 2022-23 per capita expenditures are expressed in 2017 prices. ❑ Result: Expenditure is ready to be expressed in 2017 PPPs.

4.6.1 Intra-temporal price adjustment within survey year

Within-survey price adjustments help ensure seasonal price differences do not distort household rankings. Since the CES spans 12 months, some households may report higher expenditures due to price fluctuations rather than increased consumption. Item-level expenditures are adjusted for price variations within the survey period. To implement this, each item in the 2011–12 and 2022–23 CES was mapped to one of 20 CPI subgroups, with adjustments applied at the rural/urban state level, to express expenditures in average survey year prices.

$$x_{jsy} = \frac{X_{jst}}{\frac{CPI_{Gst}}{\overline{CPI}_{Gsy}}} \qquad for\ j \in G$$

where:

x_{jsy} : Within-year adjusted expenditure on item *j* in state-sector *s* and survey year *y*
X_{jst} : Nominal expenditure on item *j* in state-sector *s* and month *t*
$\overline{CPI}_{Gsy}$: Survey-year average CPI of subgroup G, in state-sector s and survey-year y
CPI_{Gst} : Monthly CPI of subgroup G, in state-sector s and month t

For smaller states and union territories lacking subgroup-level data such as Chandigarh, Diu & Daman and Dadra & Nagar Haveli, Puducherry, the Andaman & Nicobar Islands, Goa, and the Northeastern states—group-level CPI was used.[41]

4.6.2 Spatial price adjustment

This adjustment accounts for spatial cost-of-living differences, ensuring comparability of household expenditures across regions. Fisher price indices, based on unit values, are constructed following the Tendulkar Committee (2009) methodology to adjust for spatial price variations. Originally, these indices were used to establish state-sector poverty lines. In 2011-12, the approach set the national urban poverty rate and its corresponding poverty line as the base, requiring three sets of deflators: (i) state urban relative to all-India urban, (ii) state rural relative to state urban, and (iii) state rural relative to all-India rural. Under the revised methodology, Fisher indices are applied to adjust the welfare aggregate for comparability across households in 2011-12 and 2022-23. Additionally, the latest adjustments use only rural/urban state indices relative to all-India, expressing expenditures in a single unit—average national prices.

The main steps followed in constructing the spatial deflators are as follows. See **Annex 4** for details.

I. *Classification of items*: All items were classified into commodity groups based on the item type. Commodity groups 1-15 correspond to groups defined by the Tendulkar Committee to construct spatial deflators and constitute Food, Clothing, and Footwear items, where the data allowed for the calculation of unit values.
II. *Creation of product codes:* Group item codes into product codes by consolidating the same item from different sources (PDS, free, and market purchases) under a single product code. This allows for the calculation of unit values that reflect the household's consumption pattern (e.g., one household with a high share of subsidized consumption will have lower unit values for that product).
III. *Computation of item-level statistics*: Calculate unit values for each product at the household level using the intra-temporally adjusted expenditure and the quantity consumed. Further, calculate population-weighted median prices and average quantities consumed at state-sector and all-India levels.
IV. *Exclusion of low-frequency items:* Products consumed by five or fewer households in a state-sector were excluded from the exercise and treated as not consumed in that region.
V. *Valuation of bundles*: Calculate the value of bundles for each commodity group at the state-sector level and all India levels by aggregating (median price x average quantity) of items within a commodity group.
   i. Value of state-sector bundle at state-sector median prices
   ii. Value of state-sector bundle at all-India median prices
   iii. Value of all-India bundle at all-India median prices

[41] Exact details are available in the accompanying reproducibility package.

iv. Value of all-India bundle at state-sector median prices

Population projections at the state-sector level from the Ministry of Health and Family Welfare were used to calculate the valuation of bundles.

VI. *Calculation of price indices for each commodity group, separately for rural and urban in each state*:

I. *Paasche index* = value of state-sector bundle at state-sector prices/value of state-sector bundle at all-India prices

II. *Laspeyres index* = value of all-India bundle at state-sector prices/value of all-India bundle at all-India prices

III. *Fisher index* = geometric mean of Paasche and Laspeyres index

These indices were computed separately for rural and urban areas across 15 commodity groups, resulting in 960 indices (32 states x 2 sectors x 15 commodity groups). After adjusting for survey-period price variations, Fisher indices were used to deflate expenditures at the commodity group level for each household. This process provides spatially adjusted expenditures for items within these 15 commodity groups at the household level. The implicit spatial deflator is then calculated for each household as the ratio of intra-temporally adjusted expenditure on the 15 commodity groups to spatially adjusted expenditure on the same 15 commodity groups. Since spatial deflators vary across state-sectors and household expenditure allocations differ by commodity group, each household has a unique adjustment factor that reflects variations in both consumption patterns and regional price levels.

To ensure household deflators are relative to a unified all-India price, the weighted average of deflators is scaled such that it sums to one.

Lastly, once this portion of a household's expenditure (fifteen commodity groups) has been deflated, the remaining portion, for which unit values cannot be calculated, was calculated residually. This extrapolation was done under the assumption that the share of this share of expenditure in real terms is exactly equal to the share of expenditure in nominal terms.

4.6.3 Inter-temporal adjustment across years

Real consumption aggregates are adjusted to PPP-year prices for international poverty estimation. The above constructed the intra-temporal and spatially deflated aggregate (real aggregate) is in survey-period prices, while international poverty lines are defined in PPP-year prices. Since the PPP conversion factors and international poverty lines are based in 2017, the real welfare aggregate is further adjusted using a national CPI index retrieved from the International Monetary Fund's (IMF) International Financial Statistics (IFS) database.

$$MPCE_{y,2017} = \frac{MPCE_y}{\frac{CPI_y}{CPI_{2017}}}$$

$MPCE_{y,2017} = Real\ welfare\ aggregate\ in\ 2017\ price$

$MPCE_y \quad = Real\ aggregate\ in\ survey\ year\ prices$

$CPI_y \quad = Survey\ year\ average\ General\ CPI\ index\ (All - India)$

$CPI_{2017} \quad = 2017\ calendar\ year\ average\ General\ CPI\ index\ (All - India)$

4.7 Net effect of all adjustments on the welfare aggregate

This section briefly discusses the implications of all adjustments made to construct the real welfare aggregate.

Spatial and intra-temporal price adjustments affect the expenditure distribution. Annex Figure A.7.1 plots the cumulative density functions (CDFs)[42] for the nominal and real expenditure. The orange line (real expenditure) lies to the right of the red line (nominal expenditure) in rural areas, indicating higher real purchasing power after adjusting for spatial price differences. This is because the spatial price adjustments increase consumption expenditure in rural areas when measured in national prices. The urban areas, expectedly, have a contrasting effect of adjustments.

Table 10 presents the marginal impacts on the real welfare aggregate from performing each of the proposed adjustments discussed in section 4.5.[43] The first row presents the real expenditure aggregate, which is treated as a baseline scenario—without any adjustments—for the remainder of this section. This baseline real expenditure aggregate is, on average, INR 2,435 and 3,363 in 2011-12 and 2022-23, respectively, expressed in 2017 prices.

The real expenditure aggregate is first adjusted to include the imputed value of free and subsidized food rations obtained under the PDS and free school uniforms and shoes provided to students attending public institutions. The marginal effect of this inclusion is presented as Scenario I in **Table 10**. Including the value of free and subsidized food and schooling items leads to a 2.8 percent increase in the average real aggregate relative to the baseline in 2022-23 (at INR 3,456 in 2017 prices). Subsequently, Scenarios II, III, and IV exclude lumpy expenditures (durables, taxes, furniture, and jewelry), hospitalization, and rents, respectively, relative to the baseline scenario. Each exclusion results in a downward revision of the average real expenditure aggregate. The final row presents the real welfare aggregate when all adjustments are applied simultaneously. Under this final scenario, the real welfare aggregate is lower by 6.6 percent in 2011-12 and 7.4 percent in 2022-23, relative to the baseline aggregate without any adjustments.

[42] A CDF shows the proportion of population with consumption less than a certain value in the distribution, therefore capturing the spread of the distribution.

[43] For ease of comparisons, from here on, the note presents the marginal effects of different imputations and expenditure adjustments on the final real welfare aggregate relative to a baseline of the real expenditure aggregate.

Table 10. Marginal effects of adjustments to the real welfare aggregate
(in INR, 2017 prices)

Scenario	Rural		Urban		National	
	2011-12	2022-23	2011-12	2022-23	2011-12	2022-23
Baseline real expenditure aggregate	**2,088**	**2,954**	**3,304**	**4,396**	**2,435**	**3,363**
I. Include imputed value of PDS rations and free school items	2,149	3,072	3,335	4,427	2,487	3,456
II. Exclude purchase value of durables and lumpy expenses	2,016	2,808	3,143	4,141	2,338	3,186
III. Exclude hospitalization	2,043	2,884	3,239	4,312	2,385	3,289
IV. Exclude rental expenses	2,079	2,934	3,102	4,120	2,372	3,270
Final real welfare aggregate	**2,022**	**2,836**	**2,909**	**3,817**	**2,276**	**3,114**

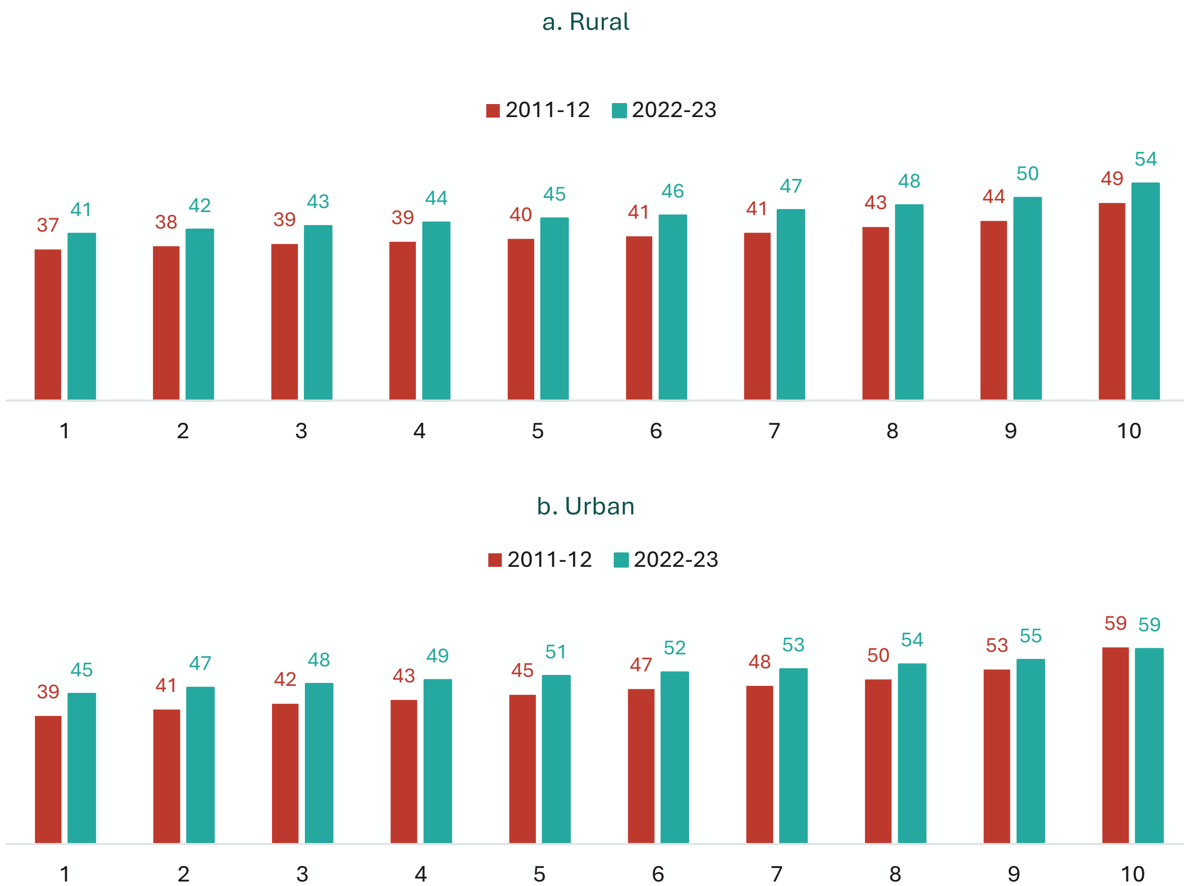
Note: Each row presents marginal effects: scenarios I to IV apply one adjustment at a time relative to the baseline, allowing the measurement of individual impacts. In all cases, corrections for spatial and temporal price differences are applied. The last row represents the final real welfare aggregate obtained by implementing all adjustments simultaneously. Since multiple adjustments may affect the same households, the marginal estimates from scenarios I to IV may not necessarily add up to the final change from baseline real expenditures to final welfare aggregate.

The shift from a real expenditure aggregate to a real welfare aggregate has uneven effects across the distribution, leading to the re-ranking of households across deciles. Annex Tables A.7.1–A.7.4 show the extent of movement between deciles in rural and urban areas for 2011–12 and 2022–23. The bottom decile remains relatively stable, with only about 1 percent of individuals (10 percent of the decile) shifting out after adjustments. The top decile is also relatively consistent, with approximately 8 percent of individuals (80 percent of the decile) continuing to remain in the top decile post-adjustments. In contrast, the middle of the distribution is less stable, with 30 percent to 70 percent of individuals in each decile switching deciles, indicating greater sensitivity of these population segments to welfare aggregate adjustments.

The transition from a real expenditure to a real welfare aggregate generally increases welfare among rural households and reduces welfare among urban households. In rural areas, these adjustments lead to increased welfare in both 2011–12 and 2022–23 **(Annex Figures A.7.1a and A.7.1b)**. This is particularly evident at the lower end of the distribution, where individuals benefit more from free and subsidized food and non-food items. In 2022-23, for instance, the increase in welfare due to the receipt of subsidized items was more, on average, than the decline in welfare due to excluding other expenditures for the bottom four deciles (**Annex Figure A.7.2b)**. However, the net effect of adjustments at the top of the rural distribution, where individuals spend more on rents and durables, is negative. Overall, the rural welfare distribution shifts rightward relative to the expenditure distribution and becomes flatter both at the lower end and the top tail of the distribution, indicative of reduced inequality. In contrast, the urban real welfare distribution generally lies to the left of the urban real expenditure distribution, except at the very bottom **(Annex Figures A.7.1c and A.7.1d)**. For most of the urban distribution, the adjustments imply lower welfare levels driven by the exclusion of lumpy expenditures and rents, resulting in a compression of the distribution (**Annex Figures A.7.2c and A.7.2d**).

Importantly, estimates of consumption-based inequality using the real welfare aggregate should be interpreted with caution. Government-subsidized food rations and schooling primarily benefit the poor and vulnerable, and including their imputed values enhances welfare, particularly at the lower end of the distribution—excluded items are less than 2 percent of the nominal expenditure aggregate of the bottom 10 percent in 2022-23 **(Figure 3)**. Conversely, excluding durables, lump expenditures (i.e., hospitalization, jewelry etc.,) and housing underestimates welfare, especially for higher-income households—excluded categories are over 10 percent of the nominal expenditure aggregate for the top 10 percent in 2022-23 **(Figure 3)**. These exclusions—necessitated by data limitations in the CES questionnaires—also result in underestimated average non-food shares across the distribution **(Figure 5)**. While the real welfare aggregate ensures a consistent ranking of households, these adjustments make the overall distribution appear less unequal.

Figure 5. Non-food share in real welfare aggregate, by decile
(% of real welfare aggregate)



Note: The figure presents share of deflated non-food aggregate in the real welfare aggregate by deciles of real expenditure aggregate. Each decile represents 10 percent of the population, where Decile 1 includes bottom 10 percent on the real welfare distribution, whereas decile 10 forms the top 10 percent in the real welfare distribution.

5. Poverty and Inequality Trends, 2011-12 and 2022-23

This section presents estimates of poverty and inequality for 2011-12 and 2022-23 based on the distribution of the MMRP-based final real welfare aggregate. Individuals whose real welfare aggregate, expressed in 2017 PPP terms per person per day, falls below the respective international poverty line ($2.15, $3.65, $6.85 per capita per day) are classified as poor by that standard.

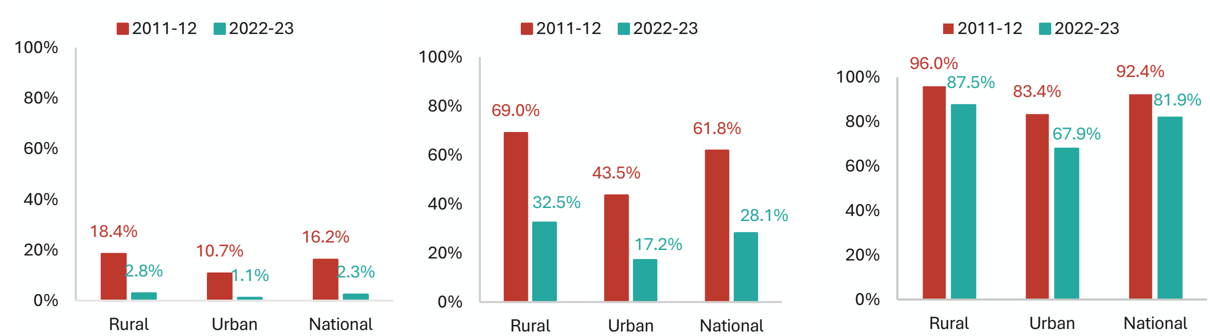
Consumption inequality is measured here by the Gini index and percentile ratios—a measure of statistical dispersion that captures the extent of consumption inequality based on the distribution of the final real welfare aggregate. This index, which ranges from 0 to 100, reflects the degree of inequality in the real welfare aggregate. A Gini index of 0 represents perfect equality, while a value of 100 indicates perfect inequality. In addition, percentile-based measures of dispersion are also presented. The P90/P10 ratio compares the consumption of the top 10 percent to the bottom 10 percent of the real per capita welfare distribution. The P75/P25 ratio compares the top 25 percent to the bottom 25 percent, which captures inequality in the middle of the distribution and is less sensitive to underreporting at the extremes.

5.1 Poverty trends using international poverty lines

Poverty rates declined significantly between 2011-12 and 2022-23 across all thresholds. Based on the real welfare aggregate against the threshold of $2.15 per person per day poverty line in 2017 PPP, extreme poverty rates fell from 16.2 percent to 2.3 percent (**Figure 6a)**. This implies that approximately 33 million people live in extreme poverty in 2022-23.[44] Poverty incidence measured against the lower-middle-income country (LMIC) poverty line ($3.65 per person per day threshold in 2017 PPP), which is considered more appropriate for India's level of development, decreased from 61.8 percent to 28.1 percent **(Figure 6b**). Upper-middle-income poverty, using the $6.85 per person per day poverty line in 2017 PPP, also declined, falling from 92.4 percent to 81.9 percent over the same period (**Figure 6c**).

Figure 6. Poverty rates using international poverty lines

a. Extreme poverty line ($2.15 per day, 2017 PPP) | b. LMIC poverty line ($3.65 per day, 2017 PPP) | c. UMIC poverty line ($6.85 per day, 2017 PPP)



[44] Based on a total population of 1,417,173,173 people in India (2022). This figure is based on United Nations projections and retrieved from the WDI (https://databank.worldbank.org/source/population-estimates-and-projections).

Over the past 11 years, extreme poverty has declined significantly in rural and urban areas, narrowing the rural-urban gap. The extreme poverty rate in rural areas fell from 18.4 percent to 2.8 percent, while urban poverty declined from 10.7 percent to 1.1 percent. The gap between rural and urban poverty has decreased from 7.7 to 1.7 percentage points. However, the speed of extreme poverty reduction was similar across urban and rural areas, with an overall annual reduction rate of 16 percent at the national level.

Moderate poverty also declined considerably. In rural areas, the poverty headcount ratio fell from 69.0 percent to 32.5 percent, decreasing from 43.5 percent to 17.2 percent in urban areas. At this threshold, the urban-rural gap in poverty also narrowed (from 25 to 15 percentage points), although the reduction rate was lower compared to extreme poverty. The annual reduction rate was approximately 7 percent nationally, with very small differences across urban and rural areas.

The UMIC poverty rate has also fallen, with a more pronounced decline in urban areas. Urban UMIC poverty decreased by 15.5 percentage points, from 83.4 percent in 2011-12 to 67.9 percent in 2022-23, representing an annualized reduction rate of 1.9 percent. In contrast, rural UMIC poverty saw a more modest decline of only 8.6 percentage points, representing an annual reduction rate of 0.8 percent.

Methodological improvements in moving from an expenditure aggregate to a real welfare aggregate resulted in lower extreme poverty, negligible impacts on LMIC poverty, and higher UMIC poverty. Extreme poverty was revised downwards in both rural and urban areas, except in urban areas in 2011 (**Table 11)**. The adjustments resulted in virtually no change in poverty at the LMIC poverty line **(Table 12**), whereas poverty at the UMIC thresholds was higher by 4.8 percentage points (**Table 13**).

Table 11. Extreme poverty rates at each adjustment stage ($2.15 per day, 2017 PPP, in %)

	Rural		Urban		National	
	2011-12	**2022-23**	**2011-12**	**2022-23**	**2011-12**	**2022-23**
Baseline real expenditure aggregate	**20.0 (0.4)**	**4.0 (0.1)**	**10.0 (0.3)**	**1.2 (0.1)**	**17.2 (0.3)**	**3.2 (0.1)**
I. Include imputed value of PDS rations and free school items	16.8 (0.3)	2.4 (0.1)	9.0 (0.3)	0.8 (0.1)	14.5 (0.3)	2.0 (0.1)
II. Exclude purchase value of durables and lumpy expenses	21.0 (0.4)	4.5 (0.1)	10.5 (0.3)	1.4 (0.1)	18.0 (0.3)	3.6 (0.1)
III. Exclude hospitalization	20.9 (0.4)	4.1 (0.1)	10.5 (0.3)	1.3 (0.1)	18.0 (0.3)	3.3 (0.1)
IV. Exclude rental expenses	20.1 (0.4)	4.0 (0.1)	11.0 (0.3)	1.3 (0.1)	17.5 (0.3)	3.2 (0.1)
Final real welfare aggregate	**18.4 (0.4)**	**2.8 (0.1)**	**10.7 (0.3)**	**1.1 (0.1)**	**16.2 (0.3)**	**2.3 (0.1)**

Note to Table 10 applies. Standard errors are presented in parentheses and are computed using the survey design where an FSU is selected from a state, sector, stratum, and substratum, and a household is chosen from one of the three Second Stage Substratum (SSS) formed within these FSU. Single sampling units (households) within any strata are treated as certainty units, indicating no variability within that state-sector-stratum-substratum-sss.

Table 12. LMIC poverty rates at each adjustment stage ($3.65 per day, 2017 PPP, in %)

	Rural		Urban		National	
	2011-12	2022-23	2011-12	2022-23	2011-12	2022-23
Baseline real expenditure aggregate	**68.3 (0.4)**	**33.4 (0.3)**	**38.9 (0.5)**	**14.3 (0.3)**	**59.9 (0.3)**	**28.0 (0.2)**
I. Include imputed value of PDS rations and free school items	65.9 (0.4)	28.0 (0.3)	37.9 (0.5)	13.1 (0.3)	57.9 (0.3)	23.8 (0.2)
II. Exclude purchase value of durables and lumpy expenses	69.8 (0.3)	36.5 (0.3)	40.3 (0.5)	16.0 (0.3)	61.4 (0.3)	30.7 (0.2)
III. Exclude hospitalization	69.6 (0.3)	34.6 (0.3)	39.8 (0.5)	15.1 (0.3)	61.1 (0.3)	29.1 (0.2)
IV. Exclude rental expenses	68.4 (0.4)	33.6 (0.3)	41.9 (0.5)	15.9 (0.3)	60.9 (0.3)	28.6 (0.2)
Final real welfare aggregate	**69.0 (0.3)**	**32.5 (0.3)**	**43.5 (0.5)**	**17.2 (0.4)**	**61.8 (0.3)**	**28.1 (0.2)**

Note to Table 11 applies.

Table 13. UMIC poverty rates at each adjustment stage ($6.85 per day, 2017 PPP, in %)

	Rural		Urban		National	
	2011-12	2022-23	2011-12	2022-23	2011-12	2022-23
Baseline real expenditure aggregate	**94.6 (0.1)**	**84.4 (0.2)**	**77.9 (0.4)**	**58.8 (0.4)**	**89.8 (0.2)**	**77.1 (0.2)**
I. Include imputed value of PDS rations and free school items	94.4 (0.1)	83.0 (0.2)	77.6 (0.4)	58.2 (0.4)	89.6 (0.2)	75.9 (0.2)
II. Exclude purchase value of durables and lumpy expenses	95.5 (0.1)	87.0 (0.2)	79.9 (0.4)	62.4 (0.4)	91.0 (0.2)	80.0 (0.2)
III. Exclude hospitalization	95.2 (0.1)	85.7 (0.2)	78.8 (0.4)	60.1 (0.4)	90.5 (0.2)	78.4 (0.2)
IV. Exclude rental expenses	94.8 (0.1)	84.7 (0.2)	80.8 (0.4)	63.1 (0.4)	90.8 (0.2)	78.6 (0.2)
Final real welfare aggregate	**96.0 (0.1)**	**87.5 (0.2)**	**83.4 (0.4)**	**67.9 (0.4)**	**92.4 (0.1)**	**81.9 (0.2)**

Note to Table 11 applies.

The revaluation of free and subsidized PDS items had the most significant impact on reducing extreme poverty. Including imputed values for subsidized and free food and non-food items significantly increased expenditures, particularly for poorer households in rural areas, as these provisions constitute a larger share of their total consumption. In 2011-12, the international extreme poverty rate dropped by 2.7 percentage points compared to the baseline scenario, using the international poverty line of $2.15 (in 2017 PPP). On the other hand, the reduction in extreme poverty was 1.2 percentage points in 2022-23. These downward revisions represented a reduction of 15 percent in the poverty headcount from the baseline in 2011-12 and a sizable 38 percent in 2022-23 (scenario I vs. baseline, **Table 11**). Revaluing subsidized and free items also reduced poverty with respect to the LMIC international poverty line of $3.65 (2017 PPP) by 4.2 percentage points in 2022-23, a 15 percent decline in poverty rates from the baseline scenario (scenario I vs baseline, **Table 12**). Both revisions are in line with the expanded scope and coverage of the PDS in 2022-23 relative to 2011-12.

Box 1. International Poverty Lines with the 2021 PPPs

A transition to 2021 PPP for international poverty estimates and an accompanying revision of IPLs is expected later in 2025. The change due to the PPP conversion factor will likely be minimal for India. Using 2017 PPP,1 PPP$ = INR 19.46904, while with the 2021 PPP,1 PPP$ = INR 19.46895. Consequently, the changes in India's estimates due to a shift towards 2021 PPP would come through the implied change in international poverty lines.

While the international poverty lines are yet to be finalized, Table 14 illustrates the likely effects on poverty rates with a range of hypothetical alternatives for the international poverty lines. This table presents a range of estimates with four poverty line values for each level of poverty (extreme, LMIC, UMIC). The darker cells in the range of alternatives for international poverty lines ($2.60, $4.40, and $8.30, respectively) are the inflation-adjusted value of the existing IPLs (in 2017 PPP, i.e., $2.15, 3.65, and 6.85, respectively) using the 2021 PPP factor. For example, the poverty line of $2.60 in 2021 PPP yields the same extreme poverty rate of 2.3 percent in 2022-23 as with the $2.15 (2017 PPP). A revised extreme poverty line of $3.00 would constitute a 15 percent higher threshold than $2.15 expressed in 2021 prices and result in a 5.3 percent poverty rate in 2022-23. A new LMIC line of $4.20 would imply a 5 percent lower threshold for poverty than $3.65 adjusted in 2021 prices and yield a poverty rate of 23.9 percent.

Table 14. Headcount poverty estimates, from a range of hypothetical IPLs in 2021 PPP (%)

		Range of alternative IPLs (in $ 2021 PPP)			
		(1)	(2)	(3)	(4)
	Year	**2.6**	**3.0**	**3.4**	**3.8**
Extreme Poverty	2011-12	**16.2**	27.1	38.0	48.7
	2022-23	**2.3**	5.3	10.0	16.3
	Year	**4.2**	**4.4**	**4.8**	**5.2**
LMIC Poverty	2011-12	57.7	**61.8**	68.2	73.7
	2022-23	23.9	**28.1**	36.2	44.0
	Year	**8.3**	**8.7**	**9.1**	**9.5**
UMIC Poverty	2011-12	**92.4**	93.4	94.2	94.9
	2022-23	**81.9**	84.4	86.4	88.2

Note: The darker cells in this table represent the rounded value of existing IPLs (in 2017 PPPs, i.e. $2.15, $3.65, and $6.85, respectively) expressed in 2021 PPP terms.

5.2 Consumption Inequality

Table 15 presents the Gini index based on the distribution of the real aggregates. In 2022-23, India's Gini index based on the final real welfare aggregate stands at 25.5, with urban areas displaying slightly higher inequality (27.9) compared to rural areas (22.8) (last row, **Table 15**).[45] This suggests greater disparities in per capita consumption spending in urban areas.

Methodological improvements and their effect on inequality estimates

As discussed in Section 4.7, the transition from an expenditure aggregate to a consistent real welfare aggregate also has implications for measured inequality in both 2011-12 and 2022-23. Compared to the baseline scenario, inequality was revised downwards by 4 Gini points due to the various adjustments to the welfare aggregate (**Table 15**). As a result, these consumption-based inequality measures likely underestimate true inequality due to the exclusion of certain expenditures (discussed in **Section 4.6**) and the underrepresentation of high-income individuals due to data collection limitations—suppressing consumption spending at the upper end of the distribution.

Table 15. Gini index

	Rural		Urban		National	
	2011-12	2022-23	2011-12	2022-23	2011-12	2022-23
Baseline real expenditures aggregate	**27.7**	**26.0**	**36.4**	**31.3**	**32.7**	**29.5**
Include imputed value of PDS rations and free school items	26.9	24.8	35.9	30.8	31.9	28.3
Exclude purchase value of durables and lumpy expenses	26.3	25.0	34.9	30.2	31.2	28.4
Exclude hospitalization	27.1	25.4	36.2	31.1	32.3	29.0
Exclude rental expenses	27.5	25.7	35.2	30.1	31.7	28.4
Final real welfare aggregate	**24.8**	**22.8**	**32.6**	**27.9**	**28.8**	**25.5**

Note to Table 10 applies.

Percentile-based measures of inequality

The percentile-based measures reflect patterns in distributional inequality observed with the Gini index. In 2022–23, the national P90/P10 ratio is 4.9, meaning the top 10 percent consume 4.9 times more than the bottom 10 percent, down from 5.8 in 2011–12 (**Annex Figure A.7.3a**). Similarly, the P75/P25 ratio fell from 3.5 to 3.1 during the same period (**Annex Figure A.7.3b**). Both ratios have decreased since 2011–12, aligning with the overall compression of the consumption distribution, particularly at both ends, as reflected in the CDFs. The narrowing of these ratios is particularly evident in urban areas, where inequality remains higher but has experienced sharper declines. This

[45] The adoption of 2021 PPPs does not affect the Gini index since the entire distribution is multiplied by the same factor.

improvement is also influenced by the earlier methodological refinements, specially constructing a welfare aggregate that includes in-kind transfers and excludes irregular spending categories.

Downward bias in estimates of consumption inequality

While the above indicators suggest a decline in consumption inequality, they are likely downward biased due to several factors. First, the underrepresentation of high-income households in the 2022-23 CES—common in middle—income country household surveys—suppresses the consumption distribution's upper tail. Second, using a sampling frame based on the 2011 census implies a relatively high rural share in the 2022-23 sample that also pulls inequality estimates lower, as rural consumption tends to be more equal and lower overall than urban consumption. Additionally, adjustments made in constructing the real welfare aggregate further compress the distribution, as shown in **Annex Figure A.7.1**. Together, these factors suggest that the survey-based Gini index underestimates the true extent of inequality in consumption.

Inequality in consumption versus income and wealth

It is important to note that consumption-based inequality, while an important welfare measure, is structurally lower than income or wealth inequality, and acts as a necessary but not sufficient measure to understand aggregate inequality in the country.[46] Consumption is generally a smoother indicator than income or wealth, as households are liquidity-constrained and may adjust expenditures over time in response to income shocks or savings. Moreover, savings and asset holdings that may affect income or wealth inequality measures are not captured in consumption data. Additionally, social transfers—many of which are not fully captured in CES data—may further compress observed consumption differences, especially at the lower end of the distribution. Consequently, even with accurate measurement, consumption-based inequality measures, such as those derived from the CES, typically reflect lower disparities than income or wealth metrics.

In summary, while the 2022-23 CES suggests a decline in consumption inequality based on the real welfare aggregate, the true extent of aggregate inequality remains underestimated due to data limitations, high rural shares, methodological choices, and the nature of consumption as an indicator. Therefore, the findings on inequality should be interpreted with caution as lower-bound estimates, highlighting the need for complementary data sources to assess broader dimensions of inequality.

6. Sensitivity checks

6.1 Correcting for outliers in unit values

Since the survey collects self-reported expenditures and quantities, recording and recall errors or extreme values may bias the poverty estimates. To estimate the sensitivity of these estimates

[46] The World Inequality Database shows income inequality rising from a Gini of 52 in 2004 to 62 in 2023. Moreover, wage disparity remains high, with the median earnings of the top 10% being 13 times higher than the bottom 10% in 2023-24 (WBG staff calculations based on PLFS 2023-24).

to such biases, outlier corrections are applied to the survey-based unit values for items for which quantities and expenditures are available in the survey.

The first step was to apply a rule to identify outliers in the implicit unit values for each item. Two rules are applied to identify these outliers—a unit value is an outlier if the absolute standardized logarithmic unit value is (a) greater than 3, or (b) greater than 4.[47] Once the outliers are flagged, the unit values are replaced with the median unit value at the most local geographical level, following the hierarchical approach described in Section 5.1. Next, the nominal and real welfare aggregate, as well as the poverty incidence are re-estimated using these outlier-corrected unit values. **The outlier correction has an insignificant impact on measures of poverty and inequality using either interval to flag the outliers (Table 16).**

Table 16. Poverty and inequality after correcting for outliers at 3 and 4 standard deviations (%)

Indicator	Sector	2011-22					2022-23				
		Scenario			Change (in p.p.) *		Scenario			Change (in p.p.) *	
		Baseline	3 SD	4 SD	3 SD	4 SD	Baseline	3 SD	4 SD	3 SD	4 SD
Poverty rate ($2.15)	Rural	18.43	18.37	18.45	-0.05	0.02	2.83	2.81	2.83	-0.02	0.00
	Urban	10.69	10.65	10.66	-0.04	-0.02	1.14	1.12	1.13	-0.02	-0.01
	National	**16.22**	**16.16**	**16.23**	**-0.05**	**0.01**	**2.35**	**2.33**	**2.35**	**-0.02**	**0.00**
Poverty rate ($3.65)	Rural	69.05	68.95	69.07	-0.09	0.02	32.45	32.36	32.45	-0.09	0.00
	Urban	43.50	43.39	43.47	-0.11	-0.03	17.15	17.04	17.15	-0.12	0.00
	National	**61.75**	**61.65**	**61.76**	**-0.10**	**0.01**	**28.12**	**28.02**	**28.11**	**-0.10**	**0.00**
Poverty rate ($6.85)	Rural	96.03	96.01	96.03	-0.02	0.00	87.47	87.41	87.47	-0.07	0.00
	Urban	83.43	83.38	83.40	-0.05	-0.03	67.90	67.82	67.92	-0.08	0.02
	National	**92.43**	**92.40**	**92.42**	**-0.03**	**-0.01**	**81.93**	**81.86**	**81.93**	**-0.07**	**0.00**
Gini Index	Rural	24.76	24.78	24.75	0.02	0.00	22.84	22.84	22.84	0.01	0.01
	Urban	32.59	32.53	32.58	-0.05	-0.01	27.87	27.85	27.87	-0.02	0.00
	National	**28.78**	**28.77**	**28.77**	**0.00**	**0.00**	**25.51**	**25.51**	**25.52**	**0.00**	**0.00**

(*) The Gini index and change are expressed in Gini points (not in percent and p.p.). All numbers are rounded up to two decimal places.

A close examination of the outlier flags suggests that high unit values for most items are reasonable, except for abnormally high values for some items like bananas, spices, and tea leaves. While the expenditures on bananas, spices, and tea leaves seem reasonable, the lower quantities, likely due to unit mismeasurement, inflate the unit values. For instance, people often purchase a dozen bananas, while the survey records quantities in numbers. This could lead to recording issues where the respondents report quantity in dozens and not numbers. However, in most cases, very high unit values seemed reasonable, requiring no differential treatment in the data. Thus, applying a blanket rule to address a few outliers did not seem like a reasonable approach. Consequently, the final estimates do not correct for outliers in unit values.

---

[47] That is, if $\frac{ln(uv) - \overline{ln(uv)}}{sd(ln(uv))}$, lies outside the interval (a) (-3, 3), and (b) (-4, 4).

6.2 Inter-temporal deflation: CPI vs. unit values

A second sensitivity check relates to testing an alternative way to construct inter-temporal deflators. This involves leveraging the implicit price information in the survey by generating inter-temporal deflators based on inflation imputed using unit values from the survey rather than CPI-based inflation rates. Since implied inflation in the survey data may be different from the CPI-based inflation, such methodological choices can potentially affect poverty estimates.

Two alternative Fisher index-based deflators are constructed from the survey data.

1. **Consider all items:** The first index compares the cost of a basket of goods and services purchased in different years using the full item lists collected in each survey round. All items are included, irrespective of the item type or source of purchase (free items are also included).
2. **Include items comparable across survey rounds**: The second index restricts the consumption basket to items available in both 2011-12 and 2022-23 surveys.[48] This ensures estimates of inter-temporal inflation have comparable underlying consumption baskets.[49]

Once the inter-temporal deflators are constructed, they are applied to the spatially deflated expenditures at the commodity group level, to adjust expenditures from the 2011-12 survey to 2022-23 price levels. After converting 2011-12 expenditures to 2022-23 prices, the General CPI index is applied to the welfare aggregates of both years to further adjust these values to the PPP-base year (2017).

The poverty and inequality measures using these survey-based inter-temporal deflators are presented in Table 17. The final scenario presented in Section 6 employs the national general CPI index for inter-temporal deflation to convert both 2011-12 and 2022-23 welfare aggregates into 2017 prices for poverty calculations. When alternative inter-temporal deflators are applied, the poverty incidence is lower in 2011-12.

[48] Differences in the two item-lists are detailed in Annex Table A.6.1.

[49] Free items are excluded from the index to prevent distortions, as they are only reported in the 2022-23 survey and have no equivalent in 2011-12.

Table 17. Poverty and inequality rates incorporating different temporal adjustments (%)

	Sector	2011-12					2022-23				
		Scenario			Change (in p.p.) *		Scenario			Change (in p.p.) *	
		Final (CPI)	UV1	UV2	UV1	UV2	Final (CPI)	UV1	UV2	UV1	UV2
Poverty rate ($2.15)	Rural	18.43	14.17	14.22	-4.26	-4.20	2.83	2.83	2.83	0.00	0.00
	Urban	10.69	9.64	9.54	-1.05	-1.15	1.14	1.14	1.14	0.00	0.00
	National	**16.22**	**12.87**	**12.89**	**-3.34**	**-3.33**	**2.35**	**2.35**	**2.35**	**0.00**	**0.00**
Poverty rate ($3.65)	Rural	69.05	63.86	64.02	-5.19	-5.02	32.45	32.45	32.45	0.00	0.00
	Urban	43.50	41.27	41.20	-2.23	-2.30	17.15	17.15	17.15	0.00	0.00
	National	**61.75**	**57.41**	**57.50**	**-4.34**	**-4.25**	**28.12**	**28.12**	**28.12**	**0.00**	**0.00**
Poverty rate ($6.85)	Rural	96.03	95.13	95.14	-0.90	-0.89	87.47	87.47	87.47	0.00	0.00
	Urban	83.43	82.05	82.10	-1.37	-1.32	67.90	67.90	67.90	0.00	0.00
	National	**92.43**	**91.39**	**91.42**	**-1.04**	**-1.02**	**81.93**	**81.93**	**81.93**	**0.00**	**0.00**
Gini Index	Rural	24.76	24.79	24.77	0.04	0.02	22.84	22.84	22.84	0.00	0.00
	Urban	32.59	32.65	32.60	0.07	0.02	27.87	27.87	27.87	0.00	0.00
	National	**28.78**	**28.59**	**28.56**	**-0.19**	**-0.21**	**25.51**	**25.51**	**25.51**	**0.00**	**0.00**

(*) The Gini index and change are expressed in Gini points (not in percent and p.p.). All numbers are rounded up to two decimal places.

6.3 Household size from different schedules

A third sensitivity check examines the impact of using household size data from alternative modules (CSQ and DGQ) instead of FDQ in the 2022–23 CES to calculate monthly per capita expenditures. While household size can change per capita estimates, and therefore, poverty rates, the differences across modules are negligible: average household sizes are nearly identical—4.268 (FDQ), 4.271 (CSQ), and 4.270 (DGQ)—with 98.6 percent of households reporting the same size across schedules. As shown in **Table 18**, poverty and inequality estimates remain robust to these variations.

Table 18. Sensitivity of estimates to household size

	Sector	2022-23				
		Scenario			Change (in p.p.) *	
		FDQ	CSQ	DGQ	CSQ	DGQ
Poverty rate ($2.15)	Rural	2.83	2.83	2.83	0.00	0.00
	Urban	1.14	1.14	1.14	0.00	0.00
	National	**2.35**	**2.35**	**2.35**	**0.00**	**0.00**
Poverty rate ($3.65)	Rural	32.45	32.45	32.45	0.00	0.00
	Urban	17.15	17.15	17.15	0.00	0.00
	National	**28.12**	**28.12**	**28.12**	**0.00**	**0.00**
Poverty rate ($6.85)	Rural	87.47	87.47	87.47	0.00	0.00
	Urban	67.90	67.90	67.90	0.00	0.00
	National	**81.93**	**81.93**	**81.93**	**0.00**	**0.00**
Gini Index	Rural	22.84	22.84	22.84	0.00	0.00
	Urban	27.87	27.87	27.87	0.00	0.00
	National	**25.51**	**25.51**	**25.51**	**0.00**	**0.00**

(*) The Gini index and change are expressed in Gini points (not in percent and p.p.). All numbers are rounded up to two decimal places.

7. Recommendations for improving the estimation of poverty and inequality in the CES

This section provides recommendations for improving the estimation of poverty and inequality in the CES. These recommendations include both areas for improving data (from questionnaire design to data collection methods) and open research questions that could support theoretical and empirical improvements in measurement. These recommendations are not exhaustive, but they seek to strengthen the CES's ability to accurately track changes in poverty, inequality, and household welfare in line with international best practices in India.

Ex-ante and ex-post testing of comparability

Changes in survey design are a natural part of efforts to improve the completeness, accuracy, and reliability of data collection. While such changes are beneficial, they also present challenges to comparability, which must be addressed through appropriate analytical methods and data treatment.

- **Going forward, reintroducing pilot studies and bridging surveys could help assess the potential implications of revisions in advance**. Conducting pilot studies and bridging surveys is crucial for systematically evaluating key revisions to survey design and ensuring best practices are followed.
- **Ex-post analyses** that attempt to estimate the likely impact of survey design changes may offer some insights, while **evidence from other indicators** could suggest the probable direction of change.
- Finally, it would be valuable to **form a new Expert Group to tackle the unresolved comparability issues and recommend methods for maintaining comparability going forward.**

Sampling frame

The 2022-23 CES sampling frame relies on 2011 Census data for selecting villages and the latest Urban Frame Survey for selecting blocks in urban areas. Among other issues stemming from outdated information, this sampling frame is likely underestimating India's urbanization share as discussed in **Section 3.4**. While the need to update the sampling frame to a recent national population census affects all household surveys and many official statistics in India, its impact on poverty estimates from the CES remains unexplored.[50]

Updating the sample frame of the CES to a national population census is key for improving accuracy.

Understanding and addressing attrition and changing household composition

The new CES design for 2022-23 introduced several practical challenges that were not present in previous rounds, understanding the impact of which requires further research. These include

[50] Chatterjee et. al (2015) study estimates of Female Labour Force Participation Rate (FLFPR). They suggest that misclassifying urban areas as rural due to old sampling frames may explain part of the decline in rural FLFPR from 2004-05 to 2011-12.

attrition, shifting household composition, and changing respondents. Since the survey spans three months, household members may move in and out, and there could be births or deaths, potentially leading to changing household composition and individual attrition between visits. In some cases, households may have temporarily moved out for work or other reasons after the first visit, when substitution was no longer allowed. Moreover, as the 2022-23 survey now collects data on household members present during each month (with potential discrepancies across months), there is no precedent for aggregating this information into a representative month. Additionally, the respondent may change across the three questionnaires, and estimates are known to be sensitive to the choice of respondent.

While international literature may provide insights from other countries, it is essential to research the specific implications of these changes in India.

Measuring incomes and inequality

Household surveys often fail to capture top income earners accurately--the 'missing the rich' problem as described by Lustig (2019)—due to methodological and data collection challenges. Sampling frames may fail to capture high-income individuals (noncoverage error) adequately, and their low population share makes them less likely to be selected in random samples (sparseness). High unit and item nonresponse rates are common among wealthy households, often due to privacy concerns and access barriers (e.g., gated communities) (Groves and Cooper, 1998). Even when included, their income is often underreported, top-coded, or entirely missing, leading to discrepancies in inequality estimates. Empirical studies show that survey participation rates decline as income increases (see Hlasny and Verme, 2015, on the US, EU, and Egypt), and item nonresponse is correlated with income (Campos-Vazquez and Lustig, 2018). The relationship between income and survey participation affects poverty and inequality measurement (Ravallion, 2021).

Improving the valuation of housing services and durables across the income distribution is a critical component of measuring welfare, and one where changes in the CES questionnaire could yield immediate improvement. For example, improving the quality of the collected data on rents—such as reducing the high rate of missing self-assessments in rural areas or revising the wording of relevant questions—along with consistently gathering data on dwelling characteristics and ownership status (i.e., owned housing, rented housing, employer-provided housing, government subsided housing rent etc.) will be essential.

At the lower end of the distribution, a large number of schemes provide households with in-kind and cash benefits, including through subsidies and publicly provided private goods. The CES currently only allows for the estimation of the welfare gain from the PDS. These transfers, to the extent that they are progressive, remain largely unaccounted for in the estimation of poverty and inequality.

A number of techniques and correction methods could be used to improve the measurement of incomes and inequality in the CES.

- Oversampling high-income households including by using tax records or other administrative data. When administrative data, such as tax records, are accessible, a hybrid method combining survey and administrative data has gained popularity (Blanchet et al., 2022; Jenkins, 2017).

- Improving the measurement of income. The CES includes very limited information on household income. Integrating a detailed income module in the CES or implementing an independent survey to measure household incomes can help pin down the differences that should be expected between household expenditure and income-based measures of inequality. Linking survey data with tax records, social benefits, transfers, or other administrative sources can obtain more reliable income information and can be used to validate survey responses.
- Given the proliferation of social schemes and public transfers over the last decade, a concerted effort needs to be made to measure the welfare contribution of public transfers better.
- The CES does not currently allow for the estimation of the flow of services for dwellings or durables due to a lack of necessary data. Annex 5 discusses some of the areas for improvement in the survey instrument.
- Researchers and statistical agencies use various correction methods, depending on the available data, to address limitations of "missing top incomes" in the data after collection. These include parametric methods, multiple imputation, and reweighting using auxiliary data.[51]

Strengthening the CES's ability to identify the drivers and determinants of poverty reduction

The CES is much closer to a household budget survey than a comprehensive multi-topic living standards survey, which limits its ability to identify the drivers and determinants of poverty reduction. For instance, the CES does not collect detailed socio-economic information such as income (wages, business earnings, taxes, transfers) and employment, health and education, asset ownership and acquisition, and detailed dwelling characteristics. Including these drivers of welfare would allow the CES to provide a comprehensive understanding of not only the trends in household welfare but also its main determinants. Moreover, the absence of this information limits a more precise measurement of welfare in the CES. An expanded CES could, therefore, (a) lead to a better measurement of poverty and inequality, (b) allow for evidence-based policy recommendations to strengthen poverty reduction further and improve equity, and (c) enable cross-validation of other surveys and economic indicators.

An expanded CES could be an invaluable foundational tool for policy and program design and for measuring the effectiveness of public programs.

[51] When only survey data are available, there are two common approaches: replacement and reweighting. The choice between them depends on whether the observed share of top-income earners is assumed to be correct. If this share is accurate, replacement methods adjust incomes above a certain threshold, either using a parametric method (typically Pareto distribution) or a nonparametric/imputation-based approach. On the other hand, if the share is incorrect, reweighting is used to add more high-income individuals to the dataset.

References

- Anand, I. 2024. "What does the Data from the Household Consumer Expenditure Survey 2022-23 tell us?", The India Forum, July 9, 2024, available at https://www.theindiaforum.in/public-policy/household-consumption-expenditure-survey-2022-23.
- Beegle, K., J. de Weerdt, J. Friedman, and J. Gibson. 2012. "Methods of Household Consumption Measurement Through Surveys: Experimental Results from Tanzania", Journal of Development Economics 98(1) 3-18.
- Beegle, K., J. de Weerdt, and J. Gibson. 2020. "What Can We Learn from Experimenting with Survey Methods?", Annual Review of Resource Economics, 12, 431-447.
- Bhalla, S., K. Bhasin, and A. Virmani. 2022. Pandemic, Poverty and Inequality: Evidence from India. WP/22/69 IMF Working Paper. International Monetary Fund.
- Blanchet, T., I. Flores, and M. Morgan. 2022. Top incomes and inequality measurement: A guide to combining survey and tax data. The Review of Income and Wealth, 68(1), 1-30.
- Campos-Vazquez, R., & Lustig, N. 2018. Labour income inequality in Mexico: Puzzles solved and unsolved (UNU-WIDER Working Paper No. 2017/186). United Nations University World Institute for Development Economics Research (UNU-WIDER).
- Ceriani, L., S. Olivieri, and M. Ranzani. 2023. Housing, imputed rent, and household welfare. *J Econ Inequal* 21, 131–168 (2023).
- Chatterjee, U., R. Murgai, and M. Rama. 2015. Job opportunities along the rural-urban gradation and female labor force participation in India. *World Bank Policy Research Working Paper*, (7412).
- Deaton, A. and S. Zaidi. 2002. Guidelines for Constructing Consumption Aggregates for Welfare Analysis. LSMS Working Paper No. 135. http://hdl.handle.net/10986/14101.
- Deaton, A. and V. Kozel. 2005. Data and Dogma: The Great Indian Poverty Debate. <u>World Bank Research Observer.</u> 20(2), 177-199.
- Groves, R. M., and M. P. Couper. 1998. Nonresponse in household interview surveys. John Wiley & Sons.
- Hlasny, V. and P. Verme. 2015. Top incomes and the measurement of inequality: A comparative analysis of correction methods using Egyptian, EU, and US survey data. Paper presented at the 6th meeting of the Society for the Study of Economic Inequality (ECINEQ), Luxembourg.
- Jenkins, S. P. 2017. Pareto models, top incomes, and recent trends in UK income inequality. Economica, 84(334), 261-289.
- Jolliffe, D. 2001. Measuring Absolute and Relative Poverty: The Sensitivity of Estimated Household Consumption to Survey Design. <u>Journal of Economic and Social Measurement</u> 27(1-2), 1-23.
- Lanjouw, J.O. and P. Lanjouw. 2001. "How to Compare Apples and Oranges: Poverty Measurement Based on Different Definitions of Consumption", <u>Review of Income and Wealth</u>, 47(1), 25-42.
- Lustig, N. 2019. The "missing rich" in household surveys: Causes and correction approaches (Stone Center on Socio-Economic Inequality Working Paper No. 08). Stone Center on Socio-Economic Inequality, City University of New York.
- Mancini, G. and G. Vecchi. 2022. On the Construction of a Consumption Aggregate for Inequality and Poverty Analysis. World Bank Group. Washington, DC.
- Ministry of Statistics and Programme Implementation (MoSPI). 2024a. Factsheet on Household Consumption Expenditure: 2022-23.

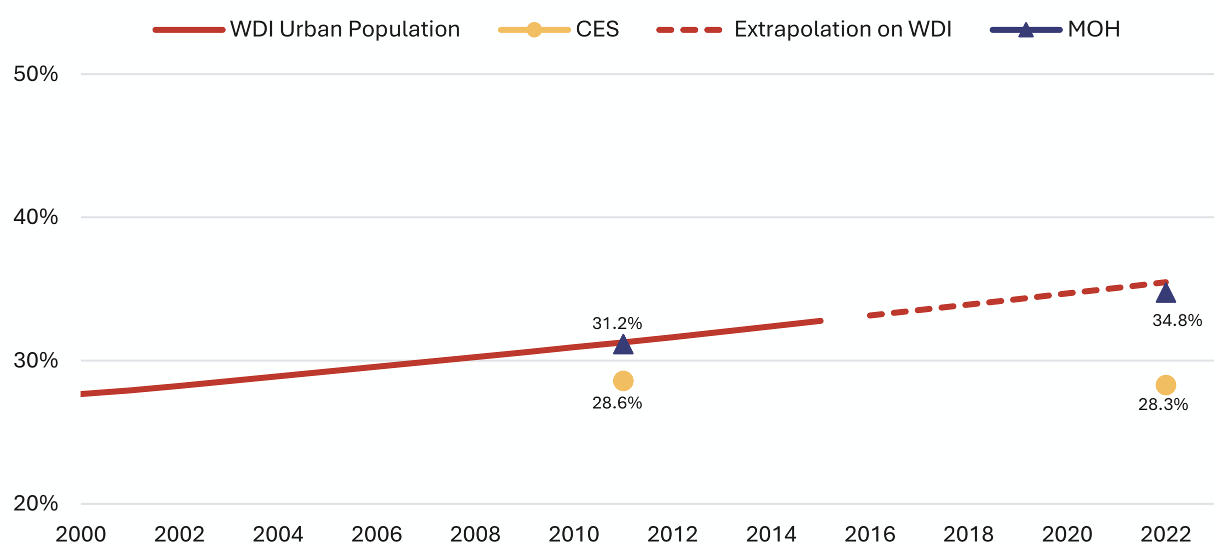
- _________. 2024b. Survey on Household Consumption Expenditure: 2022-23. NSS Office. Report no. 591.
- Narayan, A. and R. Murgai. 2016. Looking Back on Two Decades of Poverty and Well-Being in India. World Bank Policy Research Working Paper No. 7626. World Bank.
- OECD. 2024. *Housing, Wealth Accumulation, and Wealth Distribution: Evidence and Stylized Facts*. OECD Economics Department Working Papers No. 1588.
- Ravallion, M. 1998. Poverty Lines in Theory and Practice. LSMS Working Paper No. 133. World Bank, Washington, DC.
- Ravallion, M. 2021. Missing top income recipients (NBER Working Paper No. 28890). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w28890.
- Sen, A. 2000. "Estimates of Consumer Expenditure and Its Distribution: Statical Priorities after NSS 55th Round". Economic and Political Weekly. Vol. 35, No. 51, 16th December 2000.
- Sen, A. and Himanshu. 2004. "Poverty and Inequality in India-I". Economic and Political Weekly. Vol. 39, No. 38, 18th September 2004.
- Sinha Roy, S. and R. Van Der Weide. 2022. Poverty in India Has Declined over the Last Decade But Not As Much As Previously Thought. Policy Research Working Paper No. 9994. World Bank. Washington, DC.
- Singh, R. 2023. Do the wealthy underreport their income? Using general election filings to study the income-wealth relationship in India (World Inequality Lab Working Paper No. 2023/01). World Inequality Lab. Retrieved from https://wid.world.
- World Bank. 2024a. "Measuring Welfare When It Matters Most: A Typology of Approaches for Real-time Monitoring. Washington, DC.
- _________. 2024b. "Population Estimates and Projections for the World Bank: Technical Note." Edition December 2024. Available at https://datanalytics.worldbank.org/Population-Technical-Note/ (Accessed Feb 24th, 2025).
- _________. 2025a. Poverty and Inequality Platform. https://pip.worldbank.org/country-profiles/IND (Accessed Feb 18th, 2025).
- _________. 2025b. ICP 2021. DataBank. https://databank.worldbank.org/source/icp-2021 (Accessed Feb 18th, 2025).
- _________. 2025c. ICP 20217. DataBank. https://databank.worldbank.org/source/icp-2017 (Accessed Feb 18th, 2025).

Annex 1. Additional survey sampling details

Table A.1.1. Changes in selection criteria for the second stage strata (SSS)

Second stage substrata	2011-12		2022-23	
	Rural	Urban	Rural	Urban
SSS1	Relatively affluent households; based on a number of criteria such as possession of certain durables or assets	Households having MPCE of top 10 percent of urban population (MPCE>B)	Households with land possessed > X, where X is such that 5 percent of the households had land possessed more that X as per NSS 77th round SAS survey determined at state/UT level	Households owned one or more four-wheeler cars for noncommercial use as on date of listing whose combined purchase price was > INR 10 lakhs
SSS2	Of the remaining, households having principal earning from non-agricultural activity	Households having MPCE of middle 60 percent of urban population (A ≤ MPCE ≤ B); where A and B are estimated for each NSS-state region from NSS 66th round (2009-10)	Households with land possessed > Y and less than or equal to X, where Y is such that 20 percent of the households had land possessed more that Yas per NSS 77th round SAS survey determined at state/UT level	Households owned one or more four-wheeler cars for noncommercial use as on date of listing whose combined purchase price was<= INR 10 lakhs
SSS3	Other households	Households having MPCE of bottom 30 percent of urban population (MPCE < A)	Remaining households	Remaining households

Figure A.1.1. Share of urban population implied in CES survey weights vs. other sources



Annex 2. Item differences across surveys

Table A.2.1. List of items not common between 2011-12 and 2022-23 CES

Product Group	Items (Item code)	Share of households that purchased items in 2022-23 CES	Share of item expenditure in the total MPCE in 2022-23
(a) New items introduced in CES 2022-23 or split of items from those in 2011-12			
Milk and milk products	Paneer (3)	13.0%	0.3%
	Prepared Sweets (4)	15.3%	0.4%
	Cheese (gm) (5)	0.2%	0.0%
	Lassi (6)	2.3%	0.0%
	Butter Milk (7)	4.0%	0.0%
Processed foods	Cooked Meals Purchased at Subsidized Rate (No.) (76)	0.6%	0.0%
	Biscuits (13)	88.6%	1.0%
	Breakfast Cereals (Oats, Cornflakes, Muesli Etc.) (12)	4.4%	0.1%
	Health Supplements (Protein Powder, Probiotic Tablet and Drinks, Chawanprash, Etc.) (14)	3.6%	0.1%
	Noodles (15)	23.6%	0.2%
Spices	Poppy Seeds (Gm) (263)	3.4%	0.0%
Beverages	Soda Drinks (Litre) (11)	5.3%	0.1%
Clothing	Face Mask, Gloves (30)	21.8%	0.0%
	Shirts (no.) (31)	89.2%	0.6%
	Western Suit Set (No.) (32)	2.5%	0.0%
	Sherwani, Lenhnga, Gown (No.) (33)	9.5%	0.1%
	Track Suit, Track Pant, Joggers (No.) (34)	23.9%	0.1%
Sports goods	Exercise Equipment: Treadmill, Fitness Cycle, etc. (36)	0.3%	0.0%
	Weighing Machine (37)	0.4%	0.0%
	Sports Goods (Badminton Racket, Shuttle Cock, Football, Cricket Bat, etc.),toys (432)	36.5%	0.1%
	Other Sports Goods (97)	12.1%	0.0%
Cooking and other household appliances	Electric Air Heaters & Blowers (43)	0.9%	0.0%
	Air Cooler (44)	13.4%	0.1%
	Air Purifier (45)	0.2%	0.0%
	Geysers & Hot Water Systems Sterilizers (46)	2.1%	0.0%
	Ventilation System such as Exhaust, Chimney, Etc. (47)	0.4%	0.0%
	Mixer/Grinder, Juicer, Food Processor (48)	17.2%	0.0%

(a) New items introduced in CES 2022-23 or split of items from those in 2011-12			
Product Group	**Items (Item code)**	**Share of households that purchased items in 2022-23 CES**	**Share of item expenditure in the total MPCE in 2022-23**
Crockery and utensils	Water Bottle/Feeding Bottle (50)	34.5%	0.0%
Medical equipment	Wheelchair (40)	0.1%	0.0%
	Massagers (41)	0.6%	0.0%
	Nebulizer (42)	1.2%	0.0%
Other personal goods	Headphone, Earphone, Ear Pod, Airpod, Bluetooth Devices/Speakers Etc. (35)	30.7%	0.1%
Consumer services excluding conveyance	Watchmen/Security Guard, driver (27)	0.7%	0.0%
Other household consumables	Floor Cleaner, Acid, Toilet Cleaner (465)	57.3%	0.2%
	Led Bulb, Cfl Bulbs (23)	34.0%	0.2%
PDS-subsidized*	Pulses (158)	8.4%	0.0%
	Salt (178)	5.3%	0.0%
	Edible Oil (188)	5.5%	0.0%
Free items*	Rice (61)	44.6%	0.0%
	Wheat/Atta (62)	25.6%	0.0%
	Coarse grains (70)	0.9%	0.0%
	Pulses (71)	0.3%	0.0%
	Gram (72)	0.5%	0.0%
	Salt (73)	3.2%	0.0%
	Sugar (74)	0.7%	0.0%
	Edible Oil (75)	0.9%	0.0%
(b) Items in 2011-12 CES that are removed from 2022-23 CES			
Product Group	**Items (Item code)**	**Share of households that purchased items in 2011-12 CES**	**Share of item expenditure in the total MPCE in 2011-12**
Fuel and light	Coke (330)	0.7%	0.0%
Conveyance	Horse cart fare (506)	1.1%	0.0%
	Lubricants and other fuels for vehicles (511)	4.3%	0.0%

Notes: * The CES registers zero expenditures for the consumption of free items, and captures expenditures valued at government-subsidized prices for PDS-subsidized items. These expenditures do not necessarily capture the contribution of those items for household welfare (which is discussed and re-valued in section 4.5.1).

Annex 3. Definition of commodity groups

Table A.3.1. Classification of items in different commodity groups

Commodity group	Items (subtotal)	Fisher index constructed
Cereals	Cereals	Yes
Pulses	Pulses & Pulse Products	Yes
Milk	Milk And Milk Products	Yes
Oil	Edible Oil	Yes
Egg, fish, meat	Egg, Fish & Meat	Yes
Vegetables	Vegetables	Yes
Fresh fruits	Fruits (Fresh)	Yes
Dry fruits	Fruits (Dry)	Yes
Salt and sugar	Salt & Sugar	Yes
Spices	Spices	Yes
	Packaged Processed Food	
Other food	Beverages	Yes
	Served Processed Food	
	Intoxicants	Yes
Intoxicants	Tobacco	Yes
	Pan	Yes
Fuel and light	Fuel And Light	Yes
Clothing	Clothing	
	Bedding, Etc.	Yes
Footwear	Footwear	Yes
Education	Education	No
Health (Non-institutional)	Medical - Non-Hospitalization	No
Health (Institutional)	Medical - Hospitalization	No
Transports	Personal Transport Equipment	No
	Furniture & Fixtures	
	Other Household Consumables	
	Residential Building, Land and Other Durables	
Household goods and services	Crockery & Utensils	No
	Cooking & Other Household Appliances	
	Medical Equipment	
	Consumer Services Excluding Conveyance	
	Goods For Recreation	
Recreation	Sport Goods	No
	Entertainment	
	Jewelry & Ornaments	
Personal care and effects	Other Personal Goods	No
	Toilet Articles	
Rent	Rent	No
Cesses and taxes	Other Consumer Taxes & Cesses	No
Conveyance	Conveyance	No

Annex 4. Construction of spatial price deflators

Ensuring household consumption is comparable across India requires adjusting for spatial differences in the cost of living. Given significant price variations across states and between rural and urban areas, state-sector-level price indices are necessary to standardize expenditures. This section outlines the methodology used to construct spatial price deflators, leveraging Fisher indices based on unit values, following the Tendulkar Committee's (2009) approach. These deflators adjust household expenditures to a common price level, enabling more accurate welfare comparisons across regions.

The Fisher index F_{sG} for commodity group *G* and state-sector *s* was constructed relative to all-India. Below are the steps followed in the calculation of Fisher Indices.

Notation.

h:	household *h*
s:	state-sector, where sector refers to either rural or urban
I:	All-India
j:	item *j*
G:	commodity group G
pwt_h:	sample population weights for household *h*
pop_s:	census population for state-sector *s*
V:	monetary value of a bundle
$\tilde{p}$:	median unit value (proxy to price)
$\bar{q}$:	average quantity per capita

Step-by-step instructions

- In 2011-12, map the state codes to 2022-23 to make them comparable: create states such as Telangana, and merge union territories such as Daman & Diu with Dadra & Nagar Haveli.[52]
- Map items in CES 2011-12 and 2022-23 to respective CPI subgroups.
- Append all consumption blocks to create a list of all the items consumed by the households with the value of consumption and quantity consumed.
- Classify items into commodity groups – Since unit information is not available for all items; commodity groups 1-15 were used to construct the spatial price deflators following the approach adopted by Tendulkar Committee 2009). These groups include *cereals; pulses; milk; oil; egg, fish, meat; vegetables; fresh fruits; dry fruits; salt and sugar; spices; other food; intoxicants; fuel and light; clothing; footwears*.
- Exclude sub-totals from the data: In addition to the subtotals of the survey bocks, some grains also have subtotals (referenced with * at the bottom of the questionnaires.
- Drop items where quantity was not collected.
- Combine PDS and free items with corresponding non-PDS items, for example combine rice-PDS, rice-free, and rice-other sources. This applies to four items in 2011-12 – rice, wheat, sugar, kerosene, and eight items in 2022-23 – coarse grains, rice, wheat, grams, pulses, sugar, salt, and edible oil.
- Assign a notional value of one rupee to free items so that free items are not dropped from the unit values.

[52] Exact details are available in the accompanying reproducibility package.

- Arrange data in such a way that each household has an entry for every item irrespective of the consumption. That is all households have the same list of items.
- For each item, count the number of households consuming that item within each state-sector (referred to as non-zero observations).
- Calculate item-level unit value as the value divided by the quantity.
- Replace the unit value as missing if an item has less than six (non-zero or non-missing) observations within a state-sector i.e. if an item is consumed by less than six households in that state-sector, that item is excluded from the further calculations.
- Exclude the four Union Territories (UTs) due to insufficient data - Chandigarh, Daman & Diu, Dadar & Nagar Haveli, Lakshadweep, and Andaman & Nicobar Islands.
- Perform the calculations below to calculate the Fisher Index for each state-sector-commodity group. These calculations are for each of the 15 commodity groups for every state-sector and yields a total of 960 spatial deflators.

Median unit value for each state-sector

For each item belonging to the 15 commodity groups, the unit value for each household that consumed the item is calculated. Quantity and expenditures are replaced by a missing value for households that did not consume the item.

$$p_h = \frac{Temporally\ delfated\ expenditure\ on\ item\ j}{Quantity\ consumed}$$

Calculate median unit value for each item at the state-sector level $(\widetilde{p_{JS}})$ using survey population weights. Replace $(\widetilde{p_{JS}})$ to missing if less than 6 households within the state-sector consumed the item.

Average quantity per capita for each state-sector

First, replace quantity to zero for households that did not consume the item. Calculate quantity per capita for each household.

$$q_h = \frac{Quantity\ consumed}{Household\ size}$$

Calculate average per capita quantity consumed for each item at the state-sector level $(\overline{q_{JS}})$ using survey population weights.

Unit value at all-India level

The unit value for each item at all-India level is calculated as the product of average quantity consumed in state-sector $(\overline{q_{JS}})$, median price in state-sector $(\widetilde{p_{JS}})$, census population of state-sector (pop_s), and then summing these values across states. This total is then divided by the product of the average quantity consumed in state-sector and population of the state-sector, summed across states.

$$\widetilde{p_{JI}} = \frac{\sum_{s=1}^{S} \widetilde{p_{JS}} * \overline{q_{JS}} * pop_s}{\sum_{s=1}^{S} \overline{q_{JS}} * pop_s}$$

Quantity per capita at all-India level

The quantity per capita for each item at all-India level is calculated as the product of the average quantity consumed in state-sector and population of the state-sector, summed across states and then divided by all-India population.

$$\overline{q_{JI}} = \frac{\sum_{s=1}^{S} \overline{q_{JS}} * pop_s}{\sum_{s=1}^{S} pop_s}$$

State-sector bundle valued at state-sector prices

For each item *j*, the consumption value of a state-sector bundle at state-sector prices (V_{ssG}) is the product of the median unit value ($\widetilde{p_{JS}}$) and the average quantity consumed per capita in that state-sector ($\overline{q_{JS}}$). Since the bundles are defined at the commodity group level, the value of the bundle for a commodity group is the aggregation of all these products for the items belonging to that commodity group *G*.

$$V_{ssG} = \sum_{j \in G} \widetilde{p_{JS}} * \overline{q_{JS}}$$

State-sector bundle valued at all-India prices

For each item *j,* the consumption value of state-sector bundle at all-India prices (V_{sIj}) is the product of the All- India median unit value ($\widetilde{p_{JI}}$) and the average quantity consumed per capita in the state-sector ($\overline{q_{JS}}$), aggregated at the commodity group level.

$$V_{sIG} = \sum_{j \in G} \widetilde{p_{JI}} * \overline{q_{JS}}$$

All-India bundle valued at state-sector prices

For each item *j*, the value of all-India bundle at state-sector prices ($V_{Is)}$ is the product of state median price $\widetilde{p_{JS}}$ and average quantity consumed per capita at the all-India level ($\overline{q_{JI}}$ aggregated to the commodity group level.

$$V_{IsG} = \sum_{j \in G} \widetilde{p_{JS}} * \overline{q_{JI}}$$

All-India bundle at All-India prices

For each item *j*, the consumption value of all-India bundle at all-India prices ($V_{IIG)}$ is the product of the all-India median unit value $\widetilde{p_{JI}}$ and the average quantity consumed per capita at the all-India level $\overline{q_{JI}}$, aggregated to the commodity group level.

$$V_{IIG} = \sum_{j \in G} \widetilde{p_{JI}} * \overline{q_{JI}}$$

Fisher index

The Fisher Index is the geometric average of a Laspeyres Index and a Paasche Index.

Paasche index captures the relative value of state-sector bundle valued at state-sector to the value of state-sector bundle valued at all-India prices.

$$P_{sG} = \frac{V_{ssG}}{V_{sIG}}$$

Laspeyres index captures the relative value of All-India bundle valued at state-sector prices to the value of all-India bundle valued at all-India prices.

$$L_{sG} = \frac{V_{IsG}}{V_{IIG}}$$

Fisher index: geometric mean of Paasche and Laspeyres index.

$$F_{sG} = \sqrt{P_{sG} L_{sG}}$$

These indices are computed for each state-sector, and separately for the 15 commodity groups *G*.

Implicit deflators at the HH-level

The intra-temporally deflated welfare aggregate (X_{sG}) is divided by the Fisher indices (F_{sG}) for these 15 commodity groups to get the unadjusted spatially deflated welfare aggregate $\widehat{(x_{sG})}$.

$$\hat{x}_{sG} = \frac{X_{sG}}{F_{sG}}$$

In the next step, X_{sG} and $\hat{x}_{sG}$ are aggregated across the 15 commodity groups for each household, creating an implicit deflator S_h, as follows.

$$S_h = \frac{\sum_{g=1}^{15} X_{sG}}{\sum_{g=1}^{15} \hat{x}_{sG}}$$

The next step was to normalize this spatial deflator at the household level S_h with respect to all-India, because the individual Fisher Index is not guaranteed to add up to one at the national level. For this, the all-India weighted average S_I of the spatial deflators is computed, applying the survey population weights.

$$S_I = \frac{\sum_{h=1}^{H} (S_h * pwt_h)}{\sum_{h=1}^{H} pwt_h}$$

And the normalized version s_h of the spatial deflators for a given household is:

$$s_h = \frac{S_h}{S_I}$$

So that the sum of the normalized spatial deflators adds to 1 at the national level.

$$\sum_{h=1}^{H} s_h = 1$$

Real welfare aggregate (MPCE) construction

The spatial adjustment is made on the previously within-year adjusted expenditures using the adjusted spatial deflator.

$$\hat{\hat{x}}_G = \frac{x_G}{s_h}$$

For each household, compute the share of nominal expenditure X_G of the 15 commodity groups, before deflation.

$$\emptyset_1 = \frac{\sum_{G=1}^{15} X_G}{\sum_{G=1}^{25} X_G}$$

Aggregate the spatially adjusted expenditure $\hat{x}_{sg}$ of the 15 commodity groups and divide by their share of nominal expenditure $\emptyset_1$. The underlying assumption underlying this augmentation is that the share of expenditure of the 15 commodity groups is the same in nominal and real terms.

$$Real\ MPCE = \frac{\sum_1^{15} \hat{\hat{x}}_{sG}}{\emptyset_1}$$

Annex 5. Estimation of consumption flow from durables

As discussed, the choice of how durable assets are treated has important implications for the welfare aggregate. Since durable goods last over many years, it is not the *purchase* of assets but their use that contributes to the welfare of a household. Therefore, international best practices recommend against including the purchase of durable assets, and instead the value of services from using durable assets over the reference period should be included in the welfare aggregate (Deaton and Zaidi, 2002).

One way to estimate the consumption flow from services is the *rental equivalent* approach. To estimate the utility derived from using an asset over the reference year, one can use the current rental value of the asset of a specific age and type, as if it were rented from the market (Mancini and Vecchi, 2021). While this is a simple solution, rental markets may not exist for many of the durable assets owned by the household and in other cases markets may be thin, preventing from relying on prevailing market rents for assets to estimate the consumption flow from durables.

A second method, recommended as the best practice across countries, is the *user cost approach* that estimates a measure of the flow of services from the durable asset. Typically, a household that owns a durable asset has two choices at the beginning of the reference period: (i) to sell the asset, or, or (ii) to use the durable asset. If the household sells the asset (value p_t), it can save this amount such that the household receives a return, $p_t(1+i_t)$, at the end of the year, where i_t is the general nominal interest rate. If the household chooses to use the asset, it can sell the asset at the end of the year at its depreciated value adjusted for inflation $p_t(1+\pi_t)(1-\delta_t)$, where π_t is the inflation in general price levels during the year, and δ_t is the annual depreciate rate on the value of the year. The consumption flow from durables is estimated as the difference between the value of two options, the price a household is willing to pay to keep and use the asset by not selling it at the beginning of the period.

While the nominal interest rates and inflation rates are typically available from external sources, the current value of the asset and its depreciation rate are estimated based on the information collected in the survey.[53] At a minimum, the survey must collect the following information from a household to arrive at an accurate consumption flow from durable assets as per the recommended best practices.

1. The number of item units of each durable asset owned by the household during the reference period.
2. The current sale value of all units of the assets owned by the household.[54]

[53] Various methods for estimating the depreciation rates are discussed in Mancini and Vecchi (2021).

[54] A Living Standards Measurement Study (LSMS) type module typically asks a household, "If a household were to sell the unit today, what value of the item will you get in the market?" While this information is ideal if recorded for each unit of the asset separately, some countries ask this question for all units of an asset together (such as in Nepal Living Standards Survey IV).

3. The year of purchase (to estimate the age of an asset) and the purchase price of each item owned of an asset owned by the household.[55]

The age of the asset, along with its purchase and sale value, is used to estimate the depreciation rate of the asset over the year. The stock of an asset owned during the, their current value, asset-specific depreciation rate and market-based interest rate and inflation rate are together used for estimating the user cost of each asset.

In CES 2022-23, while the purchase prices of first-hand and second-hand items, as well as the associated maintenance costs during the year, are available, data on the stock of each asset and their current value is not available. Thus, to preserve welfare ranking of the households, the purchase value of assets is excluded from the consumption aggregate. Since durables are a large share of the total consumption, especially, at the top of the distribution, this exclusion reduces the spread of the welfare distribution by construction. **Figure A.5.1** provides an example module from the recent Nepal Living Standards Survey -IV (2021-22) that collects information on the stock of asset (6.06), purchase value of the most recent purchase (6.05) and expected sale value (question 6.07), that allow for estimating the use value from consumption flow of durables.

Figure A.5.1. Example module for durable assets

Section 6 Non-Food Expenditures & Inventory of Durable Goods — Part C: Inventory of Durable Goods

CODE	(6.03c) Does your household own any of the following items? ENTER 1 IF THE ANSWER IS YES, AND ENTER 2 IF NO. YES 1 / NO 2 ▼ NEXT ITEM		(6.04) How many ..[ITEM].. does your household own?	(6.05) acquire ..[ITEM] and how much did you pay for it or what would be the monetary value of the gift? (IF YEARS IS 0 GO TO 6.07) IF MORE THAN ONE ITEM OWNED, ASK ABOUT THE LATEST ITEM BOUGHT OR RECEIVED AND		(6.06) If you had sold this ..[ITEM] a year ago, how much money would you receive for it? IF MORE THAN ONE ITEM OWNED, ASK ABOUT TOTAL NUMBER AND VALUE OF ALL ITEMS		(6.07) If you wanted to sell this ..[ITEM].. today, how much money would you receive for it?
			NUMBER	a. YEARS	b. RUPEES	a. NO OF ITEMS	b. RUPEES	RUPEES
501	Radio / Tape / CD/VCD/DVD player							
502	Camera (Still / Movie)							
503	Bicycle							
504	Motorcycle / Scooter							
505	Motor car, jeep, van, etc.							
506	Refrigerator or Freezer							
507	Microwave oven							
508	Geyser (Gas / Electricity)							
509	Washing machine							
510	Fans							
511	Heater (gas/kerosene/electric)							
512	Television							
513	Air conditioner/Cooler							
514	Vacuum cleaner							
515	Inverter							
516	Solar panel (for electricity)							
517	Solar heater							
518	Iron							
519	Telephone sets / Mobile phone / Tablet							
520	Sewing machine							
521	Computer / Printer							
522	Wrist watch							
523	Furniture							
524	Kitchen utensils							
500	TOTAL							

Source: Nepal Living Standards Survey 2021-22 (NLSS-IV).

[55] Recording the purchase history of each unit of an asset may not always be feasible or ideal due to expected recall biases in values or survey length. In some countries, therefore, the survey records information on the purchase year and purchase price for the most recently purchased unit of an asset type (such as Nepal Living Standards Survey III and IV).

Annex 6. Experimental hedonic rents model[56]

An experimental hedonic model was conducted for 2022-23, to assess the potential measurement impacts of imputing housing services across the distribution. Hedonic models for housing typically rely on variables of physical characteristics, location attributes, and market conditions.[57] As a significant improvement in the questionnaire, the CES 2022-23 collected information on basic dwelling characteristics.[58] Replicating a similar hedonic model for the 2011-12 CES, nonetheless, was not feasible due to the data limitations discussed above—i.e., missing dwelling characteristics and missing self-reported (imputed) value for rural homeowners.

Following the literature, a standard hedonic rents equation was estimated using an ordinary least squares (OLS) regression. The dependent variable was the logarithm of the rental expenses. The explanatory variables included all available dwelling characteristics collected in the survey, basic socio-demographic characteristics of the household, as well as state, sector, SSS, and survey month fixed effects. The objective of the model was fully predictive. Following convention, Duan's correction was applied to retrieve the predicted value (in INR) of housing services.

To expand the training sample, the dependent variable of rental value was built by combining rental (actual) expenses reported by renters and self-reported (imputed) values for owners. The literature suggests the feasibility of using self-reported rent by homeowners to enlarge the number of rental observations, which as shown in Table 9 would double (2011) and thrice (2022) the training sample. Following Ceriani et al. (2023), a propensity score matching technique is first used to identify the sub-sample of "comparable tenants and homeowners" that can more accurately help estimate rental values among homeowners. Homeowners who cannot be matched (lying outside the common support) are discarded from the training sample.

The hedonic prediction was used for any households not reporting positive rental (actual) expenses or reporting an imputed value of rents (among homeowners). Ultimately, 78 percent of households received a value of housing based on the hedonic model prediction. Renters and matched homeowners with self-imputed rents keep those values reported in the survey. On average, the housing services would account for 12.9 percent of (pre-housing) welfare in rural areas and 22.3 percent in urban areas (**Figure A.6.1**). These housing shares are relatively similar (uniform) across all consumption levels. In rural areas, the bottom 5 percent observes the highest share (16.5 percent), whereas the highest housing shares (above 23 percent) concentrate among better-off urban households (except the urban top 5 percent).

---

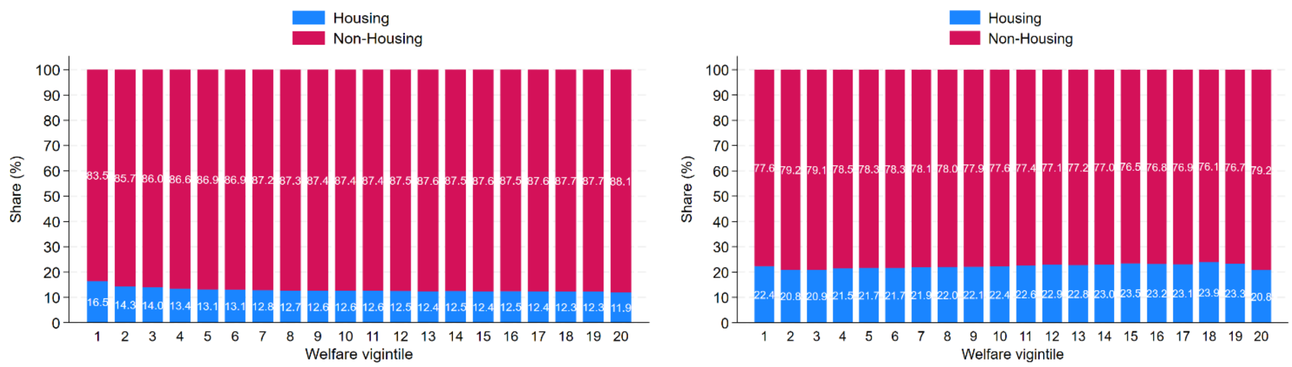
[56] The team is grateful to Sergio Olivieri and Jaime Fernandez Romero for their technical advice on the construction of hedonic rent models.

[57] Key physical characteristics include square footage, number of bedrooms and bathrooms, lot size, property age, structure type (e.g., single-family home, apartment, townhouse), and specific amenities (e.g., garage, fireplace, air conditioning). Location attributes cover neighborhood quality, proximity to schools, parks, and amenities, access to public transportation, crime rates, and environmental factors like air quality and noise levels. Market conditions reflect relevant local economic indicators.

[58] The CES 2022-23 questionnaire included dwelling characteristics: construction materials for floors, roofs and walls; primary source of drinking water; time taken by the household for a single trip to reach the source of water; and access to latrine and type of latrine; primary cooking fuel, primary energy source for lighting, and total area owned.

Figure A.6.1. Distribution of housing and non-housing shares by welfare vigintiles
(share of housing value over pre-housing welfare)

a. Rural b. Urban



Note: The plot shows the average share of predicted housing values (over the welfare aggregate) for each vigintile of the pre-housing real welfare aggregate per capita (after all adjustments proposed in section 4). Population weights are applied.

Introducing housing values based on the hedonic model would reduce the estimated poverty rates. The reduction in extreme poverty rates would be approximately 1.4 p.p. at the national level which represents a significant 60 percent reduction. The reduction in extreme poverty is more pronounced in urban areas (over 70 percent reduction) but also significant among rural households (approx. 60 percent) (**Table A.6.1).** The LMIC headcount poverty rate would decrease by 10.4 p.p. at the national level, which represents a 37 percent reduction (approx. 32 percent in rural areas and 50 percent in urban areas).

Table A.6.1. Potential effects of including predicted housing values, 2022-23 (in %)

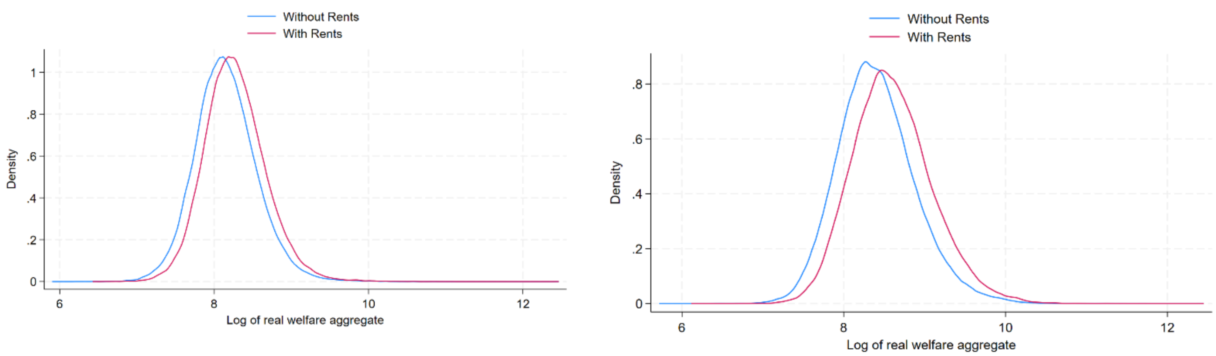
Indicator		National		Rural		Urban	
		Without housing	With housing	Without housing	With housing	Without housing	With housing
Poverty headcount	IPL 2.15	2.3	0.9	2.8	1.2	1.1	0.3
	IPL 3.65	28.1	17.7	32.5	21.3	17.2	8.6
	IPL 6.85	81.9	72.7	87.5	80.9	67.9	51.9
Inequality	Gini	25.5	26.4	22.8	22.7	27.9	28.5

Notes: All IPLs expressed in 2017 PPP. Gini is not expressed in percent. "Without housing" refers to estimates based on the real welfare aggregate, where the value of rent expenses among renters had to be excluded. "With housing" estimates re-calculate poverty and inequality after adding the hedonic model predictions to account for a value of housing services across all households in the sample. All results are preliminary and subject to change as methodological improvements become available.

However, the prediction of similar shares of housing expenses across households results in a moderate increase in measured inequality. The shifts in the welfare distribution are illustrated in **Figure A.6.2,** while negligible changes in the Lorenz curves after incorporating housing values are shown in **Figure A.6.3.** After incorporating the experimental hedonic model results, the Gini index would increase from 25.5 to 26.4, driven by increased urban inequality.

Figure A.6.2. Changes in welfare distribution, with housing value prediction

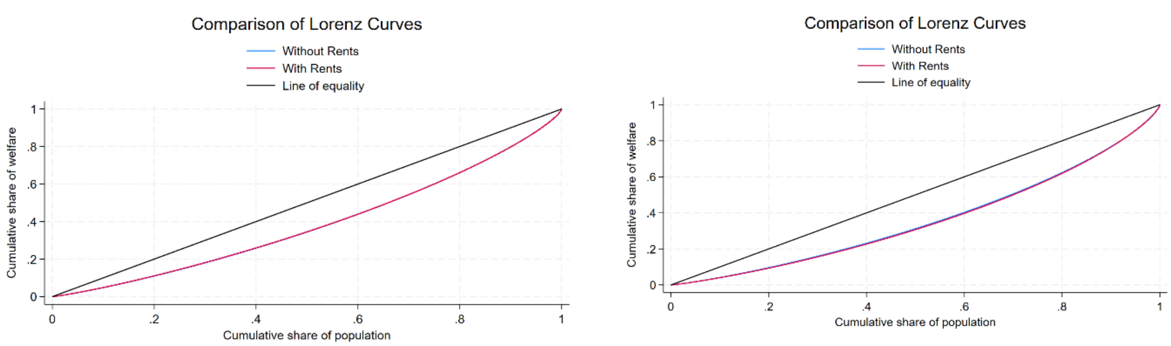
a. Rural b. Urban



Note: Population weights applied.

Figure A.6.3. Changes in Lorenz Curves, with housing value prediction

a. Rural b. Urban



Note: Population weights applied.

This prediction, nonetheless, faces the data limitations explained above, as well as methodological challenges.[59] One key limitation is the absence of critical variables for accurate predictions. Hedonic models for housing typically rely on variables including physical characteristics, location attributes, and market conditions.[60] The CES lacks relevant variables to differentiate across dwelling sizes and qualities (e.g., number of rooms). Moreover, the model cannot realistically account for geospatial location—a known major predictor of rental market values.[61] While this challenge is not unique to India's CES, improving the predictive power of the

[59] Methodologically, the linear nature of the model might not capture a complex non-linear relationship between rent and dwelling characteristics

[60] Key physical characteristics include square footage, number of bedrooms and bathrooms, lot size, property age, structure type (e.g., single-family home, apartment, townhouse), and specific amenities (e.g., garage, fireplace, air conditioning). Location attributes cover neighborhood quality, proximity to schools, parks, and amenities, access to public transportation, crime rates, and environmental factors like air quality and noise levels. Market conditions reflect relevant local economic indicators.

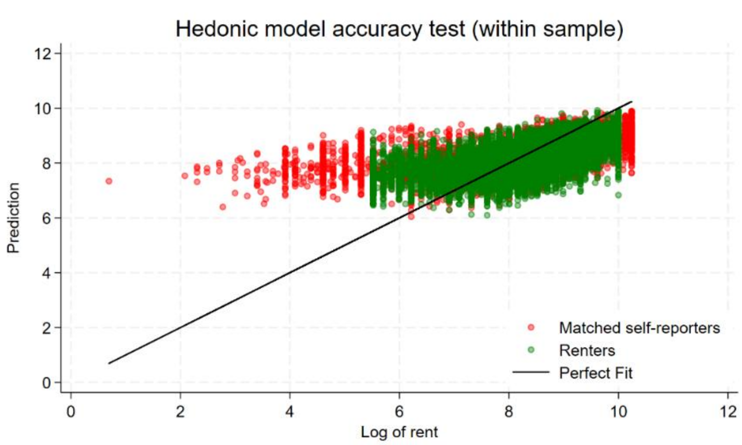
[61] The model controls for state and sector fixed effects, as well as for SSS. However, any further geographic disaggregation is not possible in the CES.

hedonic model will require collecting additional data to reflect housing location, quality and desirability.[62]

Skewness in the available rental data likely contributes to an overestimation of housing values at the bottom of the distribution and for rural dwellings. As noted above, the likelihood of reporting actual rents in the CES is low, but it is significantly higher among urban and relatively wealthier households (**Figure 2**), introducing both sectoral and socioeconomic bias. To expand the number of observations, the training sample of the model includes both the actual rents paid by renters, as well as the (imputed) rents self-reported by assimilable non-renters. However, the high non-response rate in this subset skews the sample, maintaining a similar urban bias. These concerns are further supported by an in-sample accuracy test, which suggests that the hedonic model overestimates rent values at the lower end of the renters' distribution (**Figure A.6.4**).

Ultimately, the impact of including the value of housing services across all Indian households remains an empirical question. With more detailed survey data and accurate hedonic rental imputations, it would be possible to capture rental services for both renters and homeowners consistently. However, the effect on inequality will depend on the relative significance of housing services in the overall welfare measure across different income groups. International studies have identified varied distributional patterns, and given the current data, it is not possible to predict the distributional pattern of housing values for India. If housing services constitute a larger share of welfare for poorer households compared to wealthier ones, the Gini index would likely decrease after incorporating the housing component. Conversely, the Gini index would likely increase if wealthier households allocated a greater proportion of their welfare or expenditures to housing.

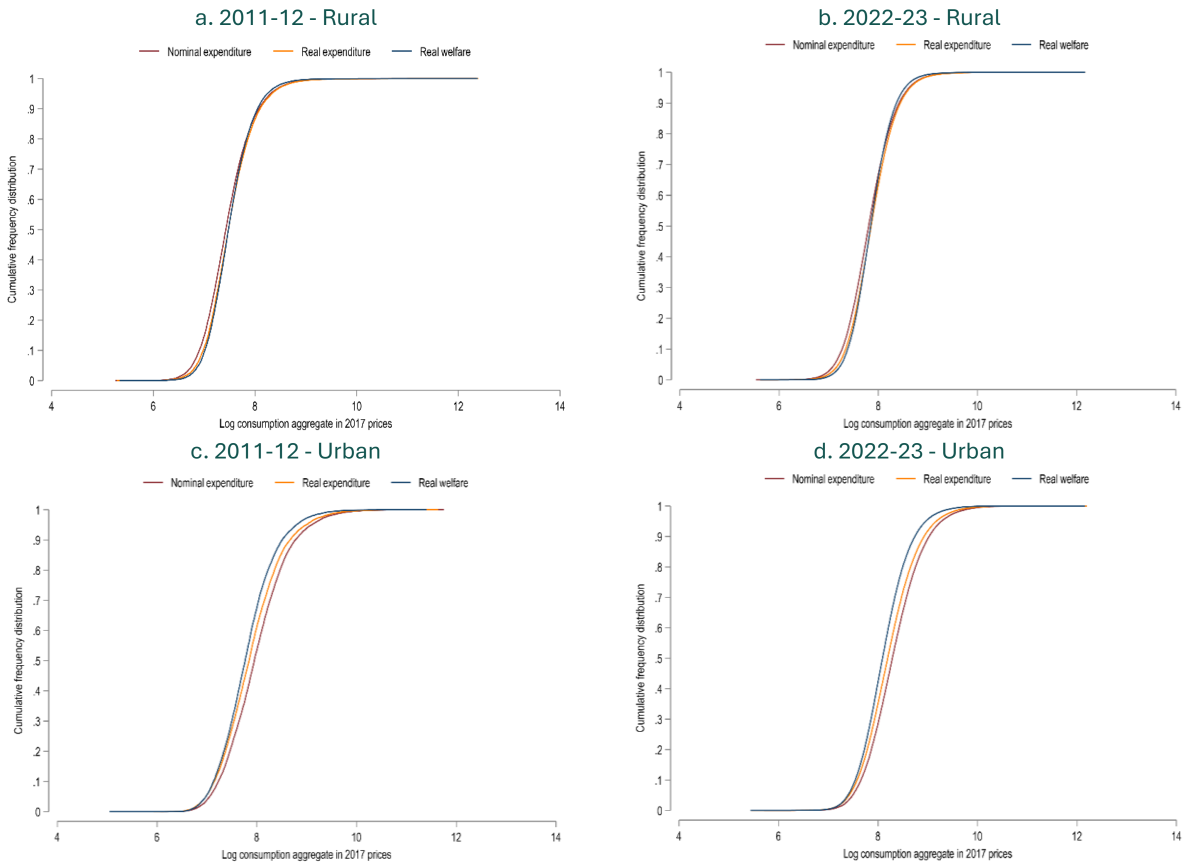
Figure A.6.4. Hedonic model's prediction accuracy



[62] Beyond traditional survey data, other variables that can help predict rent and housing prices include the historical significance or architectural style, local zoning regulations or development potential, flood risk or other exposure to natural hazards, and proximity to negative externalities (e.g., landfills, industrial sites).

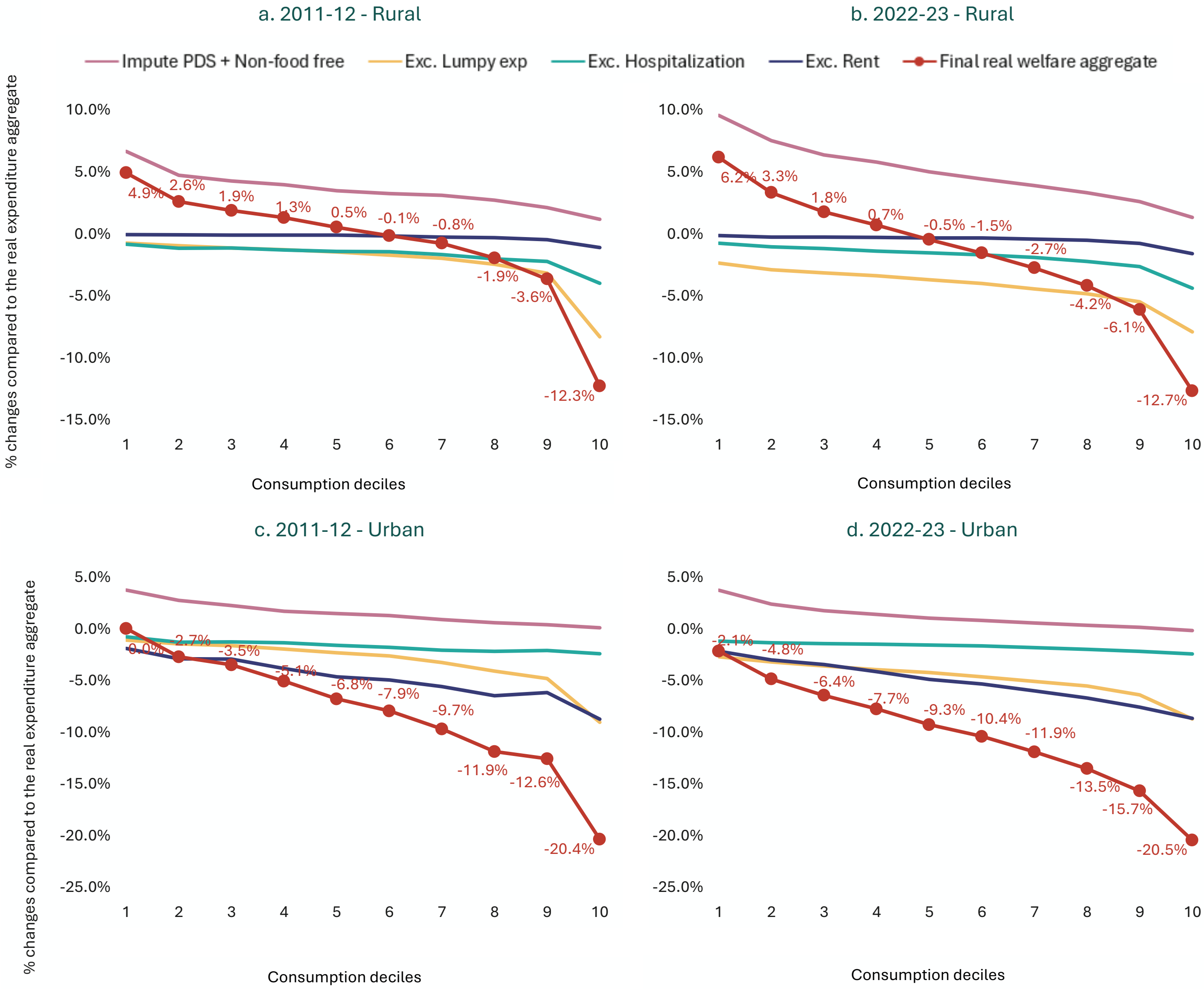
Annex 7. Additional figures and tables

Figure *A.7.1: Cumulative Density Functions for real welfare aggregate*



Note: A CDF shows the proportion of population with consumption less than a certain value in the distribution, therefore capturing the spread of the distribution. The log values of nominal expenditure, real expenditure and real welfare in 2017 prices are plotted on x-axis while y-axis shows the distribution (CDF) of each in red, orange and blue line, respectively.

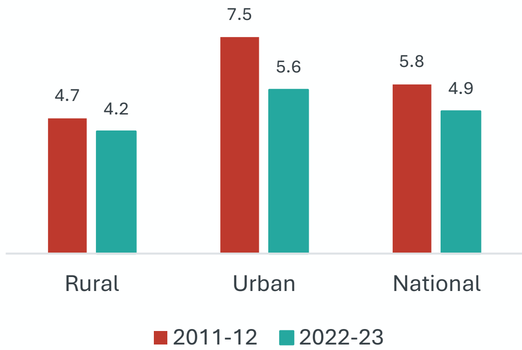
Figure *A.7.2. Impact of adjustments on welfare aggregate*



Note: The figure plots the percentage difference in the average expenditure of a decile under the given scenario relative to the average expenditure of the decile in the baseline scenario (real expenditure distribution). Deciles are created for each sector and each scenario seperately.

Figure *A.7.3. Percentile-based measures of dispersion in the final real welfare aggregate*

a. Ratio of real welfare aggregate of top decile to the bottom decile (P90/P10)



b. Ratio of real welfare aggregate of top quantile to the bottom quantile (P75/P25)

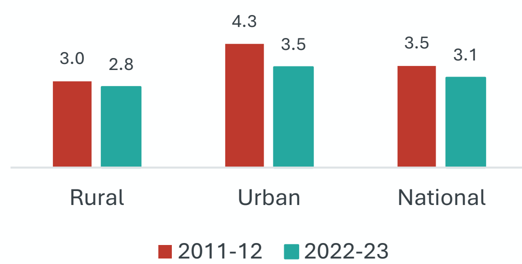


Table A.7.1. Transition of individuals between deciles after adjustments, rural (2011-12)

	Real welfare aggregate deciles									
Real expenditure aggregate deciles	1	2	3	4	5	6	7	8	9	10
1	**8.8%**	1.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	1.0%	**7.1%**	1.7%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0.1%	1.4%	**6.6%**	1.7%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%
4	0.1%	0.2%	1.3%	**6.3%**	2.0%	0.2%	0.0%	0.0%	0.0%	0.0%
5	0.0%	0.1%	0.2%	1.3%	**6.3%**	2.0%	0.1%	0.0%	0.0%	0.0%
6	0.0%	0.0%	0.1%	0.2%	1.1%	**6.4%**	2.1%	0.1%	0.0%	0.0%
7	0.0%	0.0%	0.1%	0.1%	0.3%	1.1%	**6.4%**	2.0%	0.0%	0.0%
8	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	1.0%	**6.7%**	1.8%	0.0%
9	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.3%	1.0%	**7.3%**	1.3%
10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.9%	**8.7%**

Note: The table presents the population shares across the ten deciles using real expenditure distribution and real welfare distribution.

Table A.7.2. Transition of individuals between deciles after adjustments, urban (2011-12)

	Real welfare aggregate deciles									
Real expenditure aggregate deciles	1	2	3	4	5	6	7	8	9	10
1	**8.9%**	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	1.0%	**7.3%**	1.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0.1%	1.3%	**6.3%**	2.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
4	0.0%	0.2%	1.4%	**5.5%**	2.8%	0.0%	0.0%	0.0%	0.0%	0.0%
5	0.0%	0.0%	0.4%	1.7%	**4.7%**	3.2%	0.1%	0.0%	0.0%	0.0%
6	0.0%	0.0%	0.1%	0.4%	1.7%	**4.3%**	3.4%	0.0%	0.0%	0.0%
7	0.0%	0.0%	0.0%	0.1%	0.6%	1.7%	**4.3%**	3.2%	0.0%	0.0%
8	0.0%	0.0%	0.0%	0.0%	0.2%	0.5%	1.7%	**4.9%**	2.6%	0.0%
9	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.5%	1.6%	**6.1%**	1.5%
10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	1.2%	**8.5%**

Note: The table presents the population shares across the ten deciles using real expenditure distribution and real welfare distribution.

Table A.7.3. Transition of individuals between deciles after adjustments, rural (2022-23)

	Real welfare aggregate deciles									
Real expenditure aggregate deciles	1	2	3	4	5	6	7	8	9	10
1	**8.6%**	1.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	1.2%	**6.7%**	1.9%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0.1%	1.5%	**5.9%**	2.2%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%
4	0.0%	0.2%	1.6%	**5.5%**	2.5%	0.2%	0.0%	0.0%	0.0%	0.0%
5	0.0%	0.1%	0.3%	1.5%	**5.3%**	2.6%	0.1%	0.0%	0.0%	0.0%
6	0.0%	0.0%	0.1%	0.3%	1.4%	**5.4%**	2.7%	0.1%	0.0%	0.0%
7	0.0%	0.0%	0.1%	0.1%	0.3%	1.3%	**5.5%**	2.6%	0.0%	0.0%
8	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	1.2%	**5.8%**	2.4%	0.0%
9	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.3%	1.2%	**6.5%**	1.6%
10	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	1.1%	**8.4%**

Note: The table presents the population shares across the ten deciles using real expenditure distribution and real welfare distribution.

Table A.7.4. Transition of individuals between deciles after adjustments, urban (2022-23)

	Real welfare aggregate deciles									
Real expenditure aggregate deciles	1	2	3	4	5	6	7	8	9	10
1	**8.6%**	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	1.1%	**6.4%**	2.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	0.2%	1.5%	**5.1%**	3.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
4	0.0%	0.5%	1.4%	**4.3%**	3.7%	0.0%	0.0%	0.0%	0.0%	0.0%
5	0.0%	0.1%	0.6%	1.5%	**3.8%**	3.9%	0.1%	0.0%	0.0%	0.0%
6	0.0%	0.1%	0.2%	0.7%	1.5%	**3.5%**	3.9%	0.0%	0.0%	0.0%
7	0.0%	0.0%	0.1%	0.2%	0.6%	1.6%	**3.6%**	3.8%	0.0%	0.0%
8	0.0%	0.0%	0.0%	0.1%	0.2%	0.7%	1.8%	**4.1%**	3.1%	0.0%
9	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.5%	1.8%	**5.5%**	1.9%
10	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.3%	1.4%	**8.1%**

Note: The table presents the population shares across the ten deciles using real expenditure distribution and real welfare distribution.