INCOME AND WEALTH INEQUALITY IN INDIA, 1922-2023:

THE RISE OF THE BILLIONAIRE RAJ

NITIN KUMAR BHARTI
LUCAS CHANCEL
THOMAS PIKETTY
ANMOL SOMANCHI

WORKING PAPER N°2024/09

MARCH 2024

WORLD INEQUALITY LAB

Income and Wealth Inequality in India, 1922-2023: The Rise of the Billionaire Raj *

Nitin Kumar Bharti[1], Lucas Chancel[2], Thomas Piketty[3], and Anmol Somanchi[3]

[1] New York University, Abu Dhabi and World Inequality Lab

[2] Sciences Po, Harvard Kennedy School and World Inequality Lab

[3] Paris School of Economics and World Inequality Lab

March 18, 2024

Abstract

We combine national income accounts, wealth aggregates, tax tabulations, rich lists, and surveys on income, consumption, and wealth in a consistent framework to present long run homogeneous series of income and wealth inequality in India. Our estimates suggest that inequality declined post-independence till the early 1980s, after which it began rising and has skyrocketed since the early 2000s. Trends of top income and wealth shares track each other over the entire period of our study. Between 2014-15 and 2022-23, the rise of top-end inequality has been particularly pronounced in terms of wealth concentration. By 2022-23, top 1% income and wealth shares (22.6% and 40.1%) are at their highest historical levels and India's top 1% income share is among the very highest in the world. In line with earlier work, we find suggestive evidence that the Indian income tax system might be regressive when viewed from the lens of net wealth. We emphasize that the quality of economic data in India is notably poor and has seen a decline recently. It is therefore likely that our results represent a lower bound to actual inequality levels. We call for improved access to official data and greater transparency to enhance the study of inequality and enable evidence-based public debates.

JEL codes: D31, E01, E21, N35, O15

Key words: India, income inequality, wealth inequality, top shares

* We are grateful to Jayati Ghosh for her thoughtful comments and to Rajesh Shukla for providing us with ICE360 tabulations and relevant clarifications. We acknowledge financial support from the French National Agency for Research (ANR-17-EUR-0001), ERC Synergy DINA 856455 and WISE Horizon 101095219 research grants. Lucas Chancel acknowledges support from the Stone Program in Wealth Distribution, Inequality, and Social Policy at the Harvard Kennedy School. **Emails**: Bharti: `nkb2039@nyuad.edu`; Chancel: `lucas.chancel@sciencespo.fr`; Piketty: `thomas.piketty@psemail.eu`; Somanchi: `anmol.somanchi@psemail.eu`.

Given its geographical size and population, now highest in the world, the distribution of economic growth in India has significant consequences on global inequality dynamics which in turn are crucial to our understanding of global economic arrangements. This makes the careful measurement of income and wealth inequality in India an important exercise. Between 1960 and 2022, India's average income grew at 2.6% per year in real terms (Figure 1a). Till the mid-1970s, aggregate national income experienced significant year-on-year volatility and growth remained sluggish (Appendix Figure A.1). Growth only really picked up sometime in the 1980s and then accelerated during the 1990s and 2000s. Compared to a real growth rate of 1.6% per year between 1960 and 1990, average incomes grew by 3.6% per year between 1990 and 2022. The periods 2005-2010 and 2010-2015 saw the fastest growth at 4.3% and 4.9% per year respectively (Appendix Figure A.2). Growth appears to have, however, slowed down somewhat in recent years. Placing India in comparative perspective with China and Vietnam, average incomes were similar in the three countries till about 1975, adjusting for local inflation and purchasing power differentials (Figure 1b). Subsequently, incomes in China and Vietnam grew to become 35%-50% higher than India's by 2000. At the turn of the new century, Chinese incomes began galloping ahead and are now about 2.5 times larger than Indian incomes. Vietnam on the other hand seems to have slowed down slightly after 2005 and by 2022 average incomes were about 33% higher than India's. Nonetheless, despite the absence of democracy, incomes in both China and Vietnam have grown faster than in India over the 1960-2022 period.[1]

Notwithstanding the relatively timid performance compared to China, the Indian economy did experience significant growth in absolute terms during the last three decades. The income growth of nearly 3.6% per year during 1990-2022 was accompanied by a rise in national wealth - India's aggregate wealth-to-income ratio (β) rose from 3.83 in 1995 to 5.75 in 2022 (Figure 2a). This period also saw the emergence of very high net worth individuals. As per data from Forbes billionaire rankings, the number of Indians with net wealth exceeding 1 billion USD at market exchange rate (MER) increased from 1 to 52 to 162 in 1991, 2011 and 2022 respectively (Appendix Table C.2). Over this period, the total net wealth of these individuals as a share of India's net national income boomed from under 1% in 1991 to a whopping 25% in 2022 (Figure 2b). Perhaps not surprisingly, the share of adult population that filed an income tax return, which had remained under 1% till the 1990s, also grew significantly in the three decades post the economic reforms of 1991. By 2011, the share had crossed 5% and the last decade too saw sustained growth with around 9% of adults filing a return in the years 2017-2020 (Figure 3).

What were the distributional consequences of this sustained economic growth in recent decades when looked at from a historical perspective? The Distributional National Accounts (DINA) project

[1] Both China and Vietnam have consistently invested much more on health and education than India. In 1980, India's adult (15+) literacy rate was just 41% compared to 65% in China and 84% in Vietnam; in 2010, India's expenditure on health was 1.2% of its gross domestic product (GDP) compared to 2.7% in China (Drèze and Sen, 2013).

at the World Inequality Lab (WIL) has aimed at rigorously estimating inequality dynamics in various countries using a range of data sources and cutting-edge statistical methods in a consistent and comparable manner (Blanchet et al., 2017). For developed economies like the United States (US) and France, quasi-exhaustive micro-data from income and estate taxes, covering majority of the population, forms the core basis for measurement (Piketty et al., 2018; Garbinti et al., 2018). In contrast, tracking the dynamics of inequality in a country like India is fraught with various empirical challenges relating to data coverage, quality, and availability. In this paper, we contribute to the DINA project at WIL by combining national income accounts and wealth aggregates with tax tabulations, rich lists, and a range of household surveys on income, consumption, and wealth to present long-run homogeneous income and wealth inequality series for India updated to the most recent years. Our results point to extreme levels of inequality in India compared to international standards. In 2022-23, 22.6% of national income went to just the top 1%, the highest level recorded in our series since 1922, higher than even during the inter-war colonial period. The top 1% wealth share stood at 40.1% in 2022-23, also at its highest level since 1961 when our wealth series begins. In other words, the 'Billionaire Raj' headed by India's modern bourgeoisie is now more unequal than the British Raj headed by the colonialist forces. It is unclear how long such inequality levels can sustain without major social and political upheaval. While there is no reason to believe income and wealth inequality will slow down by itself, historical evidence suggests that it can be kept in check *via* policy. Implementing a super tax on Indian billionaires and multimillionaires, along with restructuring the tax schedule to include both income and wealth, so as to finance major investments in education, health and other public infrastructure, could be effective measures in this regard.

1 What do we know so far?

Before moving to our data sources and methodology, we briefly discuss the existing literature on income and wealth inequality in India. Until the early 2000s, the question of inequality did not receive much attention in India. This may have at least partly to do with the fact that actual and perceived inequality may not have been very high prior to the 1990s.[2] The 1991 liberalization reforms and the opening of the economy to global markets seems to have made inequality a more salient issue.

1.1 Income inequality

Based on data from the National Sample Survey Organization's (NSSO)[3] consumption expenditure survey (CES) rounds in 1993-94 and 1999-2000, Deaton and Drèze (2002) found that economic inequality markedly increased during the 1990s in several forms – strong divergence across states,

[2] For instance, the Gini coefficient of per-capita consumption expenditures was fairly stable till the 1990s - see Drèze and Sen (2002), Statistical Appendix, Table A.6.

[3] In May 2019, the NSSO was merged with the Central Statistical Office and re-christened as the National Statistical Office. To avoid confusion, we shall refer to it as NSSO throughout the paper.

rising urban-rural inequality, and growing disparities *within* urban areas. Banerjee and Piketty (2005) were the first to mobilize annual tax tables provided by the income tax authorities in combination with national accounts to shed light on the dynamics of top incomes over the long run (1922-2000). They find top income shares declined post-independence but began rising during the 1980s. However, their analysis was restricted to estimating only very top income shares (top 1%, 0.1%, 0.01%) given their reliance solely on tax tabulations. Building on this, the most comprehensive analysis on income inequality in India so far is found in Chancel and Piketty (2019). They combine national accounts aggregates, household surveys on consumption and income, and tax tabulations to present harmonized long-run estimates of income inequality between 1922 and 2014. They find that income inequality began rising very sharply in the 1990s and was on the rise till 2014 when their series ends.

For the post-2014 period, given numerous data challenges, the literature on inequality has been sparse. Nonetheless, some attempts have been made using data sources from recent years to track movement in incomes. The "State of Inequality in India Report" (Kapoor and Duggal, 2022), commissioned by the Economic Advisory Council to the Prime Minister, provides a broad-based review of inequality in India in recent years along various dimensions including income, health, education, and household assets. The report acknowledges the extreme income inequality reported in Chancel and Piketty (2019) but goes on to present its own analysis using data from the Periodic Labour Force Survey (PLFS) in 2019-20. They found that the top 1% earn just over 6%-7% of the total incomes. Their analysis is however fraught with several issues - it only uses data on labour and self-employment incomes, restricted to those employed at the time of survey (non-zero incomes), and does not make use of tax tabulations thereby missing a large part of the right-tail of the distribution.[4] Nonetheless, the report draws attention to the fact that "… the benefits of growth have been concentrated and has marginalised the poor further".

In the absence of any NSSO consumption survey in recent years, some researchers have turned to the privately executed Consumer Pyramids Household Survey (CPHS), which provides monthly data on both incomes and consumption, to track trends in inequality. Using CPHS and tabulations from the suppressed 2017-18 CES round, Ghatak et al. (2022) point to a possible slow-down in economic inequality in India in recent years. Gupta et al. (2021) use monthly CPHS data to argue that income and consumption inequality declined during the COVID pandemic in India. While this may have indeed been the case, the CPHS has been shown to not be representative of the population as it misses poorer and less-educated households in its sample, with the bias seemingly growing over time (Drèze and Somanchi, 2021; Somanchi, 2021). This makes accurately assessing both levels and trends of inequality using CPHS a rather complicated task. Moreover, both studies do not use tax tabulations.[5]

[4] Moreover, capital incomes, that PLFS does not capture, are likely to be crucial for the right-tail. For instance, as per data in the income tax return statistics for 2018-19 published by the income tax authorities, 42% of total incomes in the returns came from non-labour sources.

[5] To be fair to Ghatak et al. (2022), they themselves acknowledge both limitations with their analysis.

Other innovative pieces of evidence point to serious inequalities in recent years. Based on data from legally mandated disclosures, Khera and Yadav (2020) report starkly unequal median-to-top pay ratios among NIFTY50 companies along with a complete lack of diversity in top management beyond upper-caste men.[6] Pai and Vats (2023) find a "brutal power law in most Indian consumer transactions" – as per their data, 1% of Indians take 45% of flights, 2.6% of Indians invest in mutual funds, 6.5% of users are responsible for 44% of digital transactions on the Unified Payment Interface (UPI), and 5% of users account for a third of the orders placed on Zomato (the most prominent food-delivery application). As per the annual Forbes rich lists, the net wealth of USD MER billionaire Indians has grown by over 280% cumulatively between 2014 and 2022 in real terms, 10 times the growth rate of national income over the same period (27.8%). All of this suggests that at least the very rich seem to be doing very well in recent years.

We contribute to this literature on income inequality in India in a few relatively modest ways. First, despite severe data challenges in recent years, including absence of key data sources, we are able to construct a comparable income inequality series for recent years, allowing us to extend the series in Chancel and Piketty (2019) from 2014-15 till 2022-23.[7] As amply demonstrated by the revival of poverty debates in India in recent years, the numerous challenges involved in estimating consumption and incomes for the bottom part of the distribution in recent years cannot be stressed enough. Second, using our income series for the recent years, we place India in comparative perspective. We find that by 2022-23, India's top 1% income share is among the very highest in the world (higher than even South Africa, Brazil and the US). We also present some India vs. China comparisons over the 1980-2022 period which seem to hold lessons for low- and middle-income countries facing a potential growth-inequality trade-off. Third, by combining our income and wealth series, we present wealth-to-income ratios across the distribution in India in recent years. Our results are in line with those in Singh (2023) who finds similar estimates based on income and wealth disclosures by Indian politicians, an admittedly non-representative sample of the Indian population. Both sets of estimates suggest that contrary to traditional wisdom, the Indian income tax system might be *regressive* when viewed from the point of view of net wealth – that is, the more wealth taxpayers own, the less taxes they pay as a share of their assets.

1.2 Wealth inequality

The literature on wealth inequality in India also takes off only in the 2000s. Perhaps the first comprehensive study of wealth distributions in India is found in Subramanian and Jayaraj (2006) who combine successive AIDIS rounds to study wealth inequality in India between 1991-92 and 2002-03.

[6] The NIFTY50 is a free-float market capitalization weighted index of the 50 largest Indian companies listed on the National Stock Exchange.

[7] For the recent period we also draw on the work in Anmol Somanchi's MSc thesis, see Somanchi (2023).

They report bottom 50% shares of 8% in 1991-92 and 2002-03 and top 1% shares of 15.7%. They also present an early attempt at combining surveys with data on the richest 178 households from a rich list published by 'Business Standard' to re-estimate wealth shares in 2002-03 – they find the top 1% share increased from 15.7% to 17.8%. Sinha (2006) used the Business Standard rich lists to show that the very top of the wealth distribution is well approximated by a Pareto distribution. Jayadev et al. (2007) also used the 1991-92 and 2002-03 AIDIS rounds to study trends in wealth concentration and find a small rise in wealth inequality during this period along with a growing divergence in wealth growth across states and across socio-economic groups. Anand and Thampi (2016) and Jayaraj and Subramanian (2018) study wealth inequality dynamics using three successive AIDIS rounds between 1991 and 2012. Both studies find significant concentration of wealth over these two decades. Bharti (2018) combined surveys and rich lists to produce top-corrected wealth inequality estimates till 2012. Though a wealth tax was in place in India between 1957 and 2016, it evolved to have a very low base over the years on account of exemptions for various assets like shares, mutual funds and securities. The tax was finally abolished due to the low revenue collection and high cost of tax collection. Hence, the literature has relied on rich lists to correct the surveys at the top. With the release of the 2018 AIDIS round, combining it with Hurun rich lists, Anand and Kumar (2023) find that top 10% shares declined between 2012-13 and 2018-19 while the shares within the top 1% increased. Abstracting away from the differences in all these works, there is consensus in the literature that wealth inequality was more-or-less stagnant pre-1991, after which a very clear trend of rising top shares is evident.[8]

Despite being rich and comprehensive, there are at least three limitations with the literature as it stands. First, while there is consistency in the use of the NSSO AIDIS across all papers, there is irregularity in terms of the use of rich lists, and when used, in terms of the methodology used to combine the two sources. Further, differences in the construction and use of price deflators may have non-trivial effects on the estimates. In the absence of proper deflators for assets, the literature has typically relied on the wholesale price index or the consumer price index. Second, the present inequality estimates in the literature are essentially only available on a decadal basis given the decennial nature of the AIDIS and last available estimate is for 2018. Third, no paper has consistently studied wealth inequality over the entire 1960-2022 period. We attempt to improve upon these limitations to the extent possible by presenting homogeneous long run wealth inequality series covering 1960-2022, including annual estimates starting 2002, all while using consistent definitions, data sources, deflators, and methodology.

[8] An inheritance tax was active in India between 1953 and 1985. Kumar (2020) uses inheritance tax returns and mortality tables data with an estate multiplier technique to produce the top 0.01% wealth share during 1961-85, and finds that the shares decline during this period, possibly as a result of several policies such as nationalization, abolition of privy purses, land redistribution etc.

2 Data and methodology

As noted earlier, tracking the dynamics of income and wealth inequality in a country like India is fraught with serious empirical challenges relating to data coverage, quality, and availability. To begin with, the large-scale relevance of informal employment, relatively low-income levels, and relatively high thresholds for non-taxable incomes means tax data (in whatever form it is available) covers only a tiny fraction of the adult population - less than 10% as recently as 2020. This means tax data can at best shed light on top incomes. Consequently, nationally representative household surveys on *consumption expenditure*, conducted periodically by the NSSO, have formed the basis for studying issues like poverty and inequality. However, there are two issues with using NSSO surveys for estimating income and wealth inequality. First, with regards to incomes, NSSO steered clear from measuring incomes (on the grounds that agricultural incomes - which are highly seasonal - are hard to decipher) and instead focused on measuring consumption expenditure. While certainly a good proxy for incomes, consumption dynamics significantly differ from income dynamics, and given that the rich only consume a fraction of their incomes, consumption inequality tends to understate income inequality. Second, income and wealth surveys (everywhere, not just in India) are known to be fraught with concerns of under-reporting and differential non-response by the rich and wealthy (Korinek et al., 2006; Bourguinon, 2018). Moreover, simply owing to their small numbers, unless explicitly over-sampled, the very rich and wealthy are unlikely to make it into usual samples (Vermeulen, 2016). This has meant that surveys alone are not sufficient for accurately estimating inequality where the right-tails matter a lot.

2.1 Data sources

To make progress in such a scenario, we mobilize data from a wide variety of sources to shed light on long run income and wealth inequality in India. Table 1 provides a summary of the datasets and their sources. We provide a brief description of each of them here. The exact role these datasets play in our methodology is elaborated upon in detail in the next sub-sections.

Population aggregates: Adult (20+) population figures are sourced from United Nations World Population Prospects (UN-WPP). The age cut-off for defining adults as well as the data source are chosen in line with DINA guidelines to allow for consistent cross-country comparisons (Blanchet et al., 2017).

Table 1: Summary of data sources used in this research

Data	Source	Years	Type
Adult (20+) Population	UN-WPP	1951-2022	Aggregates
National Income Accounts		1951-2022	
Pre-2014	WID	1951-2014	Aggregates
Post-2014	MoF, GoI	2015-2022	Aggregates
Price Index (GDP deflator)	WB-WDI	1960-2022	Aggregates
All-India Income Tax Returns Statistics	MoF, GoI	1922-2020	Tabulated
All-India Income Tax Time Series Data	MoF, GoI	2011-2020	Aggregates
Indian Billionaire Rankings	Forbes	1988-2022	Tabulated
India Rich Lists	Hurun	2012-2023	Tabulated
Consumption Expenditure Survey (CES)		1951-2017	
Pre-1983	WB-PGID	1951-1977	Tabulated
Post-1983	NSSO	1983-2011	Micro-data
2017-18	NSSO	2017-18	Tabulated
All India Debt and Investment Survey (AIDIS)		1961-2018	
Pre-1991	RBI/NSSO	1961-1981	Tabulated
Post-1991	NSSO	1991-2018	Micro-data
India Human Development Survey (IHDS)	NCAER/UMD	2005 & 2011	Micro-data
India's Consumer Economy 360^o Survey (ICE360)	PRICE	2016 & 2020	Tabulated
Periodic Labour Force Survey (PLFS)	NSSO	2017-2022	Micro-data

Notes: (1) UN-WPP - United Nations World Population Prospects; WID – World Inequality Database; MoF - Ministry of Finance; GoI - Government of India; WB-WDI - World Bank World Development Indicators; WB-PGID – World Bank Poverty and Growth in India Database (Özler et al., 1996); NSSO - National Sample Survey Organization; RBI – Reserve Bank of India; NCAER - National Council of Applied Economic Research; UMD – University of Maryland; PRICE – People's Research Institute on India's Consumer Economy. (2) Most datasets more-or-less correspond to financial year (so 2011 here means FY2011-12), except for adult population, Forbes billionaire rankings and Hurun rich lists which are available by calendar year. (3) In the "Type" column, "Aggregates" refer to all-India totals while "Tabulated" data contains some form of distributional information (eg. fractile thresholds and/or averages and/or rankings).

National income accounts: Aggregate totals of net national income (NNI) are obtained from the series in the WID for the years pre-2014. For post-2014, we use the figures reported in Table 1.1., Statistical Appendix of the Economic Survey 2022-23 published by the Ministry of Finance (MoF), Government of India (GoI). We assume that a set of deductions apply (eg. to account for retained earnings of corporates, non-taxable income etc.) to go from NNI to fiscal income (Atkinson, 2007).[9] There has been an ongoing debate on the quality of Indian national accounts data in the recent years, with various observers arguing that the official statistics might be exaggerating growth – for our estimation exercise, we take the official statistics as given but return to this issue later.

Price index: The two main price indices generally used are the GDP deflator and consumer price index (CPI). While both measure changes in prices, the former focuses on changes in prices of domestic production while the latter on price changes faced by consumers. Our preferred measure is the GDP deflator since its methodology has seen relatively greater progress in terms of bias reduction when measuring prices over time by switching to the chain-linking method (Piketty and Zucman, 2014).[10] We source the annual series of the GDP deflator for India from World Bank's World Development Indicators database for the period 1960-2022.

Tax tabulations: Since the establishment of the Income Tax Act in 1922, the Indian government has published annual tax tabulations with information on total number of tax filers and total income assessed for many income brackets (24 in recent years). These were available largely uninterrupted till 1999 when their publication was abruptly stopped. Their release began again in 2016 with retrospective data starting 2011 and so far tax tabulations are available till 2020. We return to the issue of release of tax tabulations later. We use these tabulations to extract a full distribution of top income earners to supplement income and consumption surveys at the top.

Rich lists: In the last few decades, two sources provide information on the wealth of richest Indians. Starting 1988, the 'Forbes' magazine has been tracking the wealth of all Indians with net wealth exceeding 1 billion USD MER and publishing an annual billionaire ranking. Similarly, since 2012, the luxury research organization 'Hurun' has been publishing an annual rich list for India, going beyond just USD billionaires, covering all Indians with a net wealth exceeding INR 1,000 crore (roughly 120 million USD MER as on March 1st, 2024). These lists have become a very useful complement to wealth surveys which tend to severely under-estimate the right tail of the wealth distribution.

Consumption Expenditure Survey (CES): Starting in 1951, the NSSO has been conducting large-scale nationally representative household surveys recording household consumption expenditure. These surveys record detailed information on quantities purchased and prices paid for over 300 commodities. In the absence of any comparable large-scale income survey, the CES has over the years become the bedrock to study a range of economic issues in India, including but not limited to poverty, inequality, and nutrition. For the period 1951-1983, we use the tabulations for rural and

[9] In practice we assume that roughly 70% of net national income accounts for fiscal income. This serves as the "control average" when extracting a distribution from the tax tabulations using generalized Pareto interpolation.

[10] See Blanchet (2017) for a summary and Blanchet et al. (2017) for full details of the price indices used in WID.

urban per-capita consumption available in the World Bank's Poverty and Growth in India database (Özler et al., 1996). From 1983-2011, we use the publicly available unit-level micro-data. Finally, for the year 2017-18, we use the leaked summary made publicly available by a journalist since the detailed report and micro-data was suppressed by the Government of India.[11]

All India Debt and Investment Survey (AIDIS): The AIDIS are decennial surveys conducted in 1961, 1971, 1981, 1991, 2002, 2012, and 2018 which collect detailed data on household wealth and debt.[12] Micro-individual survey files are available for the last four rounds. For previous surveys, only published reports are available with tabulations on average wealth and number of households by wealth brackets. The coverage of assets in these surveys has largely remained similar over the years. For e.g., up to 2002, total assets included the value of seven different types of assets- i) land ii) buildings iii) livestock iv) implements, machinery, transport equipment, etc. v) durable household assets vi) dues receivables on loans advanced and vii) financial assets. In terms of liabilities, the surveys capture cash loans. In 2012 and 2018, household durables were excluded on the pretext of valuation concerns. Hence, we remove them from earlier rounds to make the rounds comparable across time.[13] The valuation of assets is self-reported and is based on the market price prevalent in the area, with the noted exception of building values which were recorded as per book value in 2012 and land values which were recorded as per normative/book values in 2012 and 2018.[14]

India Human Development Survey (IHDS): Conducted so far in two rounds in 2004-05 and 2011-12 as a collaboration between the University of Maryland and the National Council of Applied Economic Research (NCAER), the IHDS is a nationally representative survey covering over 40,000 households in both rural and urban areas. IHDS collected data on an unusually rich set of indicators and included a panel component. The key advantage for our purpose is that it recorded information on both consumption and income of households. Consequently, IHDS plays a key role in complementing the CES in our measurement framework.

Periodic Labour Force Survey (PLFS): Starting 2017-18, the NSSO has been conducting annual rounds of the Periodic Labour Force Survey as a replacement to the erstwhile Employment and

[11] A preliminary "fact sheet" of the 2022-23 round of the consumption expenditure survey by NSSO has just been placed in the public domain (Press Information Bureau, 2024). The full report and micro-data are still awaited. Unfortunately, there appear to be key changes to the instrument and interview schedules that raise issues of comparability with previous CES rounds. Indeed, a whole page in the fact-sheet titled "Issues related to Comaparability" (pg. 4) is dedicated to cautioning the reader against hasty comparisons (National Sample Survey Office, 2024).

[12] The Reserve Bank of India (RBI) conducted the 1961 survey in rural areas only. The 1971-72 survey was conducted by NSSO and RBI together in rural and urban areas. The urban data was never published due to sampling issues. NSSO independently conducted all the later rounds in both rural and urban areas.

[13] The surveys collected a total of 159 items under different heads in the 1991-92 survey, 141 items in 2002-03, 86 in 2012-13, and 87 items in 2018-19.

[14] Enumerators had to consult 'Patwaris' (or equivalent) in rural areas and the Registrar's office in urban areas to obtain them.

Unemployment Surveys (EUS) that were being conducted at regular intervals till 2011-12. The PLFS largely follows a similar methodology to earlier the EUS, except that it records data on both labour incomes as well as incomes from self-employment, while EUS only recorded labour incomes. Additionally, the PLFS also provides data on 'usual' consumption expenditure by households. This consumption measure is at the outset not comparable to the consumption reported in the traditional CES surveys, but we are able to find a mapping between them, as described in the next section.

India's Consumer Economy 360^o Survey (ICE360): People's Research Institute on India's Consumer Economy (PRICE), a private non-profit research organization, has been conducting surveys at regular intervals starting 2014, covering over 40,000-60,000 households. While the survey covers only 23 major states and has been relatively under-used so far, it provides detailed data on household consumption expenditures, incomes, savings, and assets. Currently we only have access to percentile-level tabulations of per-capita income from the 2015-16 and 2020-21 rounds of the survey.

2.2 Methodology for income series

We are interested in estimating the distribution of per-adult pre-tax income. Very broadly speaking, the idea of the distributional national accounts (DINA) framework is to take national income accounts aggregates of total income and combine it with estimates of income at various points of the income distribution to estimate the share of national income going to different parts. For a country like India, we do not have any single source for incomes that is reliable for the entire distribution. The strength of tax data is its ability to shed light on top incomes, but its weakness is its limited coverage. On the other hand, while surveys come with much wider coverage, they tend to severely miss and/or underestimate incomes of the rich. In such contexts then, the DINA guidelines recommend combining national accounts aggregates with top incomes from tax data and middle and low incomes from surveys in a consistent and harmonious manner (Blanchet et al., 2017). This is the broad framework and approach we adopt, tweaking and tailoring it where necessary to account for India's unique data challenges. Our overall methodology can be broadly described in 4 steps.

Step 1: Estimating top incomes from tax tabulations

We use annual tax tabulations published by the Indian income tax authorities to extract the distribution of top income earners between 1922-2020. The tax tabulations provide data on the total number of income tax returns filed and the total income assessed for 25 income brackets by type of tax filers (individuals, corporation, etc.). We consider individuals and Hindu Undivided Families (HUFs) as the relevant tax units and combine them at the income tax bracket level for our analysis.[15] We use

[15] Starting 2011, the tax authorities also began publishing a separate annual publication titled "Income Tax Time Series Data" (ITSD) with data on total *effective* taxpayers, defined as the sum of units that filed income tax returns and those that paid tax at source but did not file a return. By comparing the total filers in the tax tabulations with the total effective taxpayers in the ITSD, we get an estimate of the non-filers which are missing from the tax tabulations which we use to

the adult (20+) population to define the fractiles (ranks) corresponding to the respective income thresholds in the tax data based on the number of returns in each bracket. We then use generalized Pareto interpolation techniques, developed in Blanchet et al. (2022), to extract a smooth distribution of top income earners from the tax tabulations. Generalized Pareto interpolation techniques relax the strict Paretian assumption, work without making parametric assumptions and are found to empirically outperform other "standard" Pareto interpolation methods. Given our extensive reliance on it in the paper, brief technical details are provided in Appendix D and readers are directed to Blanchet et al. (2022) for further details. As output, generalized Pareto interpolation gives us 127 generalized percentiles (g-percentiles) corresponding to the bottom 99 percentiles ($p1$, $p2$, …, $p99$), 9 deciles within the top percentile ($p99.1$, $p99.2$, …, $p99.9$), 9 deciles within the top 0.1% ($p99.91$, $p99.92$, …, $p99.99$), and finally 9 deciles within the top 0.01% ($p99.991$, $p99.992$, …, $p99.999$). This gives us a smooth distribution of top income earners annually from 1922-2020. This is except for the period 1999-2010 for which tax tabulations were never published. To fill this gap, we construct a tax distribution for 2005 by applying percentile-level growth rates observed in surveys between 1999 and 2005 to the tax data observed in 1998.

Step 2: Estimating middle and bottom incomes from surveys

Given the limited coverage of tax data, it can be considered reliable only for top incomes, covering at most about 10% of the adult population in recent years. To estimate incomes for the rest of the population, we must rely on household surveys. Historically, NSSO has steered clear of measuring incomes owing to the difficulty of accurately assessing incomes in the highly seasonal agricultural sector. Instead, they have focused on collecting data on consumption expenditures. By distributing total household consumption among all adult (20+) members, we begin by estimating an "equal-split per-adult" consumption distribution from successive rounds of CES available at regular intervals between 1951-2011. While certainly a relevant proxy for incomes, consumption expenditures are likely to under-state incomes of the rich given their relatively lower marginal propensity to consume. To work around this issue, we follow Chancel and Piketty (2019) and rely on the India Human Development Survey (IHDS) which collected data on both incomes and consumption expenditure to estimate consumption-to-income scaling ratios at each percentile of the consumption distribution. Letting y denote income, c consumption, and p percentiles, we begin by estimating 'raw' consumption-to-income scaling ratios as $\alpha_p = y_p/c_p$. One issue when do this is that we are confronted with negative savings. We therefore define an alternate set of ratios where we winsorize the 'raw' ratios at 1 (from below) such that incomes are assumed to be at least as large as consumption. Our final choice that feeds into our benchmark series is an average of these two scenarios (which only affects the ratios at the bottom as ratios in the middle and top are not affected by the winsorization). Finally, since

scale-up the total returns in the tax tabulations. Since ITSD does not provide data on effective taxpayers dis-aggregated by tax bracket, we distribute these non-filers in the same proportion as filers, essentially assuming that their distributions are identical. Chancel and Piketty (2019) test different ways of distributing the non-filers, for instance assuming that they all fall in the lowest brackets and find this does not make much difference.

we have two rounds of IHDS, we estimate and average the ratios from the two round. These final scaling ratios are then applied to each percentile of the consumption distribution estimated from CES to arrive at a survey-based income distribution. These scaling ratios are presented in Appendix Figure B.1. For all the years when NSSO CES data are not available, we interpolate survey incomes by assuming a constant compound annual growth rate between two time periods.[16] This gives us an annual distribution of survey-based incomes between 1951-2011.

As if matters were not complicated enough already, they get worse post-2011. After more-or-less uninterrupted execution and release of the CES between 1951-2011, the Government of India suppressed the report and micro-data of the 2017-18 CES round on rather unqualified grounds of poor data quality.[17] An analysis of the numbers in a leaked summary of the report suggests that the consumption-levels in real-terms may have fallen across the distribution (Subramanian, 2019b). With the suppression of the 2017-18 CES data, we are left largely in the dark about the economic progress of majority of the Indian population. Even as we write this, no CES micro-data is available post-2011.[18] Therefore, the main roadblock to estimating bottom and middle incomes for the most recent decade or so is the absence of any CES-comparable consumption survey.

To work around this, we turn to another NSSO survey, PLFS, available annually from 2017 onwards. While PLFS is primarily a labour force survey, it collects data on households' 'usual' consumption expenditure as part of its household listing exercise. Unlike the traditional approach of CES which collects detailed data on expenditures on individual commodities, the consumption data in PLFS is based on 4 or 5 very broad questions. This difference in measurement is likely to render the consumption measures in PLFS and CES incomparable (Deaton and Kozel, 2005). To resolve this, we rely on the fact that the suppressed CES 2017-18 round roughly coincided with the PLFS 2017-18 round. Further, given that both surveys were conducted by the NSSO, and their sampling designs were quite similar, we attribute any differences in reported consumption to differences in the way consumption was measured across the surveys. By point-wise comparing the CES and PLFS consumption distributions at each percentile for 2017-18 for rural and urban areas separately, we create a mapping from one consumption concept to the other.[19] These PLFS-to-CES scaling ratios

[16] In practice, for all years between t and $t+n$ for which we have data, we interpolate incomes at each percentile as $y_{p,t+1} = y_{p,t} g_p$ where $g_p = (y_{p,t+n}/y_{p,t})^{1/n}$ is the growth factor and $(g_p - 1)$ the compound annual growth rate (CAGR) at percentile p.

[17] On the heels of this decision, 108 eminent economists and social scientists made a public appeal to the government to "… restore access and integrity to public statistics ... that would feed into economic policy-making and that would make for honest and democratic public discourse" - see Kazmin (2019) for details.

[18] A very preliminary report with some early estimates has been made public just as we finish writing this (Press Information Bureau, 2024).

[19] For the year 2017-18, our procedure exactly re-creates the CES 2017-18 distribution. While strictly speaking these scaling ratios apply only to 2017-18, they are likely to be time invariant as long as we attribute differences across the surveys to differences in the survey instruments since the PLFS instrument did not change in this regard over the years. Hence, we apply these ratios for subsequent years as well (2019-2022).

for rural and urban areas separately are presented in Appendix Figure B.2 which shows that in the absence of this correction, we would be systematically over-estimating inequality since PLFS appears to systematically under-estimate consumption of the poor relative to the rich.[20] Using these PLFS-to-CES consumption scaling ratios, we construct CES-comparable consumption distributions annually between 2017-2022 from PLFS data. We then use the consumption-to-income scaling ratios as before to arrive at survey-based incomes. Finally, we interpolate for the missing years between 2011-2017 and have an annual series of comparable survey-based incomes spanning 1951-2022.

Step 3: Combining tax and survey distributions

Pre-2014: On completion of steps 1 and 2, we have two income distributions covering 1951-2022 - one generalized from the tax data (going back till 1922) and the other estimated from household surveys. The former is more reliable for top incomes while the latter for bottom incomes. To put both data sources effectively to use, we follow Chancel and Piketty (2019) and deem that for $0 < p_1 < p_2 < 1$ representing percentiles in the income distribution, survey data is reliable till p_1 while tax data is reliable between p_2 and 1. We fix p_1=0.9, i.e. we assume that the survey-based distribution is always reliable only till the 90th percentile. We let p_2 be data-driven and vary by year, determined by growth of the adult population and the fractile of the lowest income bracket covered in the tax-data. During the 1950s, we set p_2 around 0.999 and in more recent years downward adjust it to 0.91-0.92 on account of the rise in tax filing.[21] The p_2 for all years in our series are presented in Appendix Table B.2. Between p_1 and p_2, we use a convex junction profile to interpolate incomes.[22] At the end of step 3, we have a "merged distribution" that combines information from surveys and tax data. This is the approach we follow to generate the full (merged) distributions for the years 1951-2014.

Post-2014: From 2014 onwards, with tax data covering close to 10% of the population, the need for a junction profile is rendered null. Hence, we switch to a "growth rate" approach. For each year 2015 onwards, we estimate growth rates between years t and $t + 1$ at each percentile from both surveys and tax data.[23] Then for all percentiles below p_2, we apply the growth rates from surveys, and for all percentiles above p_2, growth rates from the tax data. This approach is theoretically identical to the approach followed pre-2014 but is more flexible in that it only requires representative growth rates at each percentile between two points of time. Given that tax data as of now stops in 2020-21, we

[20] This is not particularly surprising given that PLFS records consumption using just a handful of questions.

[21] While the choice of p_2 is entirely data-driven in our methodology, the choice of p_1 is somewhat arbitrary. Chancel and Piketty (2019) track inequality estimates from various combinations of p_1 and p_2 and find that the choice affects the levels of inequality (not surprisingly) but not the *trends*.

[22] Chancel and Piketty (2019) try out alternate junction profiles (linear and concave) and find that it does not matter much as majority of the correction happens above p_2.

[23] We manually corrected a kink in the growth rates at the merging point p_2 by linearly interpolating between the adjacent neighbors. For the year 2020, we observe unusually high growth rates at $p93$, $p94$ and $p95$ that appear incompatible with growth rates along the rest of the distribution. Hence, we downward correct these growth rates again using linear interpolation.

apply growth rates observed in PLFS to the full distribution, including above p_2. Surveys are known to severely under-estimate the incomes of the rich, hence we expect these survey-based growth rates to likely under-state the true income growth at the top of the distribution. We hope tax tabulations for 2021 and 2022 will be released soon allowing us to improve upon our estimates.

Step 4: Anchoring to national income aggregates

A long-standing and well-acknowledged issue in the Indian context since the 1980s is the sharp divergence of growth rates of consumption observed in NSSO household surveys and those inferred from national accounts data. Chancel and Piketty (2019) show that some of this gap can be explained by missing top incomes in surveys, but a non-trivial fraction of the gap remains unexplained. To facilitate cross-country comparisons and remain consistent with DINA guidelines, we scale our survey + tax merged distribution to make income aggregates match those in national income accounts data. This step is distribution-neutral as it involves scaling all income threshold and averages by a constant factor but leaves income shares unchanged. A significant part of this gap may be explained by undistributed profits of corporations (that is, profits part of national income estimates that are not reported in tax data or survey data). Therefore, it is likely that our method underestimates incomes of the richest segments of the population, as evidence from other countries suggests.[24] This procedure implies the growth rate of average incomes in our series matches the growth rates observed in national accounts data. At the end of it all, we have annual estimates of the full income distribution for the years 1951-2022 that are consistent with aggregate national income growth over this period.

2.3 Methodology for wealth series

Pre-2002: Our wealth inequality series spans the period 1961-2023 and the estimates for 1961-91 are based on Bharti (2018). Three points are worth noting. First, the 1961 and 1971 rounds of the survey were conducted only in rural areas - to generate all-India estimates, we re-scale the rural distribution by a multiplicative factor given by the ratio of all-India wealth to rural wealth at each percentile in 1981.[25] Second, data from the 1961, 1971, and 1981 rounds are only available in tabulated form with information on average wealth and number of households by wealth brackets. Generalized Pareto interpolation techniques were used to extract a full distribution from these wealth tabulations.[26] Third, given the non-availability of any rich list during 1961-87, the estimates for this period are based solely on surveys without any correction at the top. While the Forbes billionaire rankings started in 1988, it covered too few individuals (less than 5) in the early years to be meaningfully

[24] See section 7.5 below, as well as Section 2.2.1.1 in the DINA Guidelines (Blanchet et al., 2017).

[25] This is admittedly only a crude band-aid like fix, however, we are constrained by data availability here. This approach would under-estimate true inequality if urban wealth was more skewed in 1961-1971 than in 1981 (unlikely scenario) and over-estimate inequality if urban wealth was less skewed in 1961-1971 than 1981 (more likely scenario).

[26] A non-parametric correction is then applied to convert the household-level distribution to an individual-level distribution – see Appendix 7.3.2 in Bharti (2018) for full technical details.

used. Consequently, wealth inequality levels till 1991 are likely lower bounds and the evolution of inequality between 1991 and 2002 must be interpreted with some caution.

Post-2002: For the years 2002 onwards, the estimation challenges and our broad methodology are very similar to the case of incomes. The decennial AIDIS wealth surveys have the advantage of having largest coverage of the population but at the same time grossly miss the right-tail of the wealth distribution either due to non-responses or non-sampling of very wealth households (Subramanian and Jayaraj, 2006; Jayadev et al., 2007). To put numbers to the issue at hand, consider this: the ratio of maximum wealth in the Forbes rich list to the maximum wealth observed in AIDIS was 3279 in 2012 and 7163 in 2018. In other words, there is enormous under-estimation of wealth at the top in surveys. Further, the issue of non-representativeness of the rich population appears to be worsening over time, especially with the last round of the survey conducted in 2018 - the total net wealth from the Forbes list as a percentage of total survey-based wealth increased from 1.26% in 2002 to 2.74% in 2012 to 6.01% in 2018 (based on 5, 46, and 117 individuals respectively). Clearly, Forbes and Hurun rich lists (starting 1988 and 2012 respectively) provide much better information on the wealthiest Indians than surveys. However, as with tax tabulations, the coverage of these rich lists in terms of number of individuals is very small and hence they are reliable only for the very top of the wealth distribution. Consequently, as with incomes, we supplement wealth surveys with data from the rich lists. Our methodology for the post-2002 period can again be broadly described in 4 steps.

Step 1: Estimating wealth distribution from surveys

We begin by using the AIDIS micro-data to estimate a survey-based all-India wealth distribution for the years 2002, 2012, and 2018. As noted earlier, while the coverage of asset classes has remained largely consistent over time, household durables were excluded in the 2012 and 2018 rounds on the pretext of valuation concerns. Hence, we remove them from earlier rounds to keep our series comparable. While we acknowledge the concerns regarding the reliability of debt-levels estimated in AIDIS (Chavan, 2012; Narayanan, 1988), we nonetheless choose to work with *net* wealth rather than *total* wealth to remain consistent with rich lists that only track *net* wealth. If debt is under-estimated in AIDIS as the literature suggests, then our top shares (at least top 1% and beyond) would be lower bounds since our survey-based estimates would be over-estimating the net wealth for the bottom of the distribution.

Step 2: Simulating wealth at the top from rich lists

We use data in rich lists to top-correct the AIDIS wealth surveys. Let's assume surveys are not representative above some percentile p_0 of the wealth distribution, with an associated survey-based wealth w_0. We simulate the wealth for the remaining (1-p_0) percent of the population by assuming that a Pareto distribution describes the top of the wealth distribution starting w_0. To simulate incomes above w_0, we need the tail parameter α of the Pareto distribution. We estimate it in two ways using the rich list (with N individuals): the log-linear method and constant Pareto coefficient method. In the log-linear method, we regress the log-wealth on the log of the normalized rank and estimate α_l using ordinary least squares as the slope of the line of best-fit.[27] On the other hand, the constant Pareto method assumes a constant inverted Pareto coefficient, leading to $\alpha_c = w/(w - w_0)$, where w is the average wealth of the rich list individuals. We use both methods to generate the entire top (1-p0) percent of the distribution. We find that in comparison to the constant Pareto coefficient method, the log-linear based simulations yield total wealth at the top that are similar in levels to those reported in the rich lists, thereby proving to be more accurate for our purposes (see Appendix Table C.5). Hence, we pick the log-linear estimates for our benchmark series. The survey distribution is truncated at p_0 and replaced with the simulated wealth distribution at the top to get a complete distribution. For the survey rounds in 2002 and 2012, we use $p_0 = 0.999$ implying surveys are assumed to be representative for the bottom 99.9% the population). Since the issue of non-coverage of the rich worsened in the 2018 AIDIS round as described above, we use $p_0 = 0.995$, implying that we assume the survey is non-representative for the top 0.5% of the population. We must stress that these are likely conservative choices in that we expect surveys to be non-representative even at lower percentiles. However, given the limited size of rich lists, we prefer a conservative choice here. The chosen p_0 for each year is presented in Appendix Table C.3.

Step 3: Generalized Pareto interpolation

While we could have already estimated wealth shares at the end of Step 2, the simulations for the very top assume a *strict* Pareto law applies. Generalized Pareto interpolation techniques provides an alternative less restrictive approach that does not rely on parametric assumptions, thereby possibly improving the estimation of top wealth over the standard interpolation methods. Therefore, using the distribution generated in step 2, we create twelve wealth brackets corresponding to fractiles p0, p0.1, p0.2, p0.3, p0.4, p0.5, p0.6, p0.7, p0.8, p0.9, p0.99, and p0.999 with data on bracket thresholds and bracket averages, which serves as an input for the generalized Pareto interpolation algorithm.[28] As output, we get a smooth wealth distribution with estimated thresholds, top averages and shares for 127 g-percentiles up to $p = 0.99999$, or the wealthiest 0.001%.

Step 4: Anchoring to national wealth

From 1995 onwards, estimates of aggregate national wealth (covering household, corporate, and government sectors) are available from Kumar (2019). As with incomes, to facilitate cross-country comparisons, we scale up the final wealth distribution to bring the aggregate wealth from our wealth series in line with aggregate national wealth totals. Again, this step is distributionally neutral and involves scaling-up all bracket averages and thresholds by a scaling factor. At the end of this scaling

[27] This comes directly from the power law property of the Pareto distribution – see Appendix D for more details.

[28] The choice of number of fractiles and rank is arbitrary. The rule of thumb is that the number of fractiles should not be very small. Blanchet et al. (2022) show that choosing more than ten fractiles has no discernible effect on the estimates.

exercise, we have full wealth distributions matched to national aggregates, but only for those years in which we have *both* surveys and rich lists. For all other years, we only have rich lists. We proceed by interpolating a survey-based wealth distribution for all missing years after 2002 by applying the annual growth rate of aggregate national wealth across the entire distribution.[29] Then, we adjust the top of the distribution in each year based on annual rich lists using steps 2 and 3 described above and then scale the corrected distribution to match aggregate national wealth.[30] This leaves us with annual wealth distributions for the years 2002-2023 that are consistent with the growth of national wealth aggregates over this period.[31]

For our final benchmark wealth series, we choose the Forbes billionaire rankings over the Hurun rich lists given the much longer time coverage (2002 vs. 2012). As it turns out, both inequality levels and trends are similar when using Forbes or Hurun (see Appendix Figure C.1). As mentioned above, we assume that the AIDIS is representative till $p99.9$ ($p0 = 0.999$) till 2012 and till the $p99.5$ ($p0 = 0.995$) 2018 onwards. We believe these are conservative choices. To demonstrate the magnitude to which this choice affects our estimates, we present top shares for 2018 for different $p0$ thresholds (0.999, 0.995, 0.990, and so on) in Appendix C.4.

We now present our results. At this point it is worth re-iterating that measurement of income and wealth inequality in India are fraught with various conceptual and empirical issues. Nonetheless, we believe these are not good enough reasons to give up on this important endeavour. We see our exercise as one step towards better understanding inequality dynamics in India using the best data sources at hand in a transparent and consistent manner. There is no doubt that better data and more democratic access to it can improve the quality of our estimates. We return to this issue later.

3 Income inequality in historical perspective

Figure 4 presents a summary of long-run income inequality dynamics in India over 1951-2022. The share of national income going to the top 10% fell from 37% in 1951 to 30% by 1982 after which it began steadily rising. From the early 1990s onwards, the top 10% share increased substantially over the next three decades, nearly touching 60% in the most recent years. At the other end of the

[29] For example, we take the 2012 survey distribution and apply growth rate of aggregate wealth between 2012 and 2013 to arrive at the survey-based wealth distribution for 2013.

[30] With extremely skewed distributions (as is very much case with wealth in India), a lot of the action in terms of inequality is often happening in the right tail and these annual top corrections using rich lists speak to that fact. Indeed, as Atkinson (2007) highlighted, letting S^* be the wealth share of a tiny (infinitesimal) group at the very top, the Gini coefficient of the distribution can be approximated as $S^* + (1 - S^*)G$, where G is the Gini coefficient for the rest of the population. See also Alvaredo (2011) on the relationship between top shares and the Gini coefficient.

[31] For 2023, we only have access to the full Hurun rich list but not the Forbes billionaire rankings. Hence, to arrive at 2023 Forbes-based estimates (and remain consistent with all years prior), we generate Hurun-based estimates and multiply them by a factor given by the ratio of Forbes and Hurun estimates for 2018-22 (when both are available). As we show below, the levels and trends from both Forbes and Hurun are very similar throughout our study.

distribution, the bottom 50% were getting only 15% of India's national income in 2022-23. Table 2 summarizes income inequality in 2022-23 and presents income levels and thresholds for different income groups. The top 1% earn on average 5.3 million, 23 times the average Indian (INR 0.23 million). Average incomes for the bottom 50% and the middle 40% stood at INR 71,000 (0.3 times national average) and INR 165,000 (0.7 times national average) respectively. At the very top of the distribution, the richest ~ 10,000 individuals (of 920 million Indian adults) earn on average INR 480 million (2,069 times the average Indian). To get a sense of just how skewed the distribution is, one would have to be at nearly the 90th percentile to earn the average income in India.

3.1 The rise of Indian billionaires - Top 1% and beyond

The availability of tax tabulations going back to 1922 since when the Income Tax Act was enacted by the British administration allows us to study the evolution of the top 1% income share over an entire century (Figure 5a). From 13% in 1922, top 1% share increased significantly to over 20% in the inter-war period. They then experienced a dramatic fall during the forties to fall back to 13% by the time of India's independence. After briefly rising during the 1950s, top 1% shares consistently fell over the next two decades and reached 6.1% by 1982. This was likely the consequence of the broadly socialist policy agenda pursued by the Indian government till the 1980s. This included nationalization of various important sectors (rail, air, banking, oil), strong regulation of markets, and high tax progressivity – top marginal tax rates were as high as 97.5% in 1973. This policy mix is likely to have reduced rent seeking-behaviour at the top of the distribution (see Chancel and Piketty (2019) for a broader discussion of the 1950-1980 period). Since the early-1980s, when the Indian government began initiating a broad range of economic reforms leading up to the liberalization in 1991, the decline in top 1% shares halted. From the early-1990s onward, top 1% shares have consistently increased over the next 30 years to reach an all-time high of 22.6% in 2022.

When we zoom-in within the top 1%, we see largely similar trends for the top 0.1%, top 0.01% and top 0.001% (Figure 5b). What is interesting are the magnitudes. For instance, in 2022, just the top 0.1% earned nearly 10% of the national income. The corresponding figures were 4.3% and 2.1% for the top 0.01% and top 0.001% respectively. What explains the sharp rise in top 1% income shares starting the early 1990s? Banerjee and Piketty (2005) note that public and private sector wage growth could have played a part till the late 1990s. There are good reasons to believe capital incomes likely play a role in the subsequent years. First, the capital share in the registered manufacturing sector has dramatically increased since the early 1990s (Abraham and Sasikumar, 2017; Jayadev and Narayan, 2020).[32] Second, as we demonstrate a little later, concentration of national wealth has also considerably grown during the last 3 decades. Third, the net wealth of the richest Indians (USD MER billionaires) as a share of India's national income started booming since the early 2000s (Figure 2b).

[32] It does not matter much for our argument that the registered manufacturing sector makes up only a tiny fraction of employment as long we expect the owners of these units to fall in the top of the income distribution.

3.2 Top shares vs. bottom shares

Broadly speaking, top 10% shares followed largely similar trends as top 1% shares. Between 1950-1980, top 10% income shares declined from nearly 40% at the turn of independence to just about 30% by 1982. In the wake of the liberalization reforms of 1991, top 10% shares started galloping and have reached astonishingly high levels by 2022, closing in on 60%. In other words, in the 4 decades between 1982 and 2022, the top 10% national income share has almost doubled. We find that till the turn of the 21st century, the share of national income going to the middle 40% remained firmly higher than the top 10% (Figure 6b). For instance, middle 40% shares and top 10% share were 42.8% and 36.7% respectively in 1951, and 44.1% and 33.5% respectively in 1990. However, starting the early 2000s, the top 10% overtook the middle 40% and by 2022, they stood at 27.3% and 57.7% respectively. The exact same story applies to bottom 50% and top 1% shares (Figure 6a); starting out at 22.4% and 10.5% respectively in 1990, bottom 50% and top 1% shares stood at 15% and 22.6% by 2022. Focusing on bottom 50% shares alone (Figure 7), we see they increased somewhat from 19% in the mid-1950s and then hovered around 22%-23% during the 1960-1980s. The share dropped by 2 percentage points in 1983, stabilized there for the next years, and then has consistently fallen since 1993. In the most recent years, we see a temporary marginal rise in bottom 50% shares in the years 2018-2020 after which they begin falling again. Similarly, top 10% shares marginally fall between 2018-2020 and rise after. As we discuss later, this was likely the result of the economic slowdown in India (that began in 2017 and culminated in the COVID-19 crisis in 2020) and the pro-cyclical nature of inequality.

There are likely numerous factors related to India's growth and development process that are keeping bottom 50% and middle 40% shares depressed. The lack of quality broad-based education, focused on the masses and not just the elites, is likely to be an important one. Inter-generational mobility measured using education rank has remained constant and low since liberalization (Asher et al., 2022). In 2011, when the last population census was conducted, nearly 30% of Indians remained illiterate. Bharti and Yang (2024) show educational inequality explains a quarter of wage inequality in India between 1988-2018. On the other hand, the services-led economic growth since liberalization has surely had unequalizing effects too (Fang et al., 2023). As per NSSO data, 45.5% of the workforce was employed in agriculture, 12.4% in construction, and only 11.6% in manufacturing, with the rest in services (Press Information Bureau, 2023). India's inability to pull more of its workforce away from agriculture towards more productive and better-paying employment remains a pressing challenge.

3.3 Growth incidence curves

Using our long-run harmonized income series it is possible to study the distributional consequences of different periods of growth in India since independence. We estimate the real growth rate of incomes at each percentile of the income distribution for 4 distinct periods: 1960-1980, 1980-2000, 2000-2014, and 2014-2022. These growth incidence curves are presented in Figure 8. While the levels of the curves are not strictly comparable, since the period lengths differ, the shape of the curves are still nonetheless informative about the *distributional* nature of economic growth. We see that over the 1960-1980 period, the bottom 90% experienced significantly higher growth than the top 10%. In fact, the top 1% experienced *negative* growth rates during this period, as low as - 46.5% and - 63% at $p99.9$ and $p99.998$. On the other hand, during 1980-2000, 2000-2014, and 2014-2022 we find that growth for the top decile has been significantly higher than the rest of the population. Further, even within the top 10%, we find growth rates rising with rank such that those at the very top benefited much more than the others. This explains the widening disparities in income shares that we observe starting 1980s and 1990s. It also explains the more-or-less absence of a middle class in India, with growth being extremely concentrated at the top (Chancel and Piketty, 2019). In fact, in the last years, 2014-2022, we find that the middle 40% seem to have grown slower than the bottom 50%.

4 Wealth inequality in historical perspective

Our wealth series begins in 1961 when AIDIS began. Between 1961-1991, we only have decadal estimates. From 2002 onwards, large enough rich lists and estimates of aggregate wealth allows us to generate an annual series. Figure 9 presents a summary of long run wealth inequality. Between 1961-1981, wealth shares of different groups remained stable. From 1991, we observe an uptick in wealth concentration which accelerates through the 2000s, much as with incomes. From 45% in 1961, top 10% wealth shares increased to 65% in 2022-23. On the other hand, both bottom 50% and middle 40% shares have significantly fallen over this period.

Table 3 summarizes wealth inequality in 2022-23 and presents wealth levels and thresholds for different wealth groups. The top 1% possesses an average of INR 54 million in wealth, 40 times the average Indian. The bottom 50% and the middle 40% hold INR 0.17 million (0.1 times national average) and INR 0.96 million (0.7 times national average) respectively. At the very top of the distribution, the wealthiest ∼ 10,000 individuals out of 920 million Indian adults own an average of INR 22.6 billion in wealth, 16,763 times the average Indian.

4.1 Top 1% and top 10% shares

A key feature of the wealth accumulation process in India is the extreme concentration at the very top. Between 1961 and 2023, the top 1% wealth share increased threefold, from 13% to 39% (Figure 10a). Most of these gains came post-1991 after which point top 1% shares have been on a steep upward trend right until 2022-23. Moreover, we find that wealth is highly concentrated even *within* the top 1% (Figure 10b). Consider this: in 2022-23, the top 1% wealth share was 39.5%, 29 percentage points went just to the top 0.1%, 22 percentage points to just the top 0.01% and 16 percentage points to just the top 0.001%.

In 1961, the top 10% wealth share was 45%. It declined by 1 percentage point between 1961 and 1971, the only decade when a decline is observed.[33] Between 1961 and 1981, top 10% shares did not change much. The same applies to top 1% and top 0.1% shares as well. This is perhaps not very surprising given that this was the era when socialist policies were at its peak and we see that the wealth concentration process was more-or-less brought to a stand-still. Post-1981, with the shift away from socialist policies towards market-based reforms, we find top 10% wealth shares consistently rise over the next three decades reaching 63% in 2012. Among other factors, the growing importance of financial assets between 1981 and 2002 (Appendix Table C.6) is likely to be an important contributor to these trends. There has been greater financialization of wealth as evidenced from a growing stock market (as a % of GDP), gains from which are bound to be restricted to a few. The SENSEX (S&P Bombay Stock Exchange Sensitive Index), a free-float market-weighted stock market index of 30 companies listed on the Bombay Stock Exchange, grew by 7300% between 1990 and 2023.

From 2012 onwards, the growth of top 10% shares seems to have slowed down over the next decade. In fact, between 2012 and 2018 (last two AIDIS rounds), top 10% shares declined marginally by 2 percentage points from 63% to 61%. This contrasts with the shares of the top 1% and beyond which have continued to rise even over the last decade. These trends are consistent with those reported in Anand and Kumar (2023), albeit at lower inequality levels.[34] Nonetheless, as it happens, our annual series suggests that 2018 was the year with the lowest top 10% shares during the entire last decade. In 2019, top 10% shares rose to 63% before falling to 62% in 2020 at the time of the COVID-19 crisis. After which, top 10% shares have reverted to an upward trend over the next three years and were at 65% in 2023.

4.2 Bottom 50% and middle 40% shares

We now turn to bottom wealth shares. The sharp rise in top 10% shares from 1991 onwards came at the loss of both bottom 50% and middle 40% shares. From stagnating at 11% between 1961 and 1981, bottom 50% shares first fell to 8.8% in 1991 and further to 6.9% by 2002 (Figure 12). After which, they have hovered between 6%-7% over the next two decades with no signs of recovery. In 1961, bottom 50% and top 1% shares were identical; by 2022-23, the top 1% share was more than 5 times larger (Figure 11a). Nonetheless, while extremely low, when placed in comparative perspective, bottom 50% shares in India are comparable to those than in China (6%), United Kingdom (5%), and France (5%) in 2022, and significantly higher than United States (1%).

[33] The piecemeal land redistribution that happened in some parts of the country could have contributed to this.

[34] This is due to methodological differences. They supplement AIDIS using Hurun rich lists (our benchmark relies on Forbes) to estimate a Pareto tail parameter (using the log-linear approach) but they assume a much lower cutoff point ($p0$) from which surveys are deemed non-representative. Their choice (INR 1 million) in 2012 leads to correcting ∼ the top 10% of the distribution. In our case, we only assume a power law to apply at the very top of the distribution, correcting the AIDIS survey only for the top 0.1% before 2018 and the top 0.5% thereafter.

In the post-liberalization years of high growth and rising inequality, the middle class (middle 40%) appears to have lost out significantly. This is partly since bottom 50% wealth shares were so low to begin with (9% in 1991) that the large gains made by the top 10% in the subsequent years could only come at the expense of middle 40% shares. Between 1961 and 1981, middle 40% and top 10% were nearly identical (between 40%-45%). Over, the next three decades, top 10% shares pulled ahead while middle 40% shares consistently fell to reach 31% by 2012, marginally rising to 32% in 2018 and have since fallen to 29% by 2023 (Figure 11b).

To make some sense of the factors underlying these trends, we decompose the wealth basket of the average Indian household over time using data from the AIDIS surveys, with the important caveat that the surveys do not capture the right-tail well (Appendix Table C.6). Two observations are worth making. First, physical assets dominate household wealth - land and building jointly make up almost 90% of the total household wealth, a share that has remained very stable over the last six decades (1961-2018). One possible reason for the slightly higher bottom 50% share in India than the other countries mentioned above could be that a majority of India's rural population owns at least a small piece of homestead land. Second, as already noted above, the share of financial assets in the wealth basket increased from 4% in 1981 to 10% in 2018. The estimates for recent years are likely to be under-estimates as surveys are known to capture financial wealth only imperfectly.

5 Wealth-to-income ratios in recent years

We now combine our income and wealth series in a couple of different ways to shed further light on inequality dynamics in India. To begin with, we compare the long-run trends of income and wealth inequality (Figure 13a). Overall, top 1% and top 10% income and wealth shares track each other quite closely over the entire period of our study. Given that the two series are estimated from entirely different sources and yet exhibit such similar trends suggest that the steep rise in inequality we observe starting the mid-1990s is indeed structural and related to the underlying growth process, and not merely an artefact of improving tax compliance, reduction in marginal tax rates, or other such factors that may have also coincided since the 1990s (but none of which affect our wealth inequality estimates since we do not use tax data).

A couple of interesting facts emerge when we zoom-in and compare income and wealth inequality in the last two decades (Figure 13b). At the turn of the 21st century, in 2002, top 10% wealth shares stood at 57.1% while the top 10% income share was 42.1%, a gap of 15 percentage points. Over the next 20 years, top 10% income shares rose faster and the gap reduced considerably, effectively halved to 7.3% by 2022. On the other hand, we observe the exact opposite trend when comparing top 1% income and wealth shares over the same period. In 2002, the top 1% wealth share was 25.4% compared to 16.7%

for incomes, a gap of 8.7 percentage points. Twenty years later, by 2022, the top 1% wealth share had reached 40.1% compared to 22.6% for incomes, a gap of 13.4 percentage points. These opposing trends suggest that wealth concentration is accelerating relatively faster than incomes at the very top of the distribution.

While aggregate wealth-to-income ratios (β) at the national-level have been estimated for Western economies and India (Piketty and Zucman, 2014; Kumar, 2019), how these ratios vary across the distribution has received relatively little attention. We combine our wealth and income distributions, appropriately anchored to aggregate wealth and national income respectively, to estimate wealth-to-income ratios across the distribution.[35] Figure 14 compares the results for 2002 and 2022. In general, we find wealth-to-income ratios rise as we move up the distribution, very sharply so as we reach the very top (top 1%). At the bottom, they are very low. In 2002, at the 10th percentile of the wealth distribution, wealth made up only around 30% of incomes. At the other end of the distribution, at the 99$^{\text{th}}$ (99.999$^{\text{th}}$) percentile, wealth made up 367% (2700%) of incomes. Between 2002 and 2022, in line with trends in aggregate β, we find that wealth-to-income ratios increase at each point of the distribution, except for a few percentiles in the top decile. We also observe the largest increase for the top 0.1% whose wealth-to-income ratios have nearly doubled over the 20 years. In 2022, at the very top of the distribution, wealth stood at over 394% (4600%) of incomes at 99$^{\text{th}}$ (99.999$^{\text{th}}$) percentile.

Our results are very much in line with and complement those by Singh (2023) who uses a novel database of wealth and income disclosures from Indian politicians to estimate wealth-to-income ratios across the wealth distribution. By forensically comparing incomes and wealth reported on these disclosures (for an admittedly non-representative sample of the population), he finds significant "missing incomes" for the very wealthy that are not reflected in income tax returns. Consequently, our income inequality estimates relying on income tax returns for top incomes are bound to underestimate top shares. More importantly, both sets of estimates paint the Indian income tax schedule in new light. While traditional wisdom has it that the income tax system is progressive, these estimates suggest that it might well be *regressive* from the point of view of wealth.[36] In particular, the tax liability as a share of net wealth could be falling as we move up the wealth distribution. This is good reason why India needs a re-structuring of its tax code to account for both income and wealth. A "super tax" on the very wealthy might be a good place to start. Not only would it serve as a tool for fighting the growing inequalities we are observing today, but it would also provide additional fiscal space for the Indian government to enhance spending on essential social expenditures (health, education, nutrition) which have historically been low compared to global standards, including other countries at similar income levels (Drèze and Sen, 2013). Broad-based public investments are going to

[35] For $p \in (0,1)$ denoting fractiles, we estimate wealth-to-income ratios as $W(p)/Y(p)$, where $W(.)$ and $Y(.)$ are the quantile functions associated with wealth and income respectively.

[36] Our results are tentative given we do not observe both income and wealth for the same set of individuals and instead draw inferences based on the full distributions of income and wealth we estimate.

be crucial if much of the Indian population is to benefit from the current wave of hyper-globalization. A tax of just 2% on the total net wealth of the 162 wealthiest Indian families in 2022 would yield revenue to the tune of 0.5% of national income (more than twice the central government's budget expenditures on the National Rural Employment Guarantee Act in recent years).

6 Indian inequalities in comparative perspective

We now place income and wealth inequality levels in India in 2022 in global perspective by comparing India with Brazil, China, France, South Africa, United Kingdom, and the United States. Looking at top 10% income shares, India stands second only second to South Africa (Figure 15a). If we look at top 1% shares, however, India has the highest levels at 22.6%. As it happens, India's top 1% income share appears to be among the very highest in the world based on WID data, behind only perhaps Peru, Yemen and a couple of other small countries. In terms of top wealth shares (Figure 15b), we see that both with top 10% and top 1%, India comes out in the middle of the pack, with Brazil and South Africa standing out with their extreme wealth concentration levels (85.6% and 79.7% top 10% shares respectively). In terms of both income and wealth inequality, UK and France have the most equal distribution in the handful of countries.

An India vs. China comparison in the inequality sphere is both justified and revealing given their comparable population levels and that the two countries started out at roughly similar income and development levels in the 1950s and 1960s. Between 1970-1990, Chinese incomes began growing faster than Indian incomes and then they skyrocketed both in absolute and relative terms starting the early 2000s (Figure 1b). While growth of Indian incomes did pick-up, especially in the 2000s, they were much lower than Chinese growth rates. How did the dynamics of income and wealth inequality play out in these two countries? Comparing income and wealth inequality in India and China between 1980 and 2022 presents some interesting findings. Focusing on incomes first, top 10% and top 1% income shares were comparable in the two countries in 1980 (Figure 16a). Over the next 4 decades, we identify two key breaks in trends. Between 1980 and the early 1990s, top income shares gradually increased in both countries. The first trend break was in 1993 when a small gap opens between Indian and Chinese top shares, particularly so for top 1% shares, on the back of the sharp rise in inequality in India post the liberalization reforms of 1991. The second trend-break came in 2006 when Chinese top shares stabilize while Indian top shares continued to grow, creating a wide gap particularly for top 10% shares.[37] By 2022, top 1% income shares in India were nearly 50% larger than those in China (22.6% vs 15.7%) and top 10% shares were nearly 35% larger (57.6% vs. 43.4%). While these results are a worrying sign for India, the flip-side is that China shows it is possible for

[37] Kanbur et al. (2021) argue that the stabilization of Chinese top shares was likely the outcome of a mix of policy measures and a slow-down in structural transformation. At the same time, Piketty et al. (2019) note that data limitations for China make the estimates for more recent years tentative.

low- and middle-income economies to achieve high growth without generating the obscene income inequality levels observed in India today. They also show that a mixed economy like China's, with an important role for the public sector, can be conducive to high growth and moderate inequality, contrasted with India's experience of moderate growth coupled with extreme inequality.[38]

When looking at wealth concentration instead of income over the same period, we interestingly observe the exact opposite trend as for incomes (Figure 16b). In 1995 (when the Chinese wealth series begins), top 10% and top 1% wealth shares in China were considerably lower than India's. Between 1995 and 2006, top wealth shares in China steadily increased and caught-up with Indian levels, and from 2011 onwards top 10% wealth shares have been higher in China than India. On the other hand, top 1% shares in India have so far managed to keep ahead of China given the extreme concentration of wealth in India. The overall trends of rising top 10% wealth shares in China are possibly the consequence of its wealth-to-income ratio more than doubling between 1980 and 2022 (from 3.75 to 9.41), which in turn was largely the result of rising savings rates and asset prices (Piketty et al., 2019).

7 An evaluation of the last decade

The last decade in India has seen major political and economic developments. In 2014, the right-wing Bharatiya Janata Party (BJP) came to power at the centre with a sweeping majority after nearly un-interrupted rule by the Indian National Congress (INC) since independence. While the BJP were voted into power on a mandate of development and economic reforms, many observers believe that over its two terms, it has led an authoritarian government with centralization of decision-making power (Jaffrelot, 2021) coupled with a growing nexus between big-business and government (Banaji, 2020). We begin by briefly discussing the macroeconomic performance since 2014 before moving to inequality trends in recent years.

7.1 A grim macroeconomic picture

Official statistics suggest rather sluggish economic growth during the Modi years - real year-on-year growth rates of incomes fell from over 6% in 2015 and 2016 to 4.7% and 4.2% in 2017 and 2018 and then dramatically to 1.6% in 2019 (Figure 17a). All this was before the COVID-19 pandemic hit and incomes *fell* by 9% in 2020. There was a base-year effect in 2021 and incomes grew by 4.7% in 2022. What was driving the declining growth even before COVID? Estimates based on official statistics suggests that savings and investment rates steadily fell for over a decade till 2017-18, exports began

[38] Bharti and Yang (2024) argue that higher economic inequality in India can be at least partly attributed to differences in the way their education systems developed. The Indian system expanded in a top-down fashion (neglecting primary-level mass education for a long time) with lower diversification at the top (lower vocational graduates and a larger share of humanities graduates among college graduates). As a result, India has higher education inequality, contributing to 25% of the total wage inequality from 1988-2018 compared to 2% in 1988 and 12% in 2018 in China.

falling 2014-15, and the share of manufacturing and industry in the GDP stagnated between 2013 and 2018 (Nagaraj, 2020). Unemployment rates, especially among the youth (15-29 years), increased considerably between 2011-12 and 2017-18 (Ghatak and Mukherjee, 2019). Real wages across various sectors have more-or-less stagnated over the last decade or so (Drèze, 2023; Das and Usami, 2023). Another possible factor contributing to the economic slowdown was the harsh "demonetization" shock dealt to the economy in November 2016 when nearly 86% of the currency in circulation ceased to be legal tender overnight. While the move was supposedly aimed at fighting "black money" (unaccounted incomes) stored as currency notes, it is believed to have disproportionately hurt the informal sector, small-medium businesses, and the poor, with one set of estimates suggesting that short-term GDP fell by 2 percentage points (Chodorow-Reich et al., 2019).[39] The Modi government has certainly invested in expanding the coverage of various infrastructural benefits like housing, toilets, electricity, and banking, what some have called the "new welfarism of India's right" (Anand et al., 2020). But it is unclear if these investments have led to an improvement in purchasing power on the market. Further, in the absence of any NSSO consumption surveys in recent years, it is unclear how much progress India has made in reducing extreme poverty.[40]

7.2 Inequality dynamics, 2014-2022

In terms of inequality dynamics, the Modi years of 2014-2023 can be divided into 3 phases: 2014-2017, 2018-2020, and 2021 onwards. In the first phase, the economy was growing moderately fast and both income and wealth inequality continued to rise. In the second phase, from 2017-18 to 2020, growth slows down considerably and then plummets in 2020. In this second phase, we see top 10% income and wealth shares decline by 1-2 percentage points. The most likely explanation for which is the pro-cyclical nature of inequality, i.e. the rich tend to benefit disproportionately from boom periods and are disproportionately hurt during slumps, as Ghatak et al. (2022) also argue.[41] This seems the most likely explanation especially given that we observe similar trends for both income and wealth during this phase. Moreover, as shown in Figure 2b, the wealth of the richest Indians as a share of national income also declined between 2018 and 2020. It is hard to think of other factors that concomitantly explain these trends for both incomes and wealth.[42] Measurement error is of course a possibility, which we address shortly. Finally in the last phase, after the lock-downs were lifted

[39] As it happens, credible evidence suggests that in the 2019 general elections, BJP's vote shares were lower in areas that experienced harsher effects of demonetization (Khanna and Mukherjee, 2023). We also know from various news reports that 99.3% of the demonetized currency subsequently returned to the banking system. This implies either that the government's claims about black money in the form of large currency notes were entirely misplaced or that this erstwhile black money was effectively legalized (without having to pay appropriate taxes) via one means or another during the rather chaotic and confusing currency deposition process that followed.

[40] A preliminary "fact sheet" from the 2022-23 round of the consumption expenditure survey has just been made public. However, there are already concerns of comparability - see footnote 11 and National Sample Survey Office (2024, p. 4).

[41] Gupta et al. (2021) also find that income and consumption inequality declined during COVID-19 owing partly to the high covariance of capital incomes with aggregate fluctuations, which too relates to our argument.

[42] For the case of incomes, one possible factors could be the introduction of the PM-KISAN (Pradhan Mantri Kisan Samman Nidhi) scheme for farmers in 2018.

and the economic effects of COVID-19 dissipated, we find top shares revert to their upward trend in 2021 and 2022, while bottom shares decline back to their 2014 level. By examining the growth incidence curve for incomes and wealth between 2014-2022, we find that the real beneficiaries in the recent years appear to be the super-rich, the top 1% and beyond (Figure 17b). This is particularly so for wealth concentration at the very top. This lends some support to political economy assessments that have characterized the economic system in India in recent years as "conglomerate capitalism" (Damodaran, 2020) and a "conclave economy" (Bardhan, 2022).

The other interesting aspect to note is that both with income and wealth, the middle 40% seem to have grown slower than the bottom 50% during this period. This is likely to exacerbate the phenomenon of India's "missing middle class" (Chancel and Piketty, 2019). The bottom 50%, in return, grew at the same rate as that of the average of the population, preventing an increase in their share of total income and wealth. We must, however, emphasize that this result might be overly conservative. Recent ICE360 data indicates a significant decline in the bottom 50% income share over the period 2015-16 to 2020-21 (see Appendix Figure B.3), with growth rates well below the average. Enhanced access to household survey and administrative data sources is essential for a deeper understanding of these dynamics, as discussed subsequently.

7.3 Growing data challenges in recent years

During the last decade, various key data sources in India have either become unavailable or their quality has become suspect. This applies to all the key inputs that go into our inequality series: national income accounts, tax tabulations, and surveys. We briefly discuss these issues with the aim of drawing caution when interpreting the estimates for recent years.

National income accounts: Various concerns have been raised about validity of India's national income accounts data in recent years.[43] At least two detailed empirical exercises, one by an ex-chief economic advisor to the Government of India, point to possible over-estimation of GDP in the years post-2011 (Morris and Kumari, 2019; Subramanian, 2019a). Some concerns have also been raised regarding the possible mis-measurement of India's GDP deflator (Subramanian and Felman, 2023). More generally, the dated nature of the underlying data used to estimate GDP is very concerning – key inputs like the CPI, WPI, input-output tables, industry codes, consumption expenditure, etc. are currently based on data that might be 10-15 years old (Sapre and Bhardwaj, 2023). This is especially a worry for aspects relating to the informal sector of the economy. If it is indeed the case that GDP is being overestimated in recent years, that would imply that our inequality estimates would be slightly downward-biased.[44]

[43] As India's ex-chief statistician clarified recently, the issue with India's GDP estimates in recent years seems to be less about methodology and more about the severely outdated underlying data and unreliable proxies (Thapar, 2023).

[44] This is because (70% of) per-adult net national income serves as the "control average" for the generalized Pareto

Tax tabulations: The British colonial administration introduced an individual income tax with the Income Tax Act, 1922. Since then, data on individual incomes began being collected and the colonial administration published this data in tabulated form on an annual basis. This practice which was continued by the Indian government post-independence. Between 1922 and 1998, annual publication of these 'All-India Income Tax Statistics' provided a vital source of information on top incomes, mobilizing which Banerjee and Piketty (2005) estimated the share of national income going to the top 0.01%, 0.1% and 1% during this period. There were naturally improvements to the methodology used to generate these tabulations over the years (partly owing to technological and computational improvements), but systematic and regular release of this data was not disrupted. However, starting 1999 onwards, the government of India strangely stopped publishing these tax tabulations for reasons that remain unknown. For a whole decade when India experienced strong macroeconomic growth (2000-2010), no tax tabulations are available to date. Then in 2016, the government retrospectively released data but only starting 2011. For the next few years, data releases continued till retrospective data for 2017 was out, after which once again no tax tabulations were available. Finally in mid-2023, the government again retrospectively released data for the years 2018-2021. In short, the release of tax data has been highly erratic and incomplete in recent decades. The reason for this remains unclear. One possibility is that the analysis and release of tax data falls low on the priority list of the Income Tax department. This stands in sharp contrast to the past when, for instance, government appointed committees specially provided recommendations on ways to better analyze and report data from income tax returns.[45] Besides releasing all-India tabulations, the income tax department also used to release state-wise tabulations till 1998. These could potentially allow going beyond all-India analysis and shed light on the evolution of top incomes and inequality at the *state-level*. Given the size and population of individual states, larger than many European countries in many cases, this is an important endeavor. However, starting 1990 (to the best of our knowledge), state-wise statistics have not been released at all, even post-2011 when all-India statistics have been released. The non-availability of state-level data in recent years is strange, not only because it used to be released regularly before, but also because computerization and digitization of records in recent decades should make dis-aggregation and tabulation of returns at the state-level easier than before. This leaves the estimation of state-level income and wealth inequality an incomplete endeavor.

Income and consumption surveys: One of the key challenges when updating the income inequality series for the last decade is the absence of a comparable NSSO consumption survey after 2011-12.

interpolation algorithm used to extract a distribution of top incomes from the tax tabulations. A lower control average would mechanically increase top income shares. To what extent this issue affects our estimates depends on the extent to which national income is being over-estimated.

[45] As an example, the 'Committee on Direct Tax Statistics' recommended using a part-sampling and part-census approach for generating tabulations of income-tax statistics from 1974-75 onwards - all returns with incomes above INR 25,000 were to be covered by a census while those with incomes below INR 25,000 were to be sampled, with most states assigned a 10% sample, some 20%, and a full census in some union territories like Delhi (Directorate of Inspection, 1978). Incidentally, this was a time when the government of India was explicitly interested in curtailing the power of the elites.

As noted earlier, NSSO has historically steered clear of measuring incomes and instead focusing on consumption expenditures. Consequently, our measurement framework also relies heavily on these consumption surveys. The NSSO did conduct a round in 2017-18 but it was suppressed by the government. From 2017-18 onwards, the PLFS came to our rescue. As it turns out, even though it is primarily designed for labour market outcomes, it collects preliminary data on 'usual' consumption expenditures. By correcting these for comparability with past NSSO CES rounds, we are able to extend our income inequality series on an annual basis from 2017-18 onwards. However, this involves a correction that is bound to be only imperfect at best. This creates an additional degree of uncertainty around our estimates in the recent years. More importantly, we find that alternate data sources present contradictory trends for bottom incomes. For instance, based on percentile-level growth rates of per-capita incomes between 2015-16 and 2020-21 in the ICE360 survey, we find a steeply upward sloping growth incidence curve such that bottom 50% shares would decline from 14.4% in 2015 to 9.8% in 2020 (Figure B.3). This stands in sharp contrast to the trends in our benchmark series which suggest a relatively stable bottom 50% share over this period, besides a temporary and marginal increase (1 percentage point) between 2018 and 2020. Therefore, we see our benchmark estimates as a conservative scenario until better data emerges to improve our estimation.

Wealth surveys: It is also worth mentioning a couple of concerns relating to NSSO AIDIS that forms the basis for our wealth inequality series. First, as highlighted earlier, it appears that the issue of under-estimation at the top has worsened over the last three successive rounds in 2002, 2012 and 2018 - the total (net) wealth of USD MER billionaires in the Forbes list as a percentage of the total survey wealth increased from 1.26% in 2002 to 2.74% in 2012 to 6.01% in 2018. The issue of under-estimation and under-representation of the very rich and wealthy in sample surveys is not unique to India but the fact that the issue is getting worse over time deserves closer attention by the NSSO. More stratification and purposive over-sampling at the top could be ways to counteract the current trend of increasing non-representativeness of the right tail. Further, with all its surveys (CES, AIDIS, PLFS, etc.), NSSO should release non-response rates by some variable like, say, the 'usual' consumption expenditure variable that it could collect at the household listing stage - this would allow decomposing the non-representativeness of surveys for the right tail more clearly into response-related and measurement-related issues. It is also worth highlighting that we are likely to under-estimating wealth at the top of the distribution due to off-shore wealth. Of the total foreign-owned off-shore real-estate in Dubai, 20% is owned by Indians (Alstadsæter et al., 2024), amounting in total value to to 1.1% of India's GDP (Alstadsæter et al., 2022). The second concern relating to AIDIS relates to the timing of the release of the latest round of the data. Starting in 1961-62, these surveys were meant to be decennial surveys and indeed they were conducted every 10 years, in 1971-72, 1981-82, 1991-92, 2002-03 and 2012-13. It is unclear why the last round was conducted within a shortened gap of 6 years in 2018-19. If this is part of a broader plan of more regular AIDIS rounds, then it is a welcome change. If, on the other hand, this was the result of political considerations, then there is a cause for worry. Coincidentally (or not), our estimates suggest that the top 10% wealth share may have been at its lowest during the last decade in 2018 (Appendix Table C.1).

7.4 Call for democratic access to data

Tax tabulations: As we have stressed above, the publication of tax tabulations has been highly erratic and inconsistent in India between 1999 and 2022-23. Our first call is for stability and continuity: it is high time that India's tax administration publishes a consistent set of tax tabulations *every year* that can be used by the academic community and the civil society to ascertain the evolution of inequality in India (and also the functioning of its tax system). Ideally, it would be even better if researchers could additionally also have access to anonymous micro files based upon tax declarations, as they do in other democratic countries, and hopefully this time will come in India. In the meantime it is particularly critical to publish consistent and yearly tax tabulations, which play an important role even in countries where micro files are available.

Enriching tax tabulations: In addition, existing tax tabulations should be enriched. In recent years, the tabulations provide the total number of tax returns filed and total incomes reported for 26 income brackets, starting from 1.5 lakh going all the way to 500 crore and beyond. These tabulations are reported separately for 5 categories of tax filers of which we focus on individuals and Hindu undivided families (HUF) for our purposes. For both individuals and HUF, bracket-wise tabulations for gross-incomes are provided which we use for our analysis. In addition, however, separate (un-linked) tabulations are provided for different sources of income (salary, house property, business, capital gains, interest, other sources, and losses set off). While these separate tables are meaningful, they are individually not particularly helpful in understanding the dynamics of top incomes. The reason is because we do not know if individuals in the highest income bracket of gross incomes are the same individuals in the highest income bracket of salary incomes (or business incomes or property incomes, and so on). In other words, we do not know the break-up of gross incomes by the various sources. This severely limits the ability to shed light on the sources of income dynamics at the top of the distribution. In an ideal scenario, we would have access to a sample of appropriately anonymized micro-data of tax returns. This would allow a much greater degree of flexibility and richness to the analysis. However, if that remains unacceptable, then in addition to publishing separate tables with distributions of incomes from different sources, it would be very helpful to also publish a single all-encompassing "composition table". This would, for each bracket of *gross incomes*, provide the break-up of gross incomes by different sources. In addition, if provided with the break-up of returns by gender in each bracket of gross incomes, researchers would be able to shed sharper light on gender inequality at the top of the distribution.[46] An example is shown in Appendix Table E.1.

[46] To the best of our understanding, tax authorities do not collect information on religion or caste on the tax returns.

Inclusion of wealth information in tax tabulations: Finally, it would be highly desirable if tax tabulations could include information on not only income flows but also wealth stocks and asset ownership. Even in the absence of a comprehensive and well-administered wealth tax, which would be the ideal situation, it is common practice for tax administrations in most countries to collect a large quantity of information about wealth. This includes information on real estate assets through the registration of real-estate transactions. This also includes information on company ownership and financial portfolios that should be automatically transmitted from banks to tax administrations in order to properly control and audit the taxation of financial income flows. One problem in most developing countries – as well as in many developed countries – is that this information is typically not used in a systematic manner. We recommend that India's tax administration releases, along with its income tax tabulations, a detailed set of wealth tabulations indicating the numbers of asset holders and amounts of their assets for a large number of wealth brackets and asset classes (real estate, equity, bonds, etc.). It is sad that researchers and citizens alike resort to wealth rankings published by magazines in order to study wealth patterns in India, especially given that such rankings involve little transparency about the concepts and methods they use (and rely in practice on limited and unsystematic information sources). Unfortunately, this situation is likely to persist as long as public authorities do not start fulfilling their role as providers of reliable, transparent and systematic statistical information.

Surveys: To the Government of India and NSSO's credit, the Periodic Labour Force Surveys (PLFS) has been conducted annually starting 2017-18 as a replacement to the earlier NSSO Employment-Unemployment Surveys. More importantly, the data has been released regularly without interruption since the beginning. However, the absence of information on capital incomes is a serious drawback with the PLFS. We urge the Indian authorities to consider including questions in the PLFS that cover all key sources of incomes. Besides allowing better use of PLFS for inequality measurement, this would surely also enhance the ability of PLFS to shed light on labour market trends. Various large-scale surveys conducted by non-governmental organizations over the last two decades have collected data on incomes (IHDS, CPHS, ICE360) and it is perhaps time the NSSO also catches up with this trend. At this point, it is worth noting that various opposition parties have been demanding a caste-based census in India in the run up to the 2024 elections. This is indeed an important demand worth rallying around. More than 75 years after India's independence, we still do not have a carefully quantified understanding of the interplay of economic inequality with social factors, of which caste is undoubtedly most important. The state of Bihar recently conducted a socio-economic caste census and released some findings in the public domain. There are indications that the Government of Jharkhand is also preparing for a similar survey. These are positive signs. Access to anonymized micro-data or detailed tabulations from these surveys would provide a wealth of information to study a range of key policy-relevant questions. Ultimately, there is an urgent need for more democratic access to statistical data that is free from political interference.

7.5 A brief methodological defence

Since the release of the initial income inequality results in Chancel and Piketty (2019), some objections have been raised regarding certain aspects of our methodology. We take this opportunity to address some of them. One concern is that since we use surveys for the bottom 90% and tax data for the top incomes, our methodology "… rules out—by assumption—any under-estimation (or under-reporting) of incomes by the bottom 90% of the population" (Kundu, 2017). A lot of ink is spilt by Aiyar (2017) in creating a similar impression. Except, this is simply false since we make no such assumption. Instead, what our method assumes is that across the entire distribution, the same fraction of gross incomes is under-reported or hidden.[47] How likely is it that this assumption holds? On the one hand, surveys may indeed be under-estimating incomes and consumption simply because it is easy to hide from surveyors and households have incentives to do so for various reasons. On the other hand, as noted earlier, there is evidence that tax tabulations also miss a non-trivial fraction of incomes as we reach the top of the wealth distribution (Singh, 2023). Thus, we have no *prima facie* reason to believe surveys necessarily under-estimate a *larger* fraction of incomes than tax data. On the contrary, if the super-rich can tightly control the amount of incomes reported in their tax returns (for example through income shifting), then we may be severely under-estimating incomes at the very top.[48]

A more valid concern that has been raised is that the sharp rise in inequality we observe in recent decades may reflect the reduction in tax rates and the improved ability of the tax authority to track incomes, and not necessarily a structural increase in inequality (Aiyar, 2017; Kundu, 2017). This is certainly possible. However various pieces of evidence suggest that the rise in inequality we observe is real. First, as we demonstrate earlier in the paper, despite being estimated from entirely different sources, we observe a concomitant rise in wealth and income inequality over the same period (Figures 13a and 13b). Second, for the rise in inequality in the 1980-2000s period to be attributable to changes in income tax rates, the elasticity of incomes would need to be enormously (and unrealistically) high. Third, as Banerjee and Piketty (2005) show, the rise in inequality starting the mid-1980s and 1990s shows up not just for gross incomes, but also for wages which are presumably less susceptible to tax evasion and less sensitive to improvements in tax surveillance. For these reasons we believe that the broad trends captured by our series are robust. However, given the very nature of this exercise, there is some degree of uncertainty about the exact *levels.* This is reflected in the 54 different long-run variants tracked in Chancel and Piketty (2019) accounting for various combinations of assumptions along the way. With better and more transparent data, there is no doubt these series can be improved upon. In the meanwhile, we see our exercise as one step towards a clearer and more transparent understanding of the dynamics of income and wealth inequality in India over the long run.

[47] Moreover, this fraction is allowed to vary across years. However, in each year, the same fraction is assumed to apply to the full distribution.

[48] For instance, the compensation of the Chairman and Managing Director of Reliance Industries Limited reportedly remained *fixed* at INR 15 crore (150 million) for over 12 years (Khera and Yadav, 2020).

8 Conclusion

We combine national income accounts, wealth aggregates, tax tabulations, billionaire rankings, rich lists, and surveys on income, consumption, and wealth in a consistent framework to present long-run homogeneous income and wealth inequality series going back till 1922 for incomes and 1961 for wealth. Our estimates suggest that inequality levels declined post-independence till the early 1980s, after which both top income and wealth shares began rising and have skyrocketed since the early 2000s. Trends of top income and wealth shares closely track each other over the entire period of our study, including the most recent decades. By 2022-23, top 1% income and wealth shares are at their highest historical levels at 22.6% and 40.1% respectively and India's top 1% income share is among the very highest in the world, higher than even South Africa, Brazil and US.

By combining our wealth and income series, we also find tentative evidence that wealth-to-income ratios in the recent years could be as low as 30%-40% at the lower end of the wealth distribution compared to over 4600% at the very top of the distribution. In line with earlier work, this suggests that the Indian tax system might well be regressive when viewed from the lens of net wealth. This is perhaps good reason to consider a restructuring of the tax code to account for both income and wealth. Moreover, broad-based public investments in health, education and nutrition are needed to enable the average Indian, and not just the elites, to meaningfully benefit from the ongoing wave of globalization. Besides serving as a tool to fight inequality, a "super tax" of 2% on the net wealth of the 167 wealthiest families in 2022-23 would yield 0.5% of national income in revenues and create valuable fiscal space to facilitate such investments.

As per our benchmark estimates, the Billionaire Raj headed by India's modern bourgeoisie is now more unequal than the British Raj headed by the colonialist forces. One reason to be concerned with such high levels of inequality is that extreme concentration of incomes and wealth is likely to facilitate disproportionate influence on society and government. This is even more so in contexts with weak democratic institutions. After *largely* being a role model among post-colonial nations in this regard, the integrity of various key institutions in India appears to have been compromised in recent years. This makes the possibility of India's slide towards plutocracy even more real.[49] If only for this reason, income and wealth inequality in India must be closely tracked and challenged.

[49] It would only be fair to say that the wealthy already wield excessive power and influence in India. This was on unabashed display in the city of Jamnagar in Gujarat leading up to the "pre-wedding" celebrations of the youngest son of India's (and Asia's) richest family. Take the case of Jamnagar's airport - not only was the primarily defence air station given a temporary 10-day international airport status to accommodate foreign flights, but it appears that the Indian Air Force also permitted access to its sensitive areas and provided additional military personnel to staff the air traffic control tower (Chandra, 2024).

References

Abraham, V. and Sasikumar, S. K. (2017). "Declining wage share in India's organized manufacturing sector: Trends, patterns and determinants". ILO Asia-Pacific Working Paper Series.

Aiyar, S. S. A. (2017). "Piketty's Botched Analysis of Inequality in India". Economic Development Bulletin No. 28, CATO Institute.

Alstadsæter, A., Godar, S., Nicolaides, P., and Zucman, G. (2024). "Global Tax Evasion Report 2024". EU Tax Observatory.

Alstadsæter, A., Planetrose, B., Zucman, G., and Økland, A. (2022). "Who Owns Off-Shore Real Estate? Evidence from Dubai". EU Tax Observatory Working Paper No. 1.

Alvaredo, F. (2011). "A Note on the Relationship between Top Income Shares and the Gini Coefficient". *Economic Letters*, 110(3):274–277.

Anand, A., Dimble, V., and Subramanian, A. (2020). "New welfarism of India's right". *The Indian Express*, 22 December.

Anand, I. and Kumar, R. (2023). "The Sky and the Stratosphere: Wealth Concentration in India During the Last (Lost) Decade". *Review of Income and Wealth*.

Anand, I. and Thampi, A. (2016). "Recent Trends in Wealth Inequality in India". *Economic and Political Weekly*, 51(50):59–67.

Asher, S., Novosad, P., and Rafkin, C. (2022). "Intergenerational Mobility in India: New Measures and Estimates Across Time and Social Groups". *American Economic Journal: Applied Economics*.

Atkinson, A. B. (2007). "Measuring Top Incomes: Methodological Issues". In: Atkinson, A. B. and Piketty. T. (eds.), "Top Incomes over the Twentieth Century: A Contrast Between European and English-Speaking Countries", Oxford University Press, Oxford.

Atkinson, A. B., Piketty, T., and Saez, E. (2011). "Top Incomes in the Long Run of History". *Journal of Economic Literature*, 49(1):3–71.

Banaji, J. (2020). "Indian Big Business: The evolution of India's corporate sector from 2000 to 2020". *Phenomenal World*, 20 December.

Banerjee, A. and Piketty, T. (2005). "Top Indian Incomes, 1992-2000". *World Bank Economic Review*, 19(1):1–20.

Bardhan, P. (2022). "The 'New' India: A Political-Economic Diagnosis". *New Left Review*, 136 July/August.

Bharti, N. K. (2018). "Wealth Inequality, Class and Caste in India, 1961-2012". WID.world Working Paper No. 2018/14.

Bharti, N. K. and Yang, L. (2024). "The Making of China and India in 21st Century: Long Run Human Capital Accumulation from 1900 to 2020". WIL Working Paper (forthcoming).

Blanchet, T. (2017). “Prices and currency conversion in WID.world”. WID.world Technical Note Series No. 2017/3.

Blanchet, T., Chancel, L., Flores, I., and Morgan, M. (2017). “Distributional National Accounts Guidelines: Methods and Concepts Used by the World Inequality Database”. World Inequality Lab.

Blanchet, T., Fournier, J., and Piketty, T. (2022). “Generalized Pareto Curves: Theory and Applications”. *Review of Income and Wealth*, 68(1):263–288.

Bourguinon, F. (2018). “Simple adjustments of observed distributions for missing income and mission people”. *Journal of Economic Inequality*, 16:171–188.

Chancel, L. and Piketty, T. (2019). “Indian Income Inequality, 1992-2015: From British Raj to Billionaire Raj?”. *Review of Income and Wealth*, 65(S1):S33–S62.

Chandra, J. (2024). “Jamnagar airport gets international status for Anant Ambani’s pre-wedding bash”. *The Hindu*, 1 March.

Chavan, P. (2012). “Debt of Rural Households in India: A Note on the All-India Debt and Investment Survey”. *Review of Agrarian Studies*, 2(1).

Chodorow-Reich, G., Gopinath, G., Mishra, P., and Narayanan, A. (2019). “Cash and the Economy: Evidence from India’s Demonetization”. *The Quarterly Journal of Economics*, 135(1):57–103.

Damodaran, H. (2020). “From ‘Entrepreneurial’ to ‘Conglomerate’ Capitalism”. *Seminar*, 734.

Das, A. and Usami, Y. (2023). “Downturn in Wages in Rural India”. *Review of Agrarian Studies*, 13(2).

Deaton, A. and Drèze, J. (2002). “Poverty and Inequality in India: A Re-Examination”. *Economic and Political Weekly*, 37(36):3729–4748.

Deaton, A. and Kozel, V. (2005). “Data and Dogma: The Great Indian Poverty Debate”. *The World Bank Research Observer*, 20(2):177–199.

Directorate of Inspection (1978). “State-wise Income Tax Statistics 1974-75”. Directorate of Inspection (Research, Statistics and Publications), New Delhi.

Drèze, J. (2023). “Wages of distress”. *The Indian Express*, 13 April.

Drèze, J. and Sen, A. K. (2002). “India: Development and Participation”. Oxford University Press, New Delhi.

Drèze, J. and Sen, A. K. (2013). “An Uncertain Glory: India and its Contradictions”. Penguin India, New Delhi.

Drèze, J. and Somanchi, A. (2021). “Not having the have nots”. *Economic Times*, 21 June.

Fang, T., Peters, M., and Zilibotti, F. (2023). “Growing Like India—the Unequal Effects of Service-Led Growth”. *Econometrica*, 91(4):1457–1494.

Garbinti, B., Goupille-Lebert, J., and Piketty, T. (2018). “Income inequality in France, 1900-2014: Evidence from Distributional National Accounts (DINA)”. *Journal of Public Economics*, 162:63–77.

Ghatak, M. and Mukherjee, U. (2019). "The Mirage of Modinomics". *The India Forum*, 22 February.

Ghatak, M., Raghavan, R., and Xu, L. (2022). "Trends in Economic Inequality in India". *The India Forum*, 19 September.

Gupta, A., Malani, A., and Woda, B. (2021). "Inequality in India declined during Covid". NBER Working Paper No. 29597.

Jaffrelot, C. (2021). "Modi's India: Hindu Nationalism and the Rise of Ethnic Democracy". Princeton University Press, Princeton NJ.

Jayadev, A., Motiram, S., and Vakulabharanam, V. (2007). "Patterns of Wealth Disparities in India during the Liberalization Era". *Economic and Political Weekly*, 42(38):3852–2863.

Jayadev, A. and Narayan, A. (2020). "The Evolution of India's Industrial Labour Share and its Correlates". *Development and Change*, 51(4):998–1017.

Jayaraj, D. and Subramanian, S. (2018). "The Distribution of Household Assets in India: 1991-1992 to 2012-2013". *Indian Journal of Human Development*, 12(2):181–203.

Kanbur, R., Wang, Y., and Zhang, X. (2021). "The great Chinese inequality turnaround". *Journal of Comparative Economics*, 49(2):467–482.

Kapoor, A. and Duggal, J. (2022). "The State of Inequality in India Report". Institute for Competitiveness, New Delhi.

Kazmin, A. (2019). "Economists condemn politicization of Modi government data". *Financial Times*, 15 March.

Khanna, G. and Mukherjee, P. (2023). "Political accountability for populist policies: Lessons from World's largest democracy". *Journal of Public Economics*, 219.

Khera, R. and Yadav, M. (2020). "What pay ratios in NIFTY50 companies tell us about income inequality in India". *Ideas for India*, 20 November.

Korinek, A., Mistiaen, J. A., and Ravallion, M. (2006). "Survey nonresponse and the distribution of income". *Journal of Economic Inequality*, 4:33–55.

Kumar, R. (2019). "The evolution of wealth-income ratios in India 1860-2012". WID.world Working Paper No. 2019/07.

Kumar, R. (2020). "Top Indian wealth shares and inheritances, 1966-1985". *Cliometrica*, 14:551–580.

Kundu, T. (2017). "Is Thomas Piketty right about inequality in India?". *Mint*, 15 September.

Kuznets, S. (1953). "Shares of Upper Income Groups in Income and Savings". National Bureau of Economic Research, Cambridge MA.

Morris, S. and Kumari, T. (2019). "Overestimation in the Growth Rates of National Income in Recent Years? An Analyses Based on Extending GDP04-05 Through Other Indicators of Output". IIM-A Working Paper No. 2019-01-01.

Nagaraj, R. (2020). "Understanding India's Economic Slowdown". *The India Forum*, 20 February.

Narayanan, M. P. (1988). "Debt Versus Equity Under Asymmetric Information". *The Journal of Financial and Quantitative Analysis*, 23(1):39–51.

National Sample Survey Office (2024). "Household Consumption Expenditure Survey: 2022-23 Fact Sheet". Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

Özler, B., Datt, G., and Ravallion, M. (1996). "A Database on Poverty and Growth in India". mimeo, Poverty and Human Resources Division, The World Bank.

Pai, S. and Vats, A. (2023). "Indus Valley Annual Report". Blume Ventures.

Pareto, V. (1896). Cours d'économie politique.

Piketty, T. and Saez, E. (2003). "Income Inequality in the United States 1913-1998". *The Quarterly Journal of Economics*, 118(1):1–39.

Piketty, T., Saez, E., and Zucman, G. (2018). "Distributional National Accounts: Methods and Estimates for the United States". *The Quarterly Journal of Economics*, 133(2):553–609.

Piketty, T., Yang, L., and Zucman, G. (2019). "Capital Accumulation, Private Property and Rising Inequality in China, 1978-2015". *American Economic Review*, 109(7):2469–2496.

Piketty, T. and Zucman, G. (2014). "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010". *The Quarterly Journal of Economics*, 129(3):1155–1210.

Press Information Bureau (2023). "Agriculture has highest estimated percentage distribution of female workers followed by manufacturing as per the Annual Periodic Labour Force Survey (PLFS) Report 2021-22". Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

Press Information Bureau (2024). "Percapita Monthly Household Consumption Expenditure more than doubled during 2011-12 to 2022-23". Ministry of Statistics and Programme Implementation, Government of India, New Delhi.

Sapre, A. and Bhardwaj, V. (2023). "Status and Compilation Issues in National Accounts Statistics: A Short Summary". NIPFP Working Paper No. 297.

Singh, R. (2023). "Do the Wealthy Underreport their Income?" Using General Election Filings to Study the Income-Wealth Relationship in India". WIL Working Paper No. 2023/01.

Sinha, S. (2006). "Evidence for power-law tail of the wealth distribution in India". *Physica A: Statistical Mechanics and its Applications*, 359:555–562.

Somanchi, A. (2021). "Missing the Poor, Big Time: A Critical Assessment of the Consumer Pyramids Household Survey". *SocArxiv*.

Somanchi, A. (2023). "Income Inequality in India, 2000-2020". MSc thesis, Paris School of Economics.

Subramanian, A. (2019a). "India's GDP Mis-estimation: Likelihood, Magnitudes, Mechanisms and Implications". CID Faculty Working Paper No. 354.

Subramanian, A. and Felman, J. (2023). "Is the economy surging or decelerating: Understanding India's growth rate". *Business Standard*, 15 September.

Subramanian, S. (2019b). "What is Happening to Rural Welfare, Poverty and Inequality in India?". *The India Forum*, 27 November.

Subramanian, S. and Jayaraj, D. (2006). "The Distribution of Household Wealth in India". WIDER Working Paper Series RP2006-116.

Thapar, K. (2023). "Full Text: Pronab Sen Explains Why Data on Which GDP is Calculated Is A Major Concern". *The Wire*, 29 November.

Vermeulen, P. (2016). "Estimating the Top Tail of the Wealth Distribution". *American Economic Review: Papers & Proceedings*, 106(5):646–650.

Table 2: Income inequality in India, 2022-23

Income Group	Adults	Income share (%)	Threshold (INR)	Average income (INR)	Ratio to average
Average	92,23,44,832	100.0	0	2,34,551	1.0
Bottom 50%	46,11,72,416	15.0	0	71,163	0.3
Middle 40%	36,89,37,933	27.3	1,05,413	165,273	0.7
Top 10%	9,22,34,483	57.7	2,90,848	13,52,985	5.8
Top 1%	92,23,448	22.6	20,73,846	53,00,549	22.6
incl. Top 0.1%	9,22,345	9.6	82,20,379	2,24,58,442	95.8
incl. Top 0.01%	92,234	4.3	3,46,06,044	10,18,14,669	434.1
incl. Top 0.001%	9,223	2.1	20,01,98,548	48,51,96,875	2,068.6

Notes: The table presents a summary of income inequality in India in 2022-23. All INR values in current 2022 prices. Adult population estimates for 2022 from UN World Population Prospects. Average income scaled to match national income accounts totals as per WID data (differs marginally from official sources). See section 2.2 for details.

Sources: Authors' estimates combining national income accounts aggregates, tax tabulations and surveys on income and consumption.

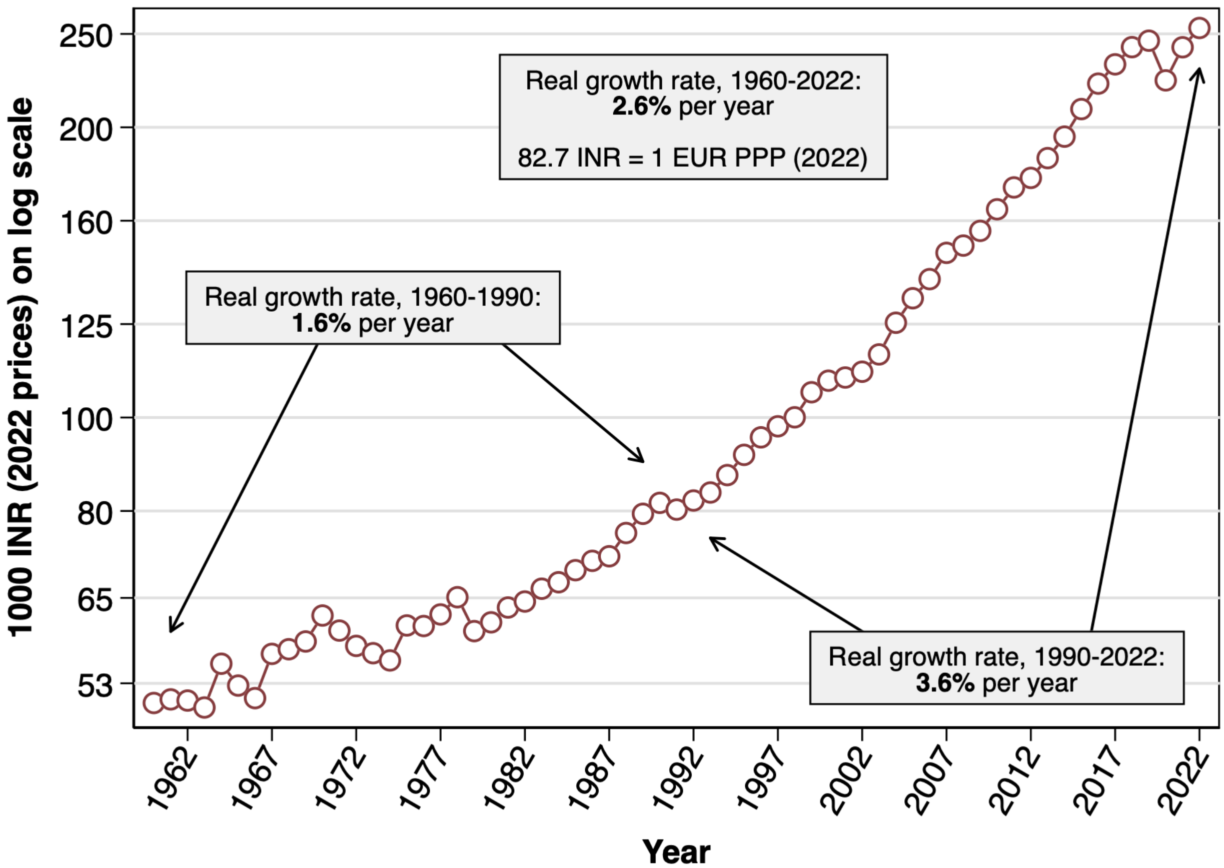
Table 3: Wealth inequality in India, 2022-23

Wealth Group	Adults	Wealth share (%)	Threshold (INR)	Average wealth (INR)	Ratio to average
Average	92,23,44,832	100.0	-4,10,00,000	13,49,029	1.0
Bottom 50%	46,11,72,416	6.4	-4,10,00,000	1,73,184	0.1
Middle 40%	36,89,37,933	28.6	4,31,138	9,63,560	0.7
Top 10%	9,22,34,483	65.0	21,98,344	87,70,132	6.5
Top 1%	92,23,448	40.1	81,60,022	5,41,41,525	40.1
incl. Top 0.1%	9,22,345	29.7	5,26,17,860	40,04,54,807	296.8
incl. Top 0.01%	92,234	22.2	36,86,80,160	2,99,67,73,491	2,221.4
incl. Top 0.001%	9,223	16.8	2,75,66,99,904	22,61,33,54,928	16,762.7

Notes: The table presents a summary of wealth inequality in India in 2022-23. All INR values in current (2022) prices. Adult population estimates for 2022 from UN World Population Prospects. Average wealth scaled to match aggregate national wealth as per WID data. See section 2.3 for details. The threshold of INR -4.1 crore for the bottom 50% is driven by one observation in AIDIS with enormous debt. The $p2$ in the distribution is INR -4,000 and the average wealth for the bottom 50% after dropping the extreme negative outlier is 1,93,031.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys (AIDIS) and Forbes billionaire rankings.

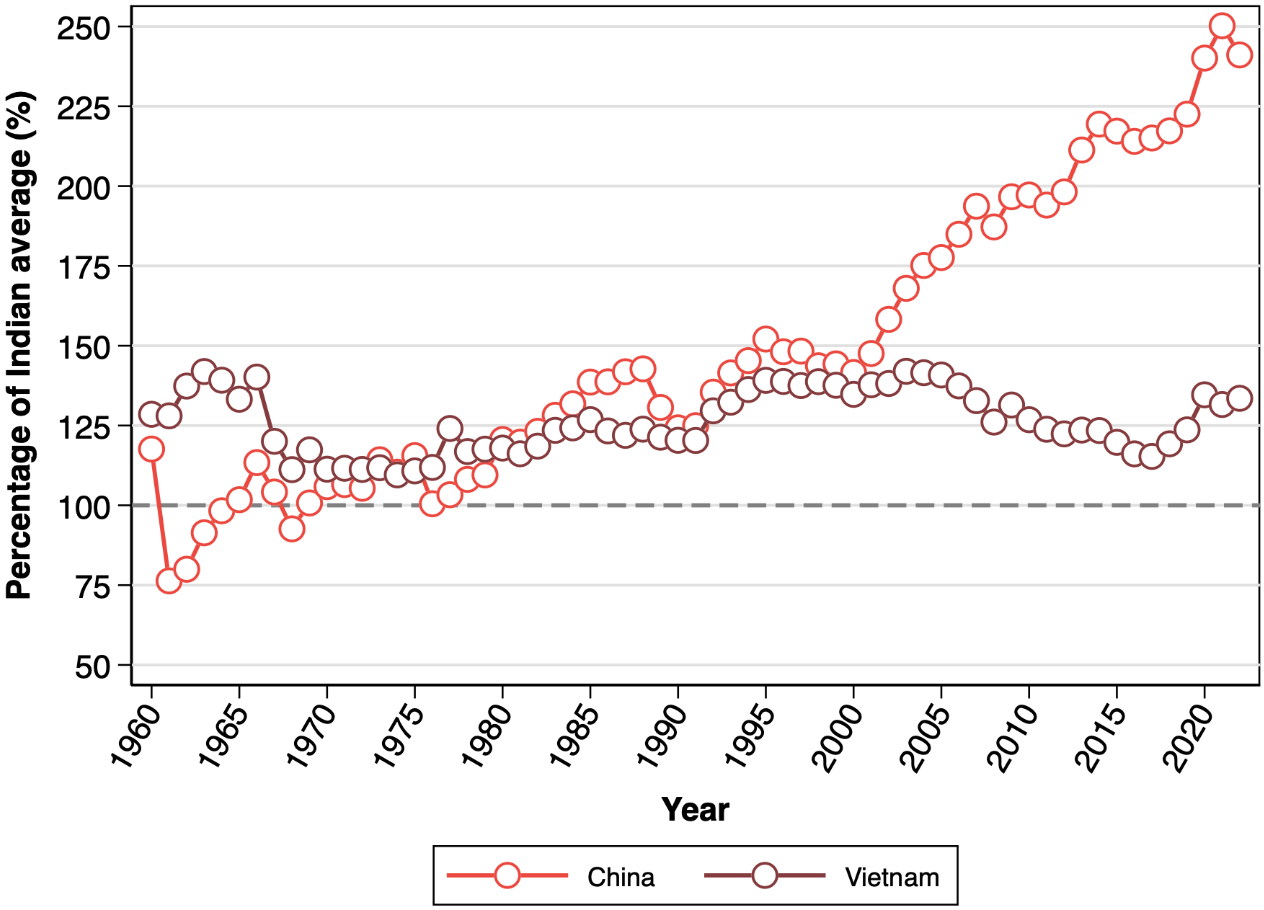
Figure 1a: Evolution of average income in India, 1960-2022



Note: Average income = net national income / adult population. Incomes are plotted on log-scale.

Sources: Authors' estimates combining national income accounts data from WID for pre-2014 and Table 1.1, Statistical Appendix, Economic Survey 2022-23 for post-2014, adult population from United Nations World Population Prospects, and price index (GDP deflator) from World Bank's World Development Indicators database.

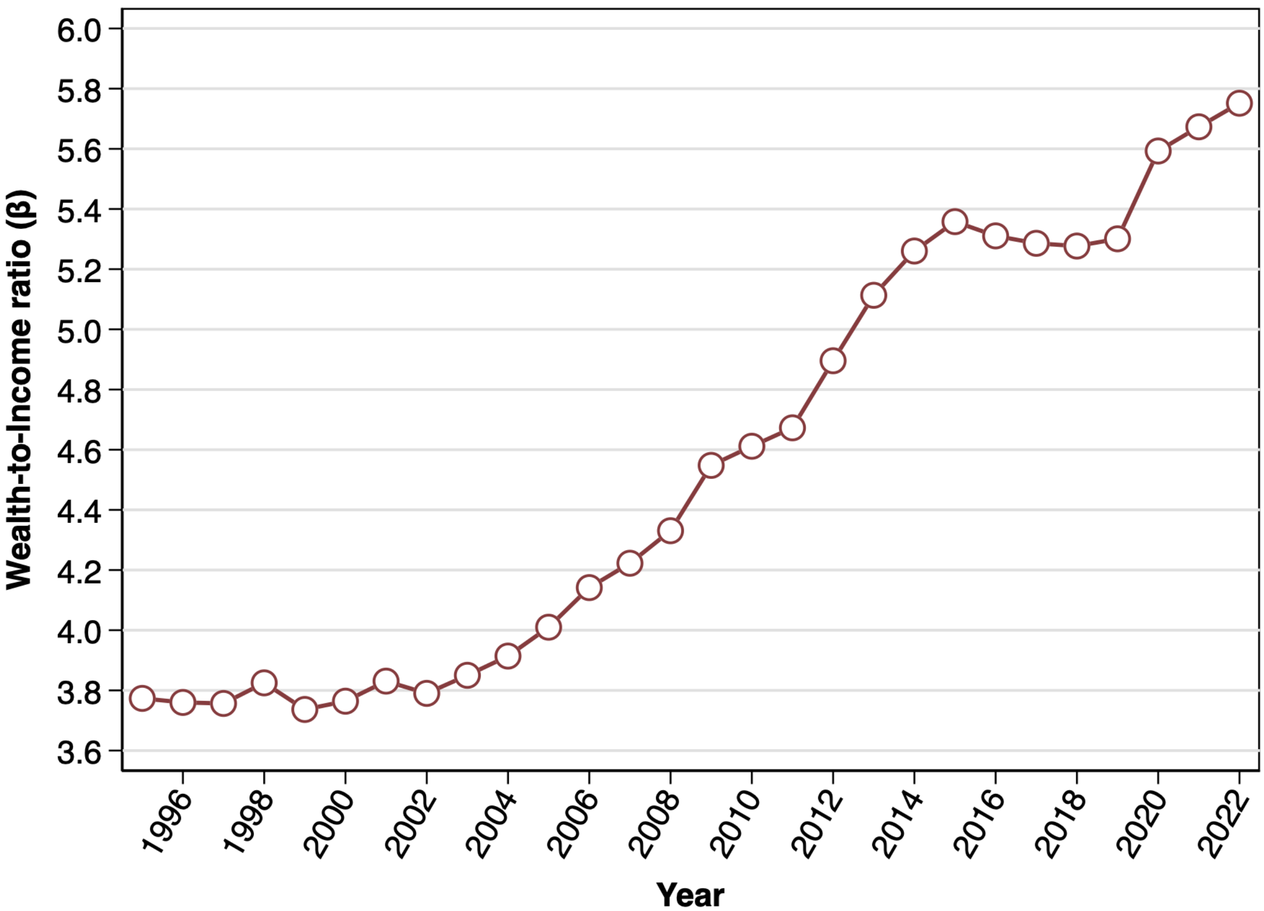
Figure 1b: Indian incomes in comparative perspective, 1960-2022



Note: The figure presents average incomes in China and Vietnam as a percentage of India's average. Incomes in nominal local currency converted to 2022 Euros PPP to account for local inflation and purchasing power differentials.

Sources: Authors' estimates combining national income accounts data, price index (GDP deflator) and adult population from WID.

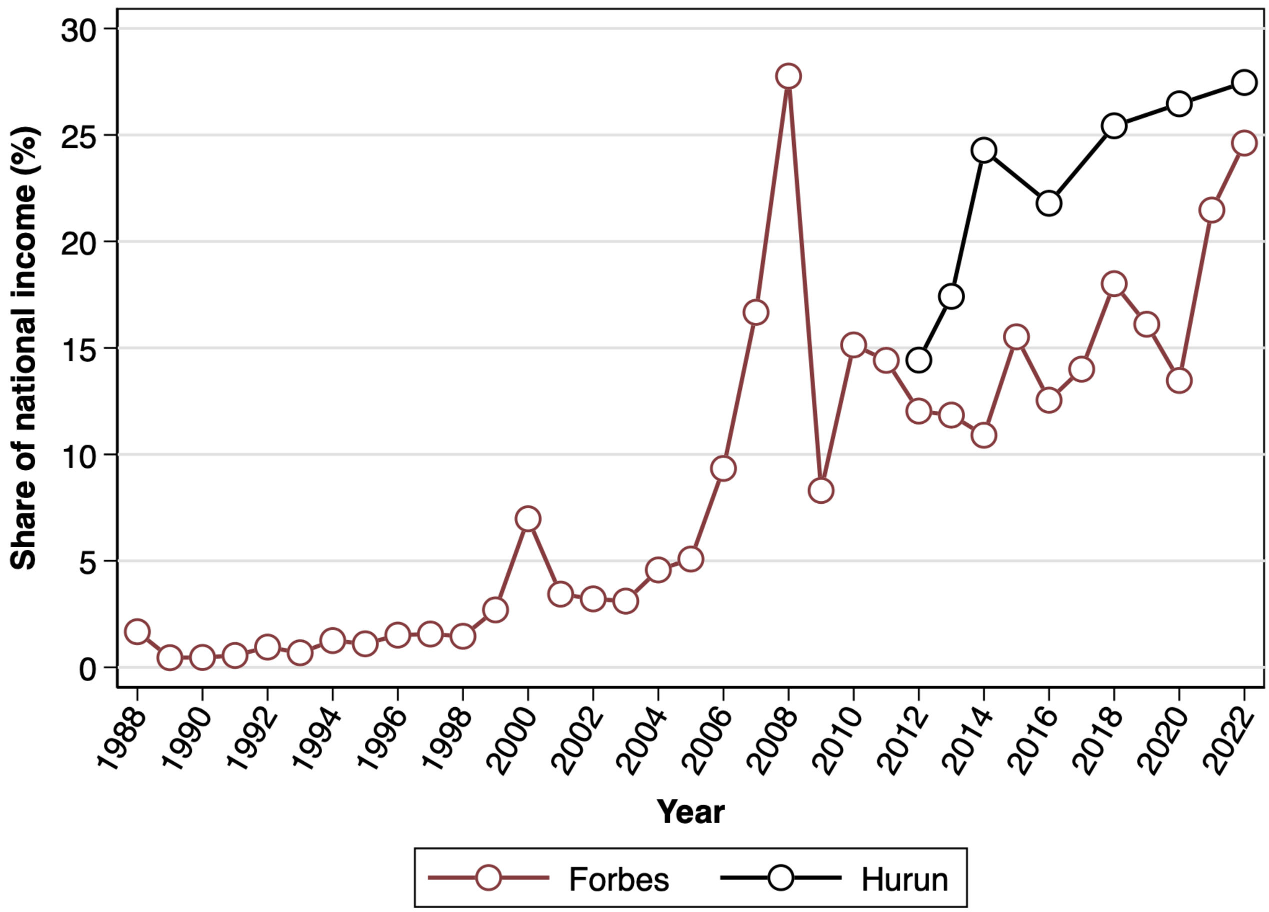
Figure 2a: National wealth-to-income ratio, 1995-2022



Note: The figure presents India's aggregate wealth-to-income ratio (β) for the period 1995-2022.

Sources: WID data based on Kumar (2019).

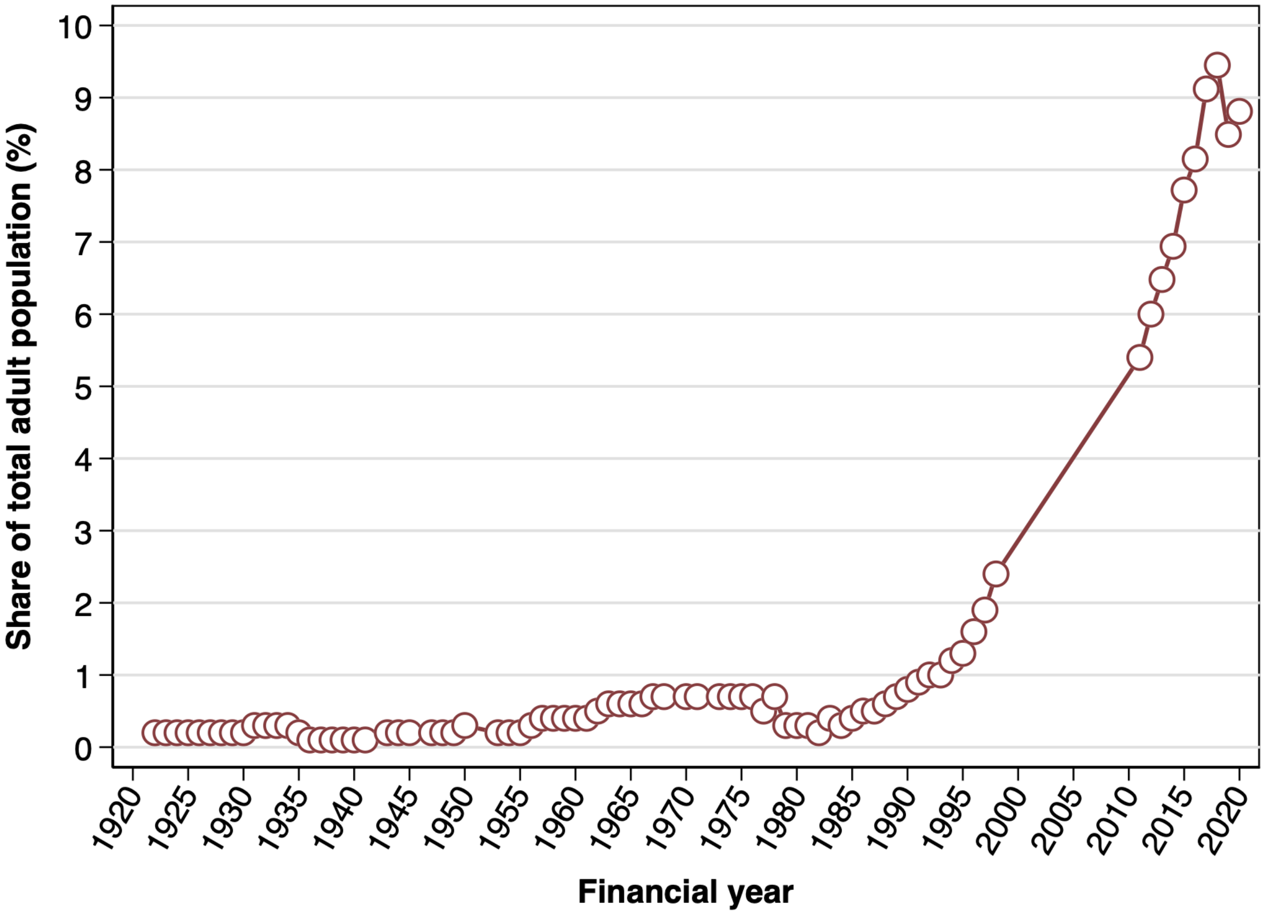
Figure 2b: Wealth of richest Indians, 1988-2022



Note: The Forbes billionaire rankings track all individuals with net wealth exceeding 1 billion USD MER. The number of such individuals increased from 1 in 1988 to 7 in 1999 to 46 in 2012 to 162 in 2022. The Hurun rich list tracks all individuals with net wealth exceeding 1000 crore INR (roughly 120 million USD MER as of March 2024). The number of such individuals increased from 100 in 2012 to 831 in 2018 to 1103 in 2022. Appendix Table C.2 presents annual figures for both sources.

Sources: Authors' estimates combining national income accounts data from WID data for pre-2014 and Table 1.1, Statistical Appendix, Economic Survey 2022-23 for post-2014, with data from Forbes billionaire rankings and Hurun rich lists.

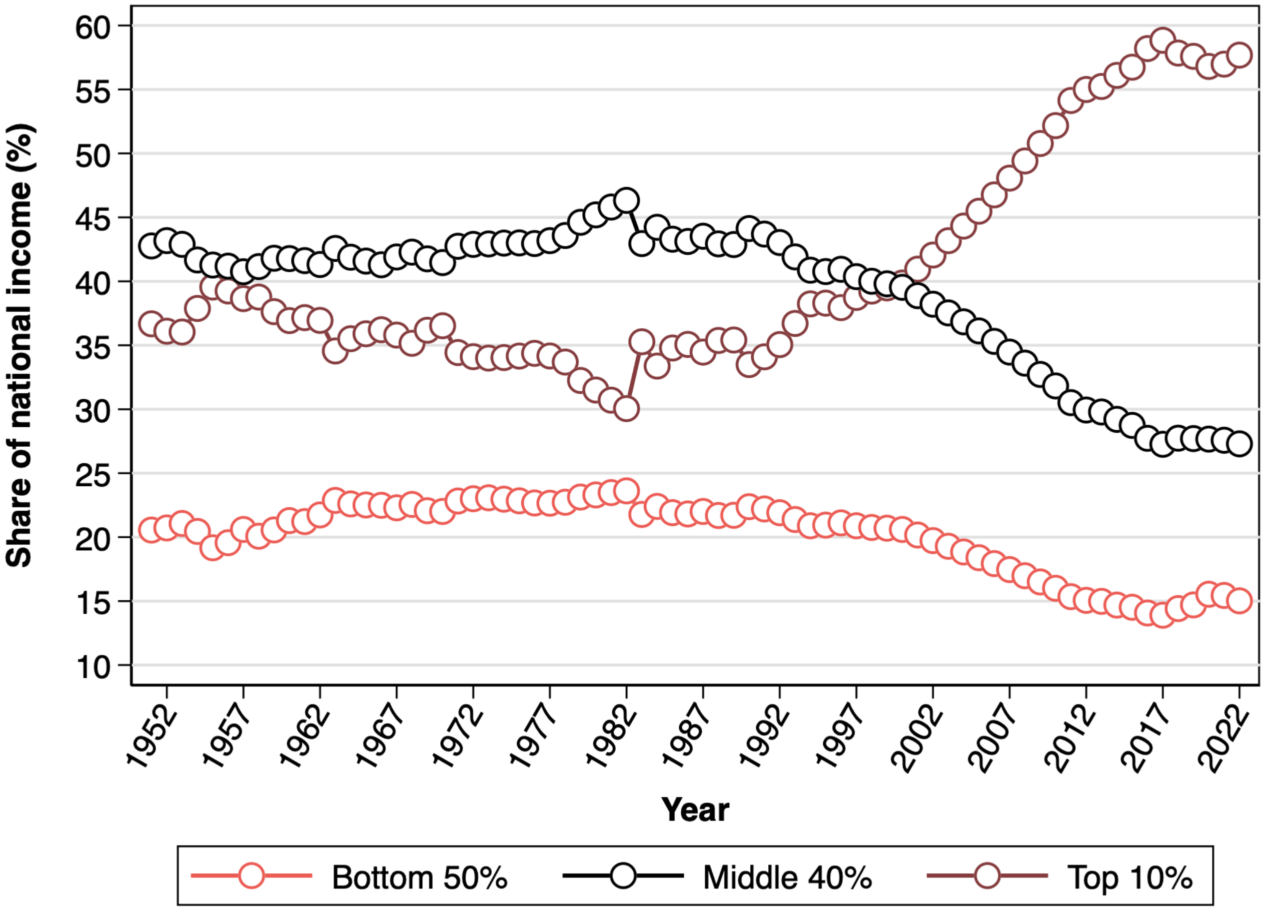
Figure 3: Proportion of income-tax return filers in India, 1922-2020



Note: Figures relate to individuals and Hindu Undivided Families (HUF) and include those that paid income tax at source but did not file a return and those that filed a return but did not pay income tax because they fell below the taxable threshold and/or on account of deductions. Figures excludes returns with zero or negative gross incomes.

Sources: Authors' computations based on Indian income tax administration statistics, adult population figures from UN World Population Prospects, and Banerjee and Piketty (2005).

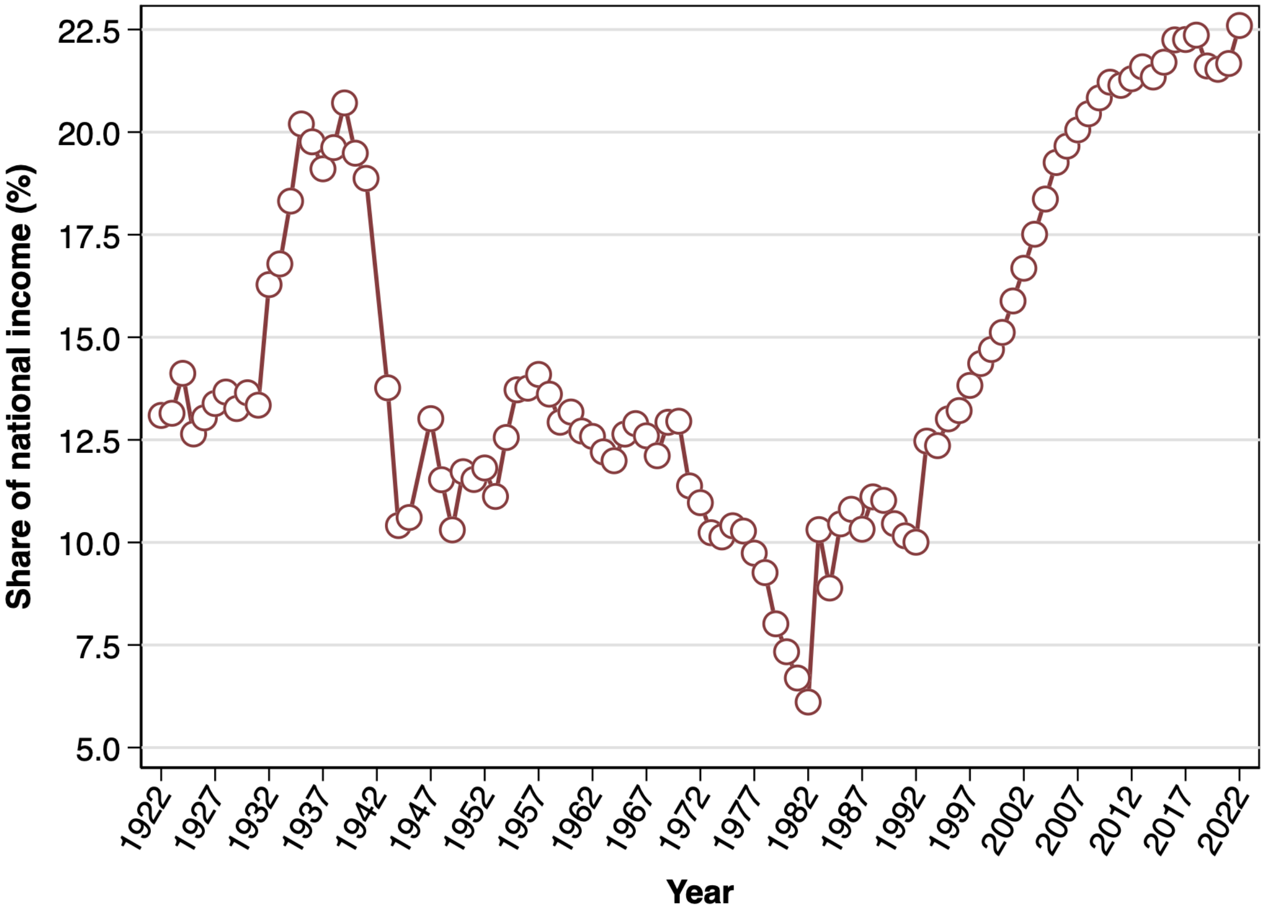
Figure 4: Long-run income inequality in India, 1951-2022



Note: The figures presents the distribution of pre-tax per-adult net national income.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

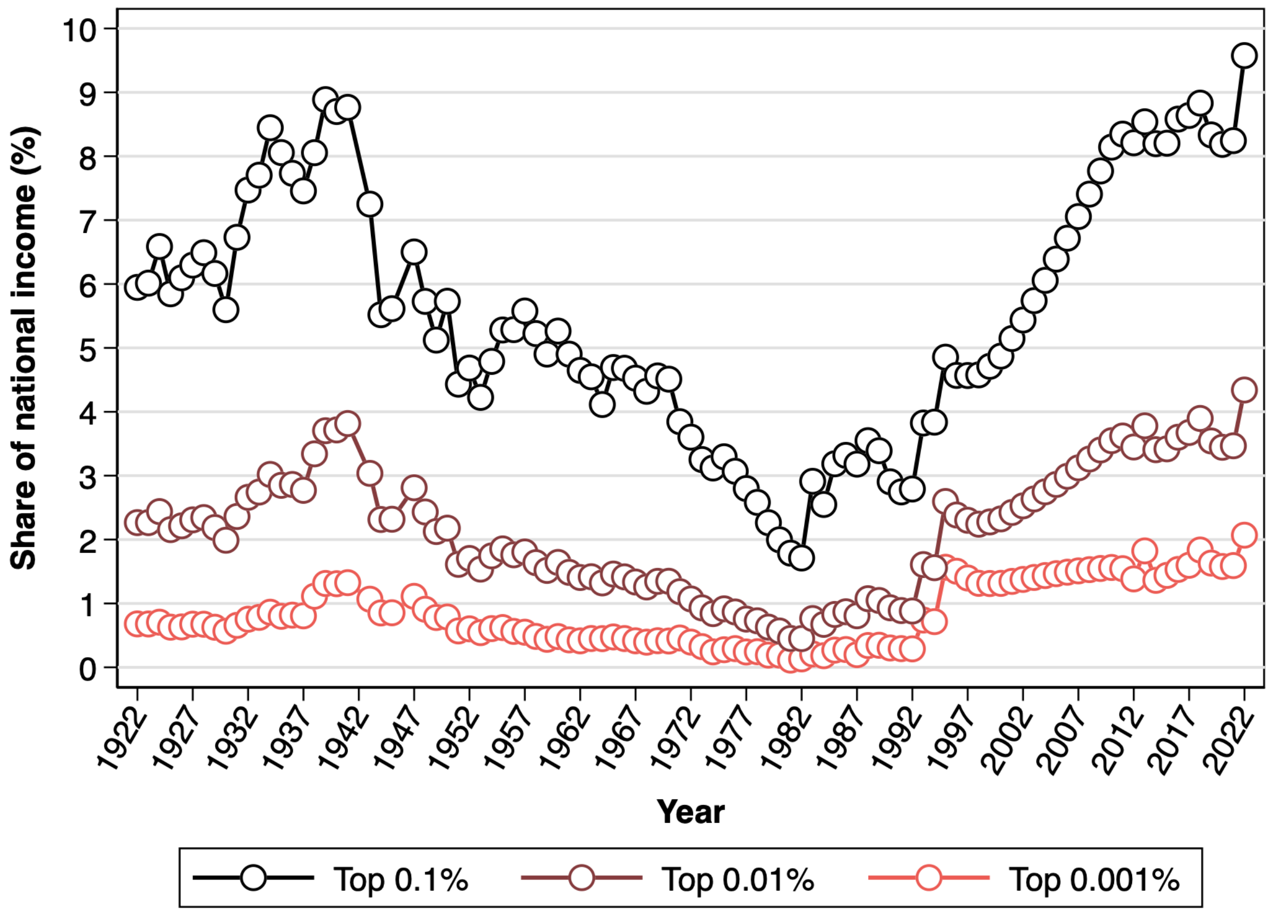
Figure 5a: Top 1% national income share, 1922-2022



Note: The figure presents the distribution of pre-tax per-adult net national income.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

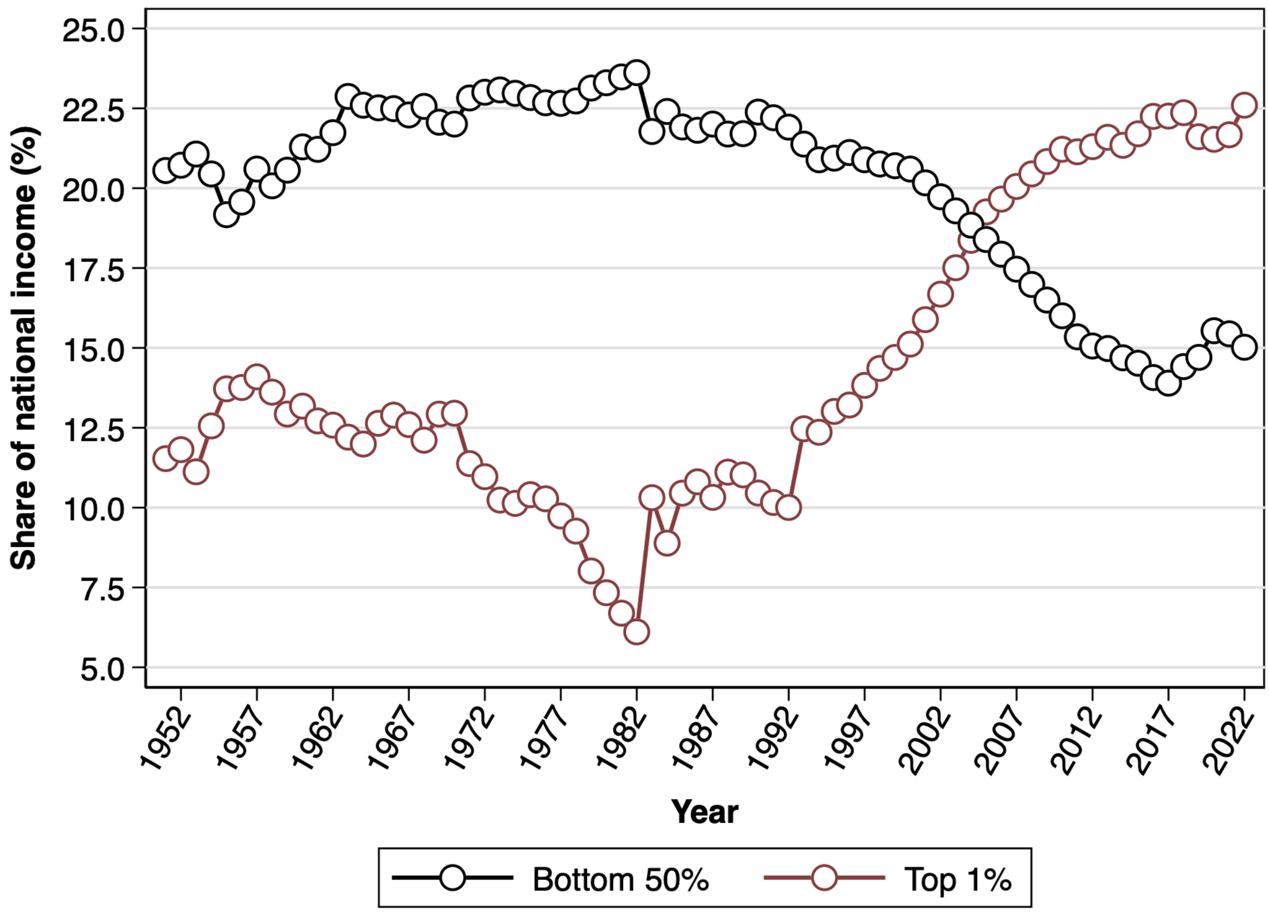
Figure 5b: Inequality within the richest 1%, 1922-2022



Note: The figure presents the distribution of pre-tax per-adult net national income.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

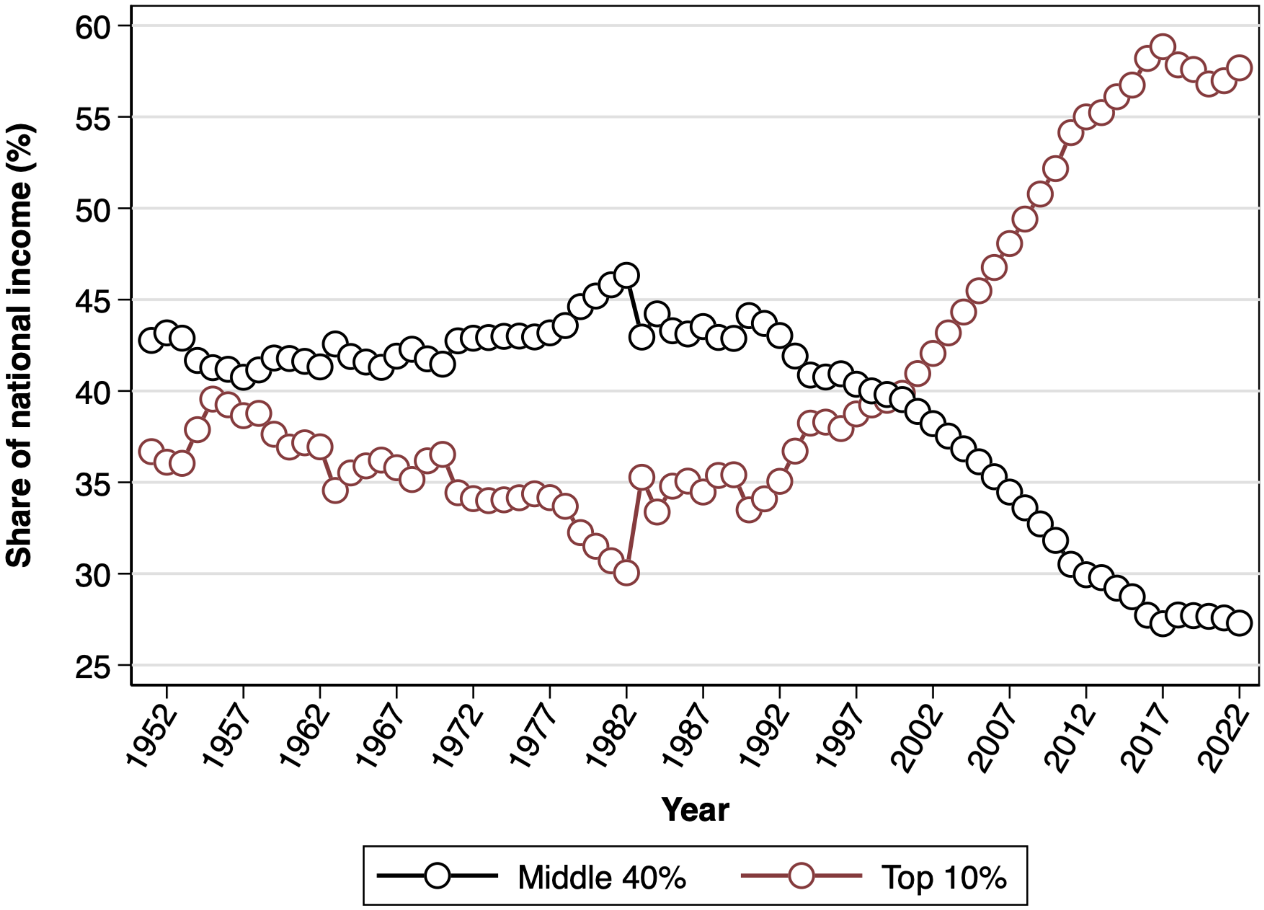
Figure 6a: Bottom 50% vs Top 1% national income shares, 1951-2022



Note: The figure presents the distribution of pre-tax per-adult net national income.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

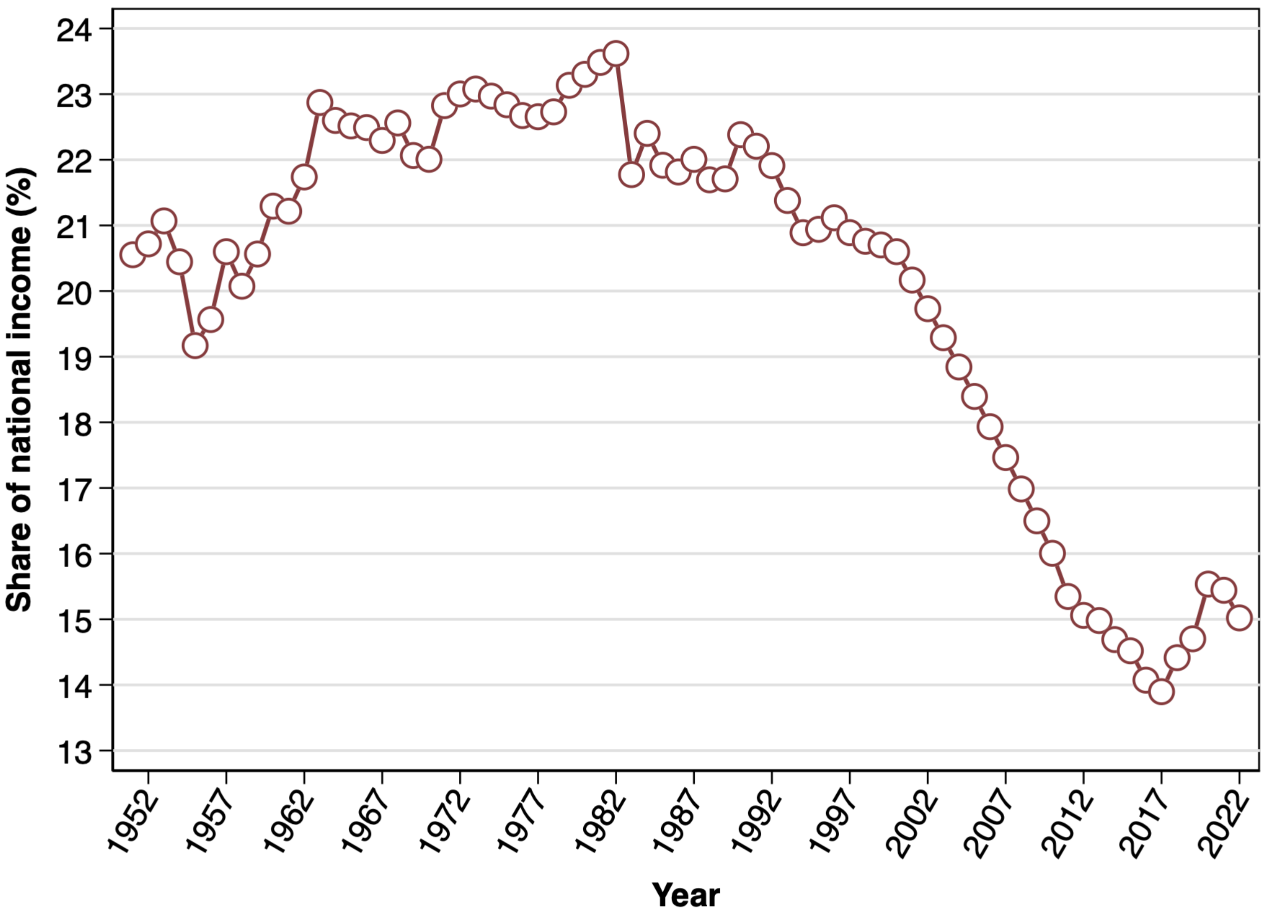
Figure 6b: Middle 40% vs Top 10% national income shares, 1951-2022



Note: The figure presents the distribution of pre-tax per-adult net national income.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

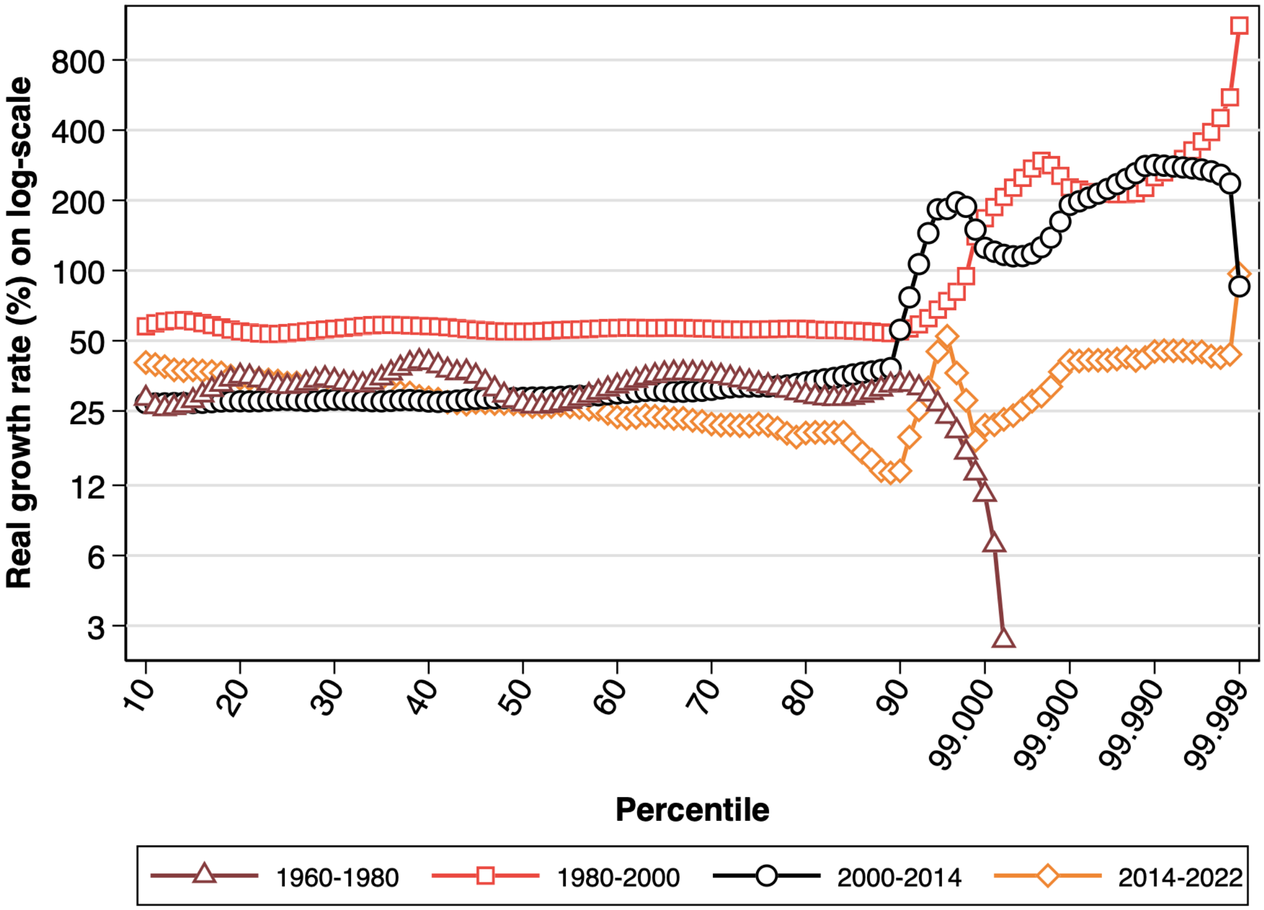
Figure 7: Bottom 50% national income share, 1951-2022



Note: The figures presents the distribution of pre-tax per-adult net national income.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

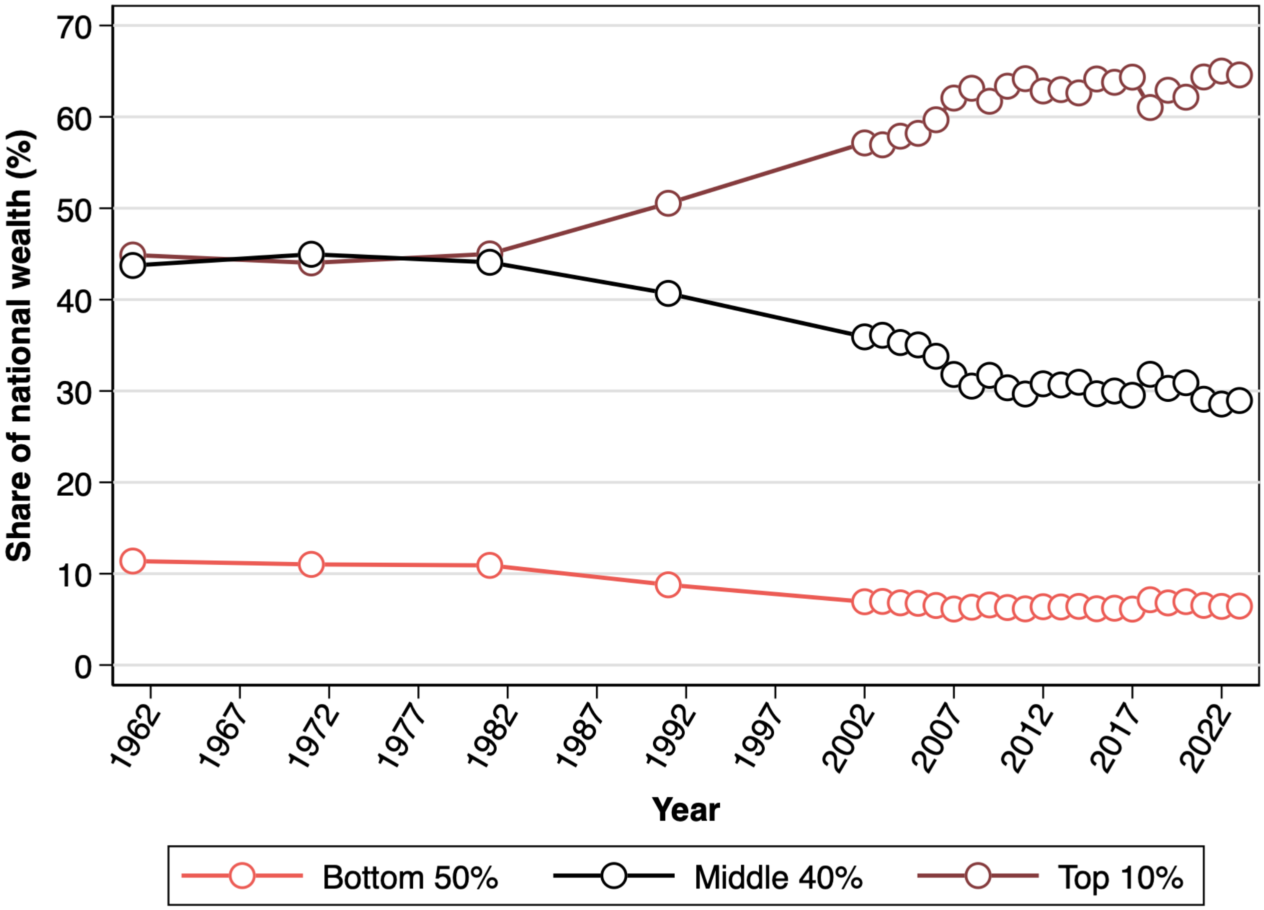
Figure 8: Income growth incidence curves, 1960-2022



Note: The figure presents the real cumulative growth rate of incomes at each percentile of the distribution starting $p10$ for 4 different periods. Growth rates are plotted on log-scale. The estimates for 1960-1980 do not go much beyond $p99$ as the top 1% experienced severely negative real growth rates (-1.6% at $p99.3$, -46.5% at $p99.9$, -63.0% at $p99.998$) which cannot be plotted on log-scale.

Sources: Authors' estimates combining income and consumption surveys, tax tabulations and national income accounts aggregates.

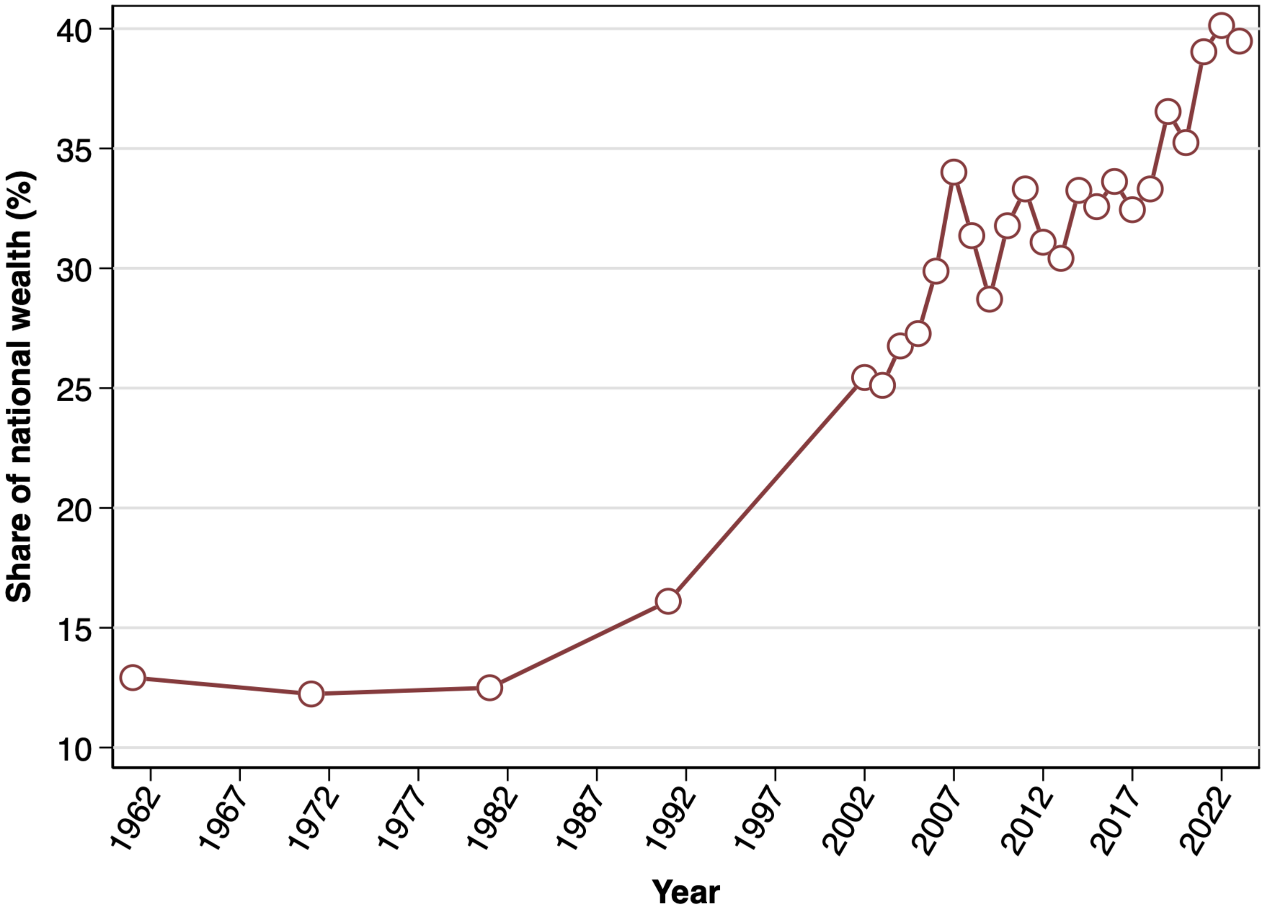
Figure 9: Long-run wealth inequality in India, 1961-2023



Note: The figure presents the distribution of per-adult national wealth.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys and Forbes billionaire data.

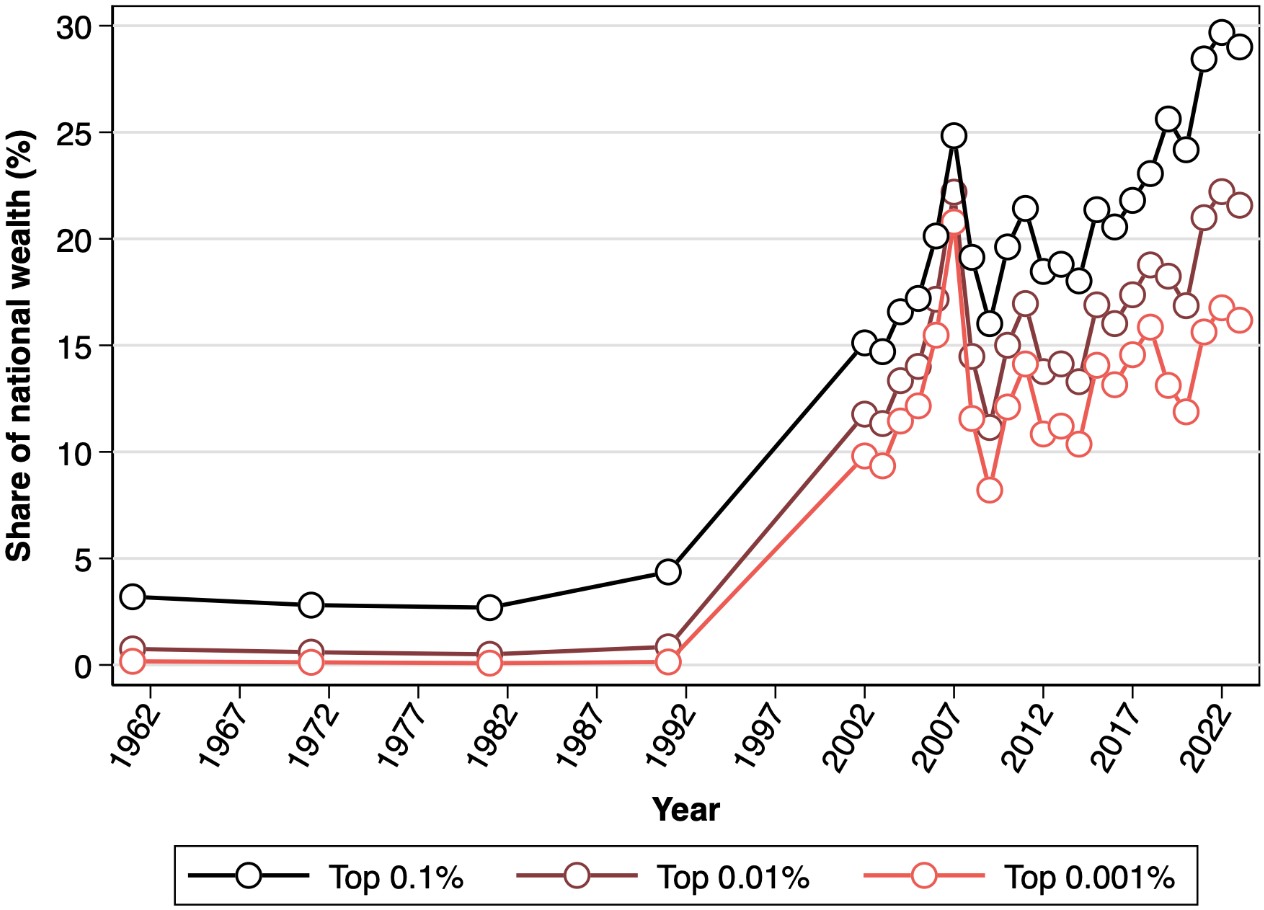
Figure 10a: Top 1% National wealth share, 1961-2023



Note: The figure presents the distribution of per-adult national wealth.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys and Forbes billionaire data.

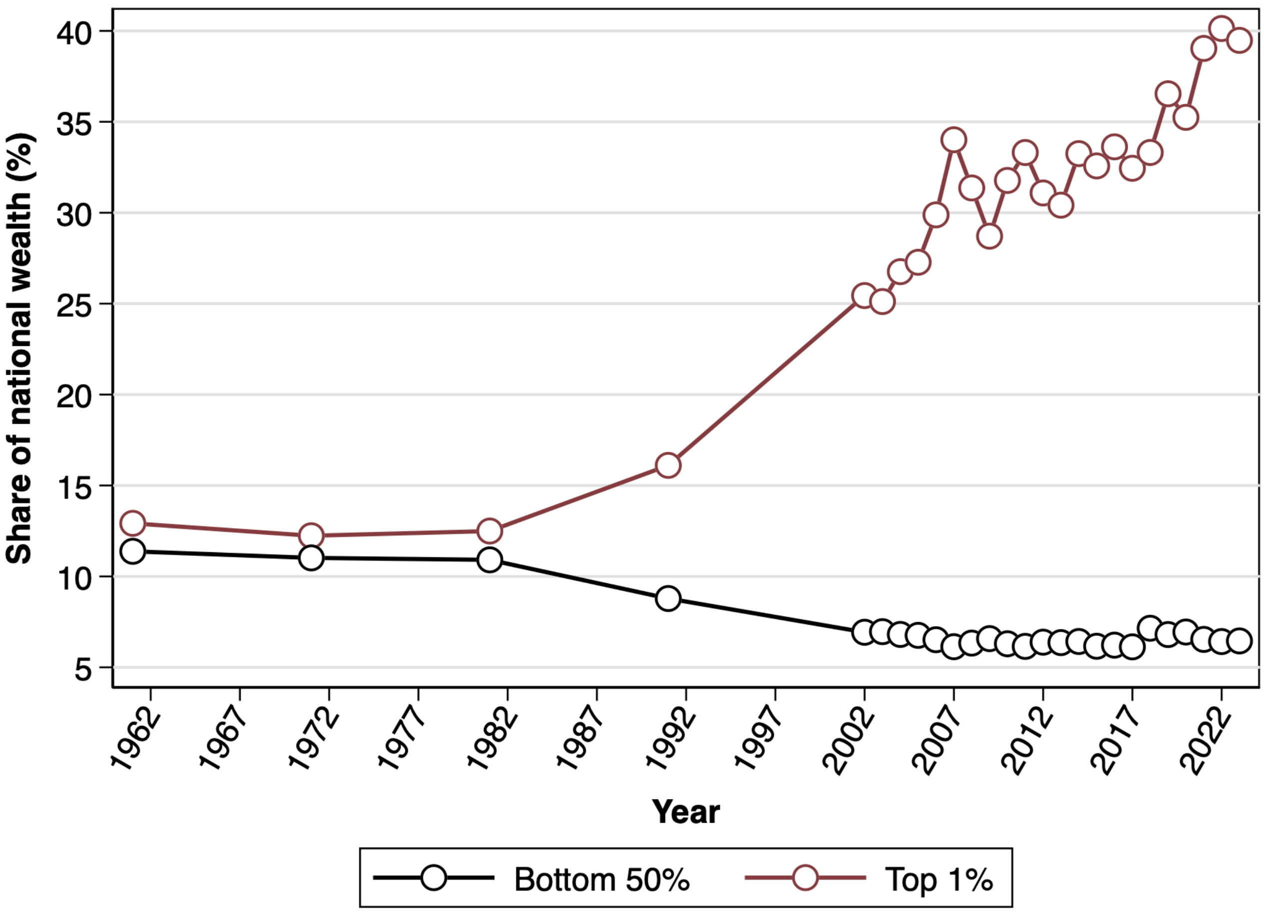
Figure 10b: Inequality within the wealthiest 1%, 1951-2023



Note: The figure presents the distribution of per-adult national wealth.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys and Forbes billionaire data.

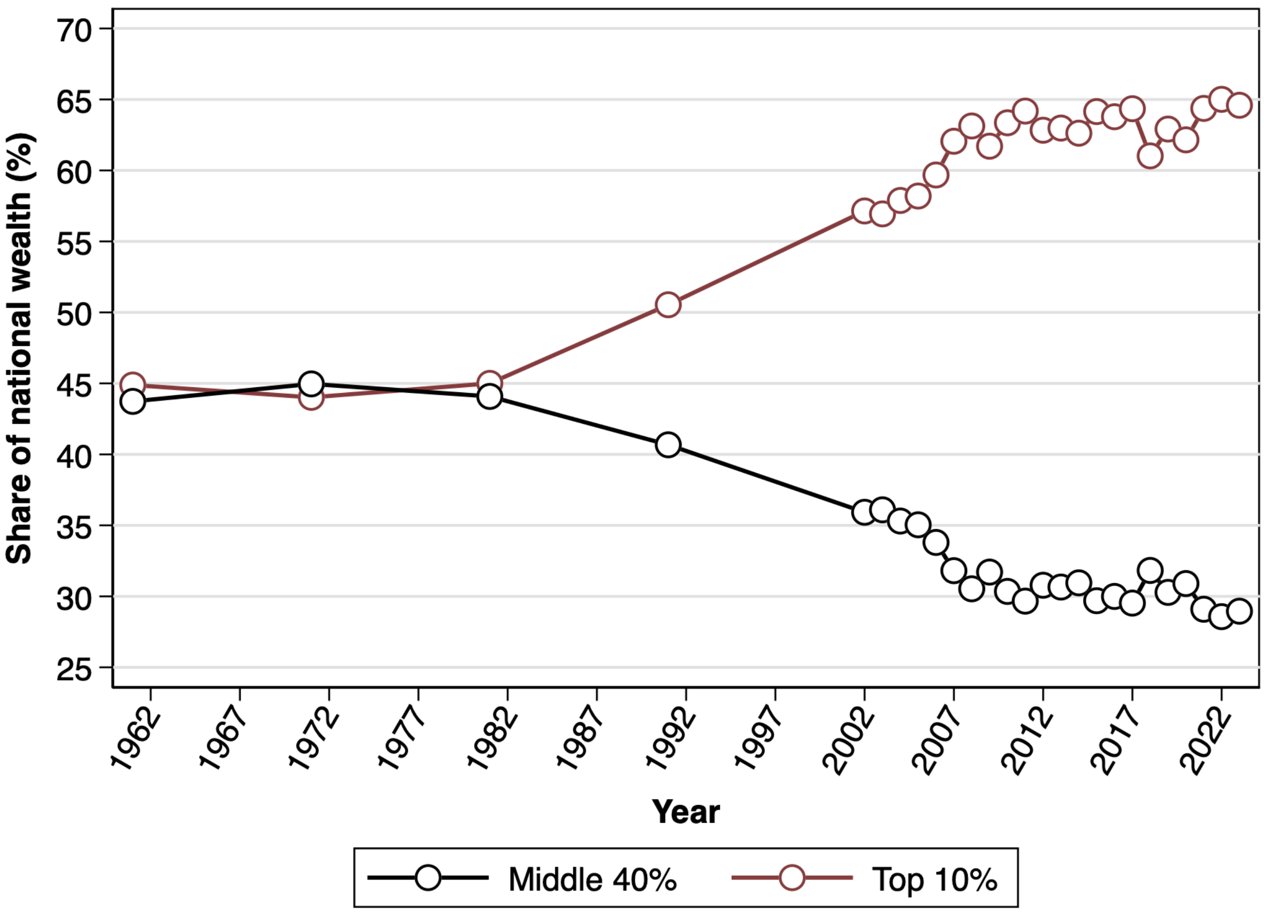
Figure 11a: Bottom 50% vs Top 1% wealth shares, 1961-2023



Note: The figure presents the distribution of per-adult national wealth.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys and Forbes billionaire data.

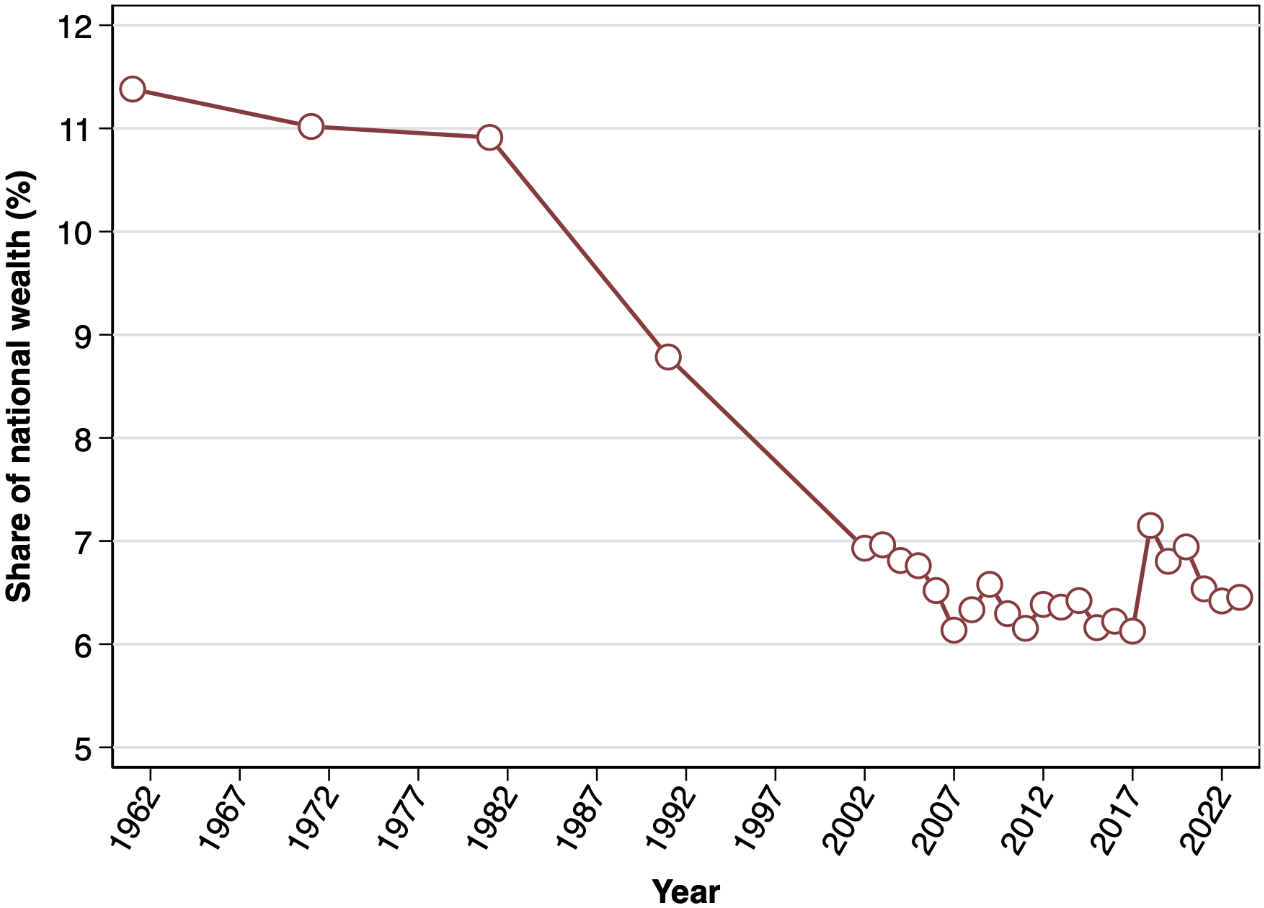
Figure 11b: Middle 40% vs Top 10% wealth shares, 1961-2023



Note: The figure presents the distribution of per-adult national wealth.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys and Forbes billionaire data.

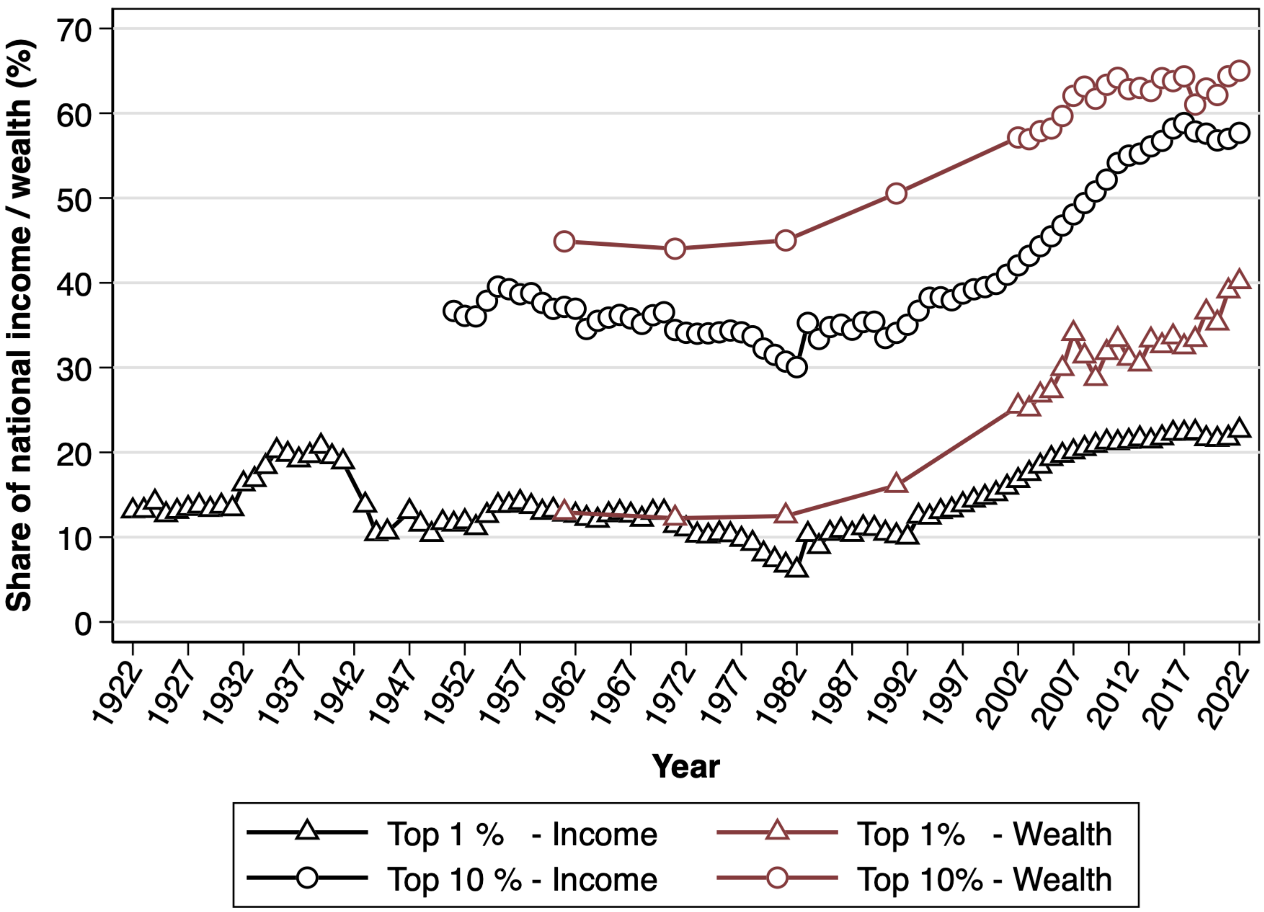
Figure 12: Bottom 50% national wealth share, 1961-2023



Note: The figure presents the distribution of per-adult national wealth.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys and Forbes billionaire data.

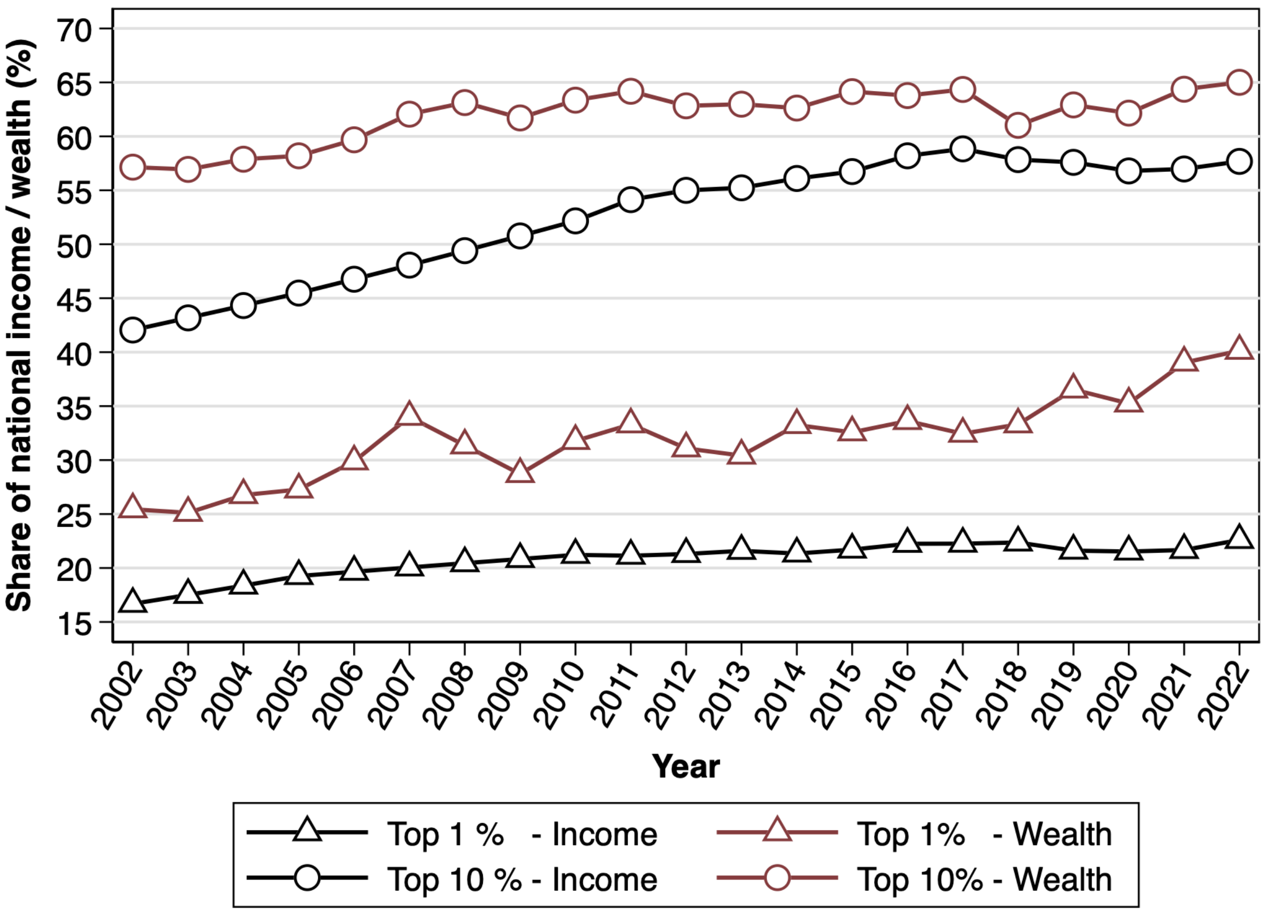
Figure 13a: Income and wealth inequality in the long run, 1922-2022



Note: The figure presents the distribution of per-adult net national income and national wealth.

Sources: Authors' estimates combining national income and wealth aggregates, surveys (consumption, income, wealth), tax tabulations, and Forbes data.

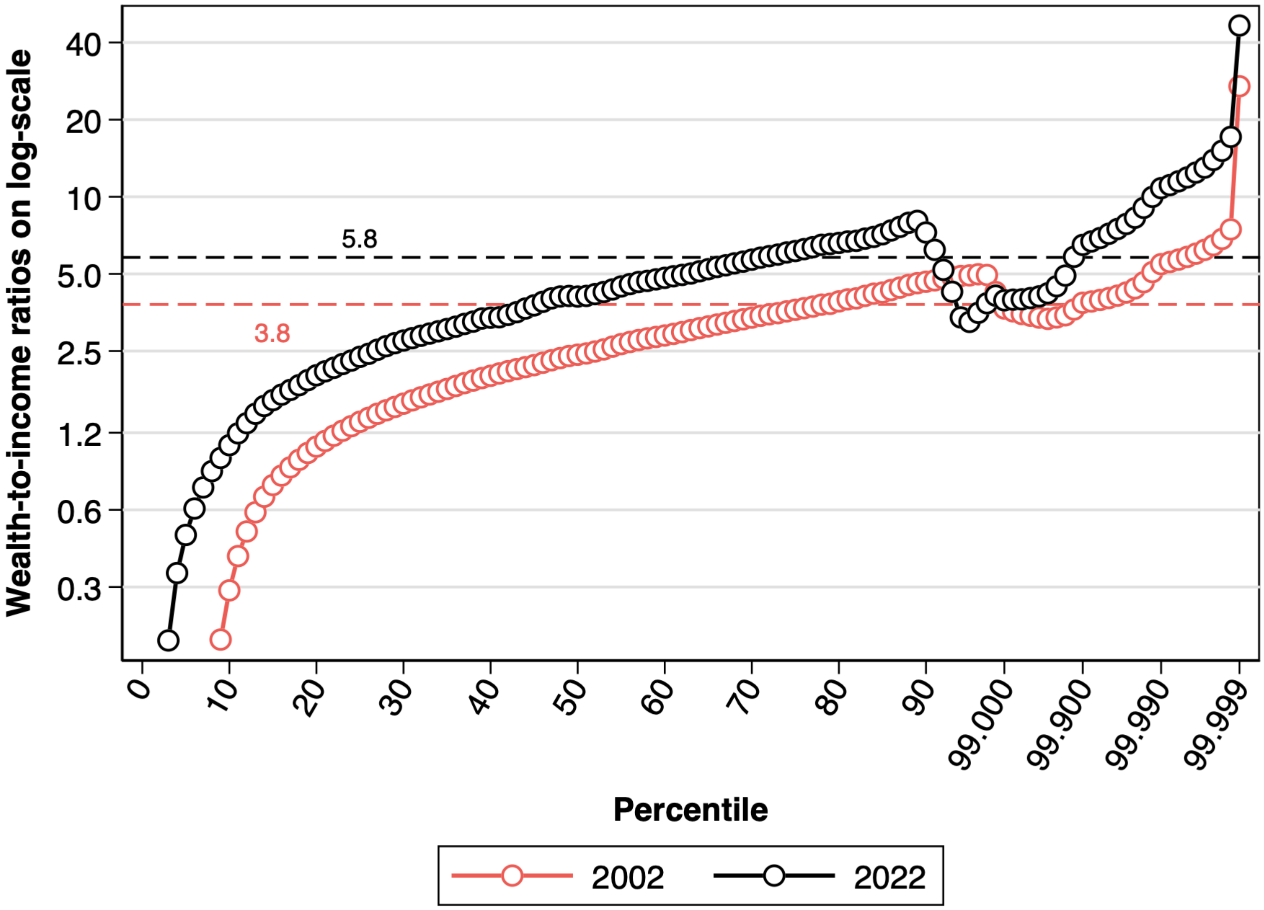
Figure 13b: Income and wealth inequality in the short run, 2002-2022



Note: The figure presents the distribution of per-adult net national income and national wealth.

Sources: Authors' estimates combining national income and wealth aggregates, surveys (consumption, income, wealth), tax tabulations, and Forbes data.

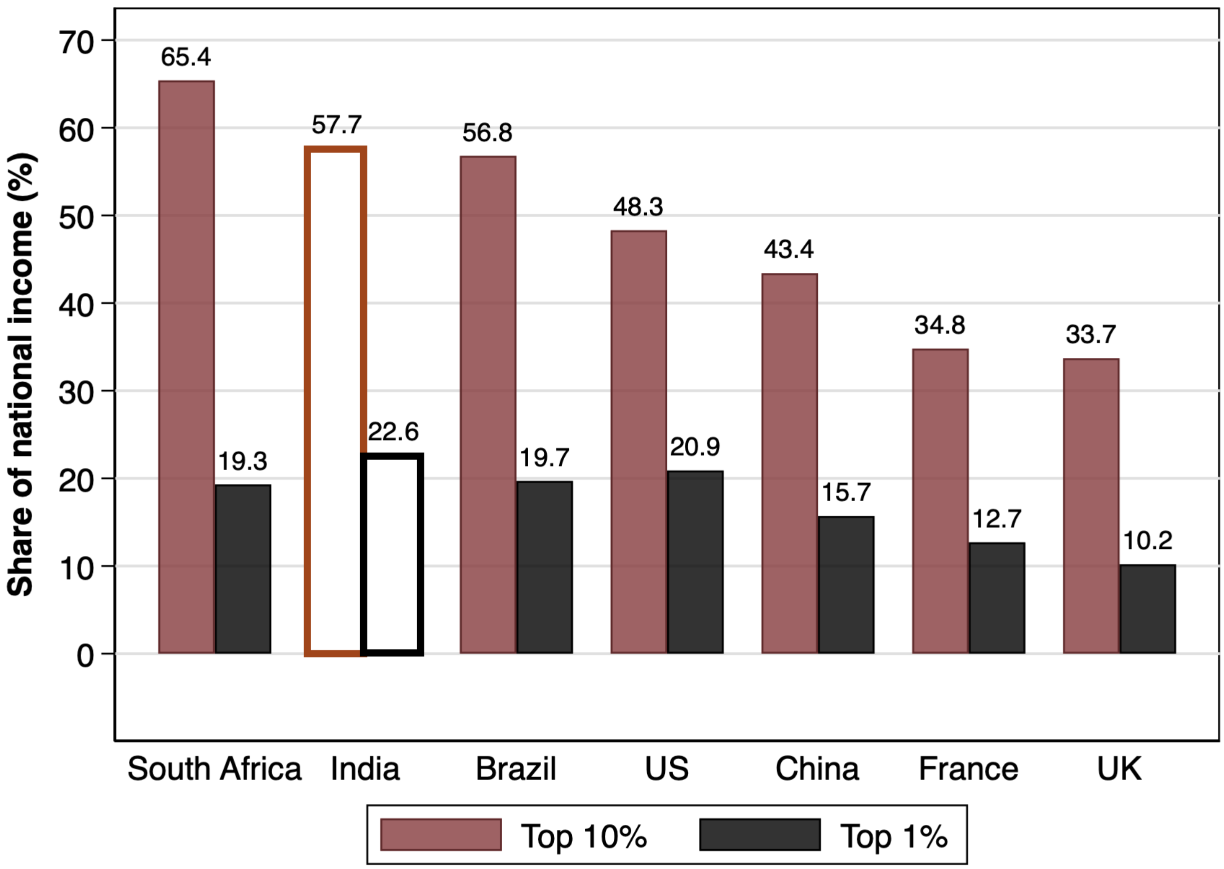
Figure 14: Wealth-to-income ratios across the distribution, 2002 & 2022



Note: Wealth-to-income ratios estimated as $W(p)/Y(p)$, where $p \in (0,1)$ denotes fractiles and $W(.)$ and $Y(.)$ are the associated quantile functions of the wealth and income distributions respectively. They are plotted on log-scale. The dashed lines show India's aggregate wealth-to-income ratio (β) in 2002 and 2022. The estimates for 2002 begin at $p0.08$ because the ratios are negative below that point.

Sources: Authors' estimates combining national income accounts aggregates, national wealth aggregates, tax tabulations, Forbes data and surveys (income, consumption, wealth).

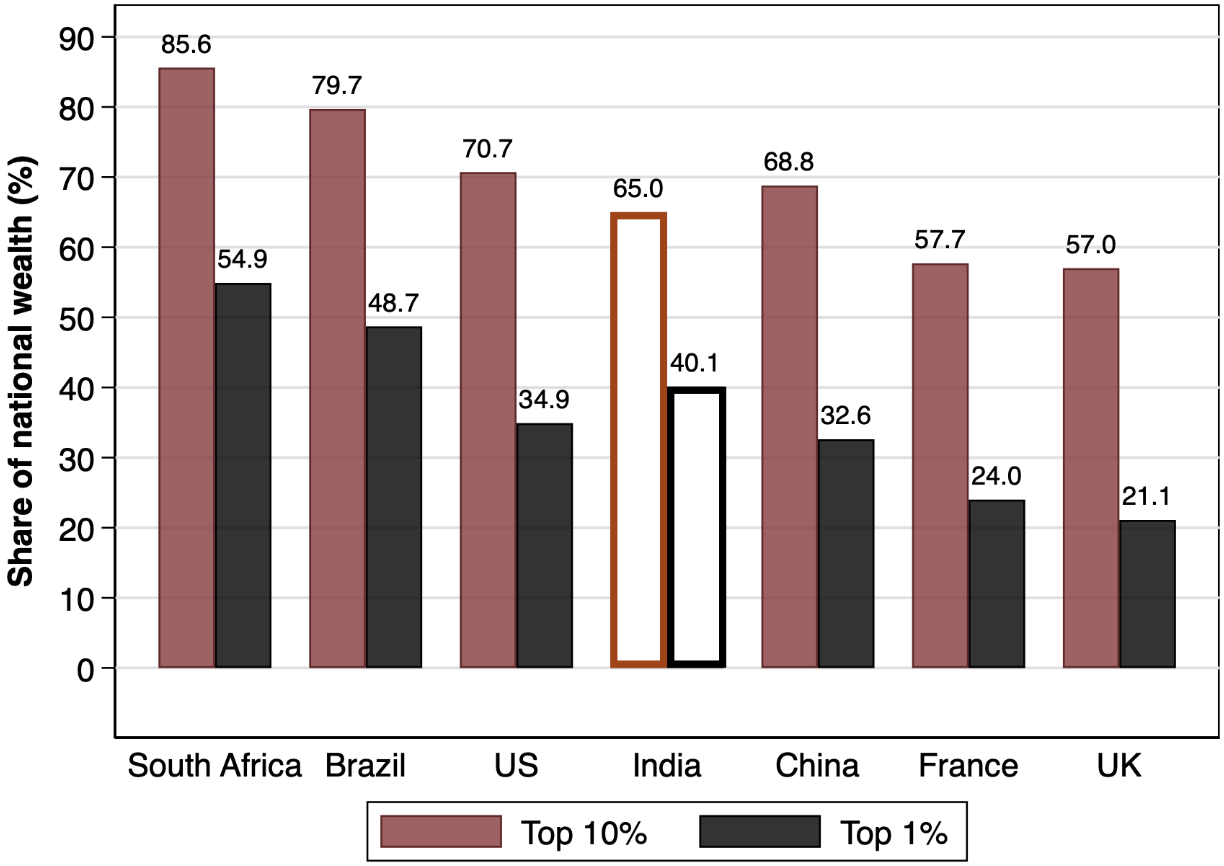
Figure 15a: Top income shares in global perspective, 2022-23



Note: The figure compares India's top 10% and top 1% income shares with a handful of countries that include some of the most unequal ones. In 2022-23, India's top 1% income was not only the highest among these countries but among the very highest anywhere in the world.

Sources: Authors' estimates for India and WID data based on country-specific studies for rest.

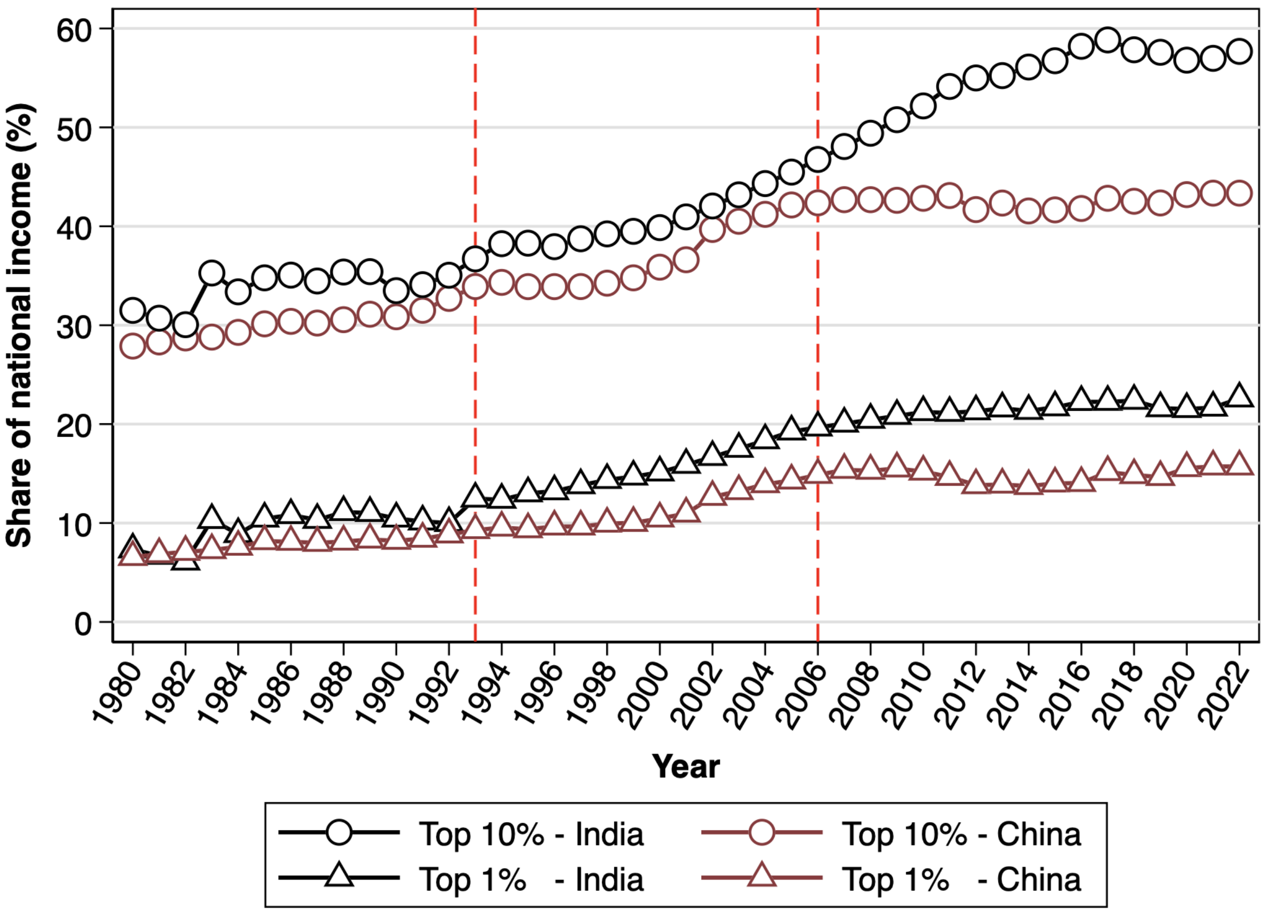
Figure 15b: Top wealth shares in global perspective, 2022-23



Note: The figure compares India's top 10% and top 1% wealth shares with a handful of countries that include some of the most unequal ones. In 2022-23, India's top 1% wealth share was higher than the US and China and closing in fast on Brazil.

Sources: Authors' estimates for India and WID data based on country-specific studies for rest.

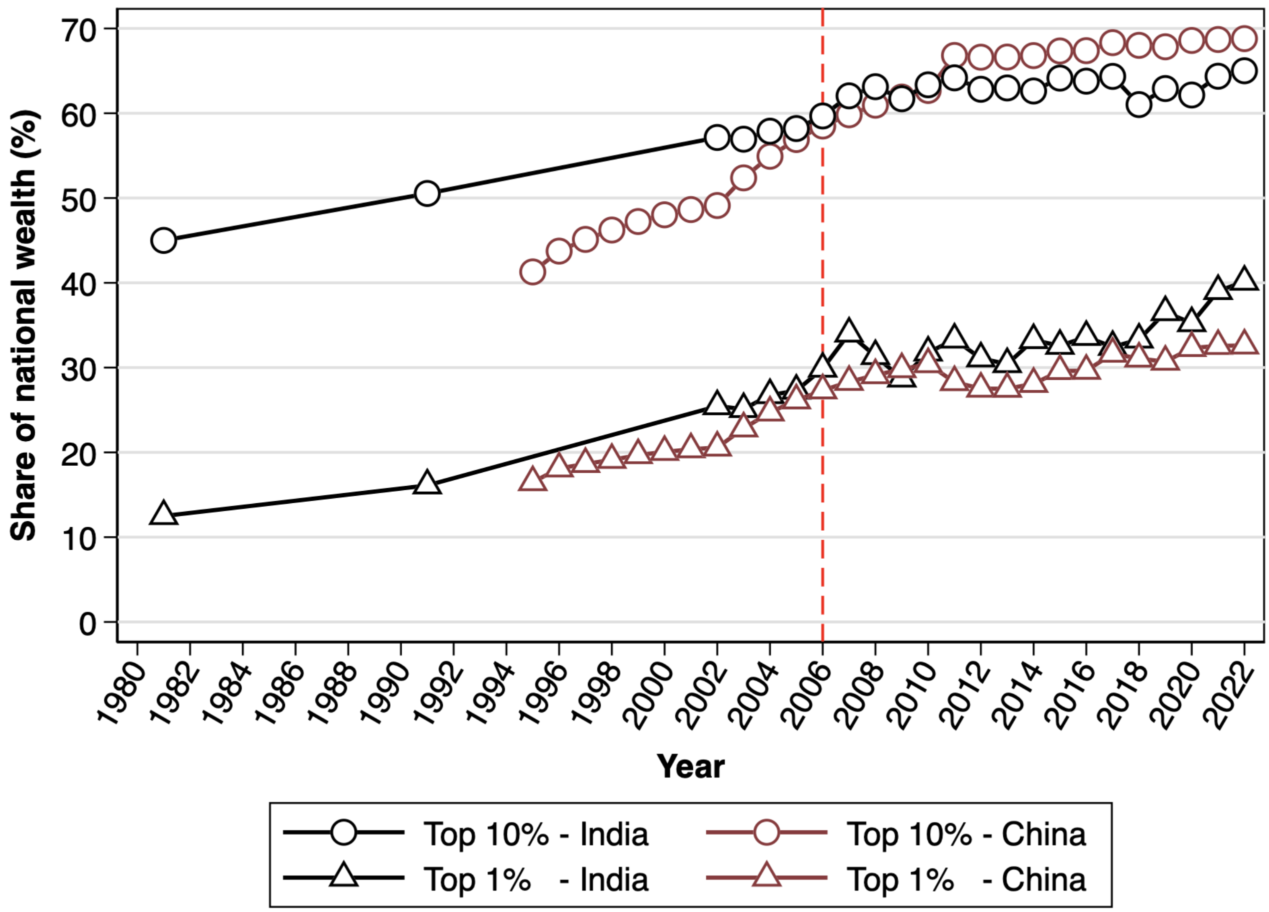
Figure 16a: Top income shares, India vs. China, 1980-2022



Note: The figure compares top 10% and top 1% income shares in India and China for the period 1980-2022. The red lines highlight the two trend-breaks we identify: 1993 when Indian top shares started slipping ahead and 2006 when Chinese top shares stabilized (see Section 6 for details).

Sources: Authors' estimates for India and WID data for China based on Piketty et al. (2019).

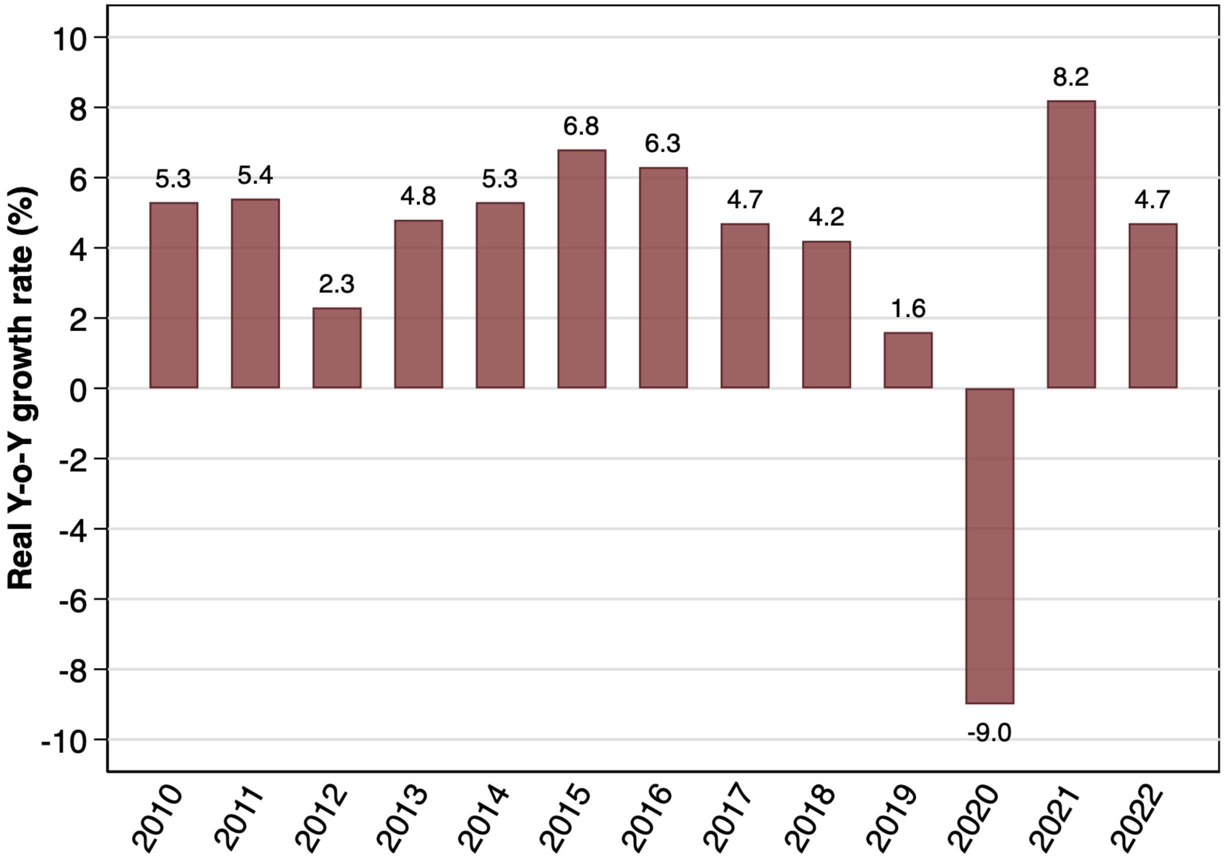
Figure 16b: Top wealth shares, India vs. China, 1980-2022



Note: The figure compares top 10% and top 1% wealth shares in India and China for the period 1980-2022. The red line points to one of the trend-breaks we identify in the income series: 2006 when Chinese top income shares stabilized (see Section 6 for details).

Sources: Authors' estimates for India and WID data for China based on Piketty et al. (2019).

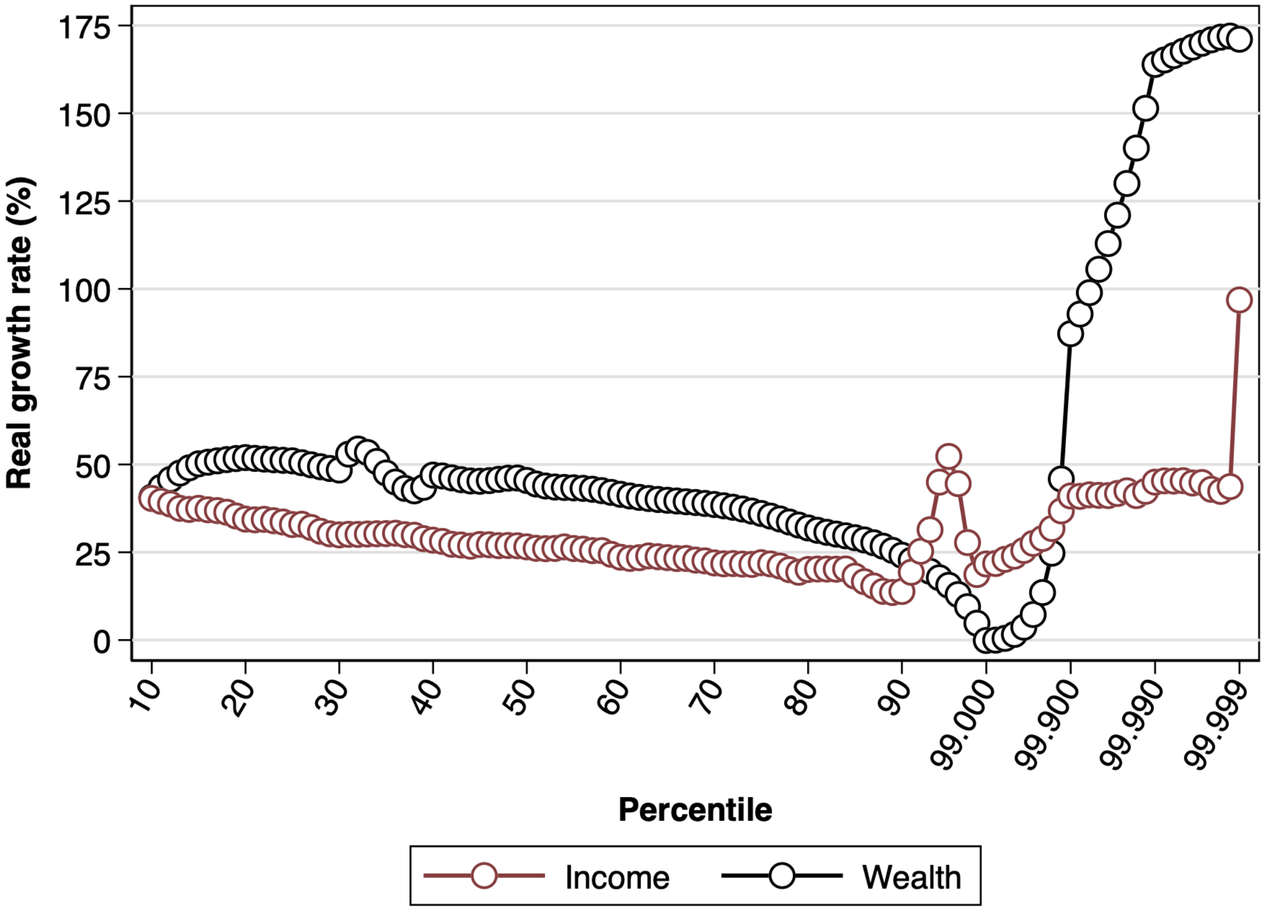
Figure 17a: Annual income growth rates, 2010-2022



Note: The figure presents real year-on-year growth rates of average incomes, defined as net national income divided by the adult population, between 2010 and 2022.

Sources: WID data for pre-2014 and authors' estimates combining net national income aggregates from Table 1.1, Statistical Appendix, Economic Survey 2022-23 for aggregate net national incomes with adult population from United Nations World Population Prospects and the GDP deflator from World Bank's World Development Indicators database.

Figure 17b: Income and wealth growth incidence curves, 2014-2022

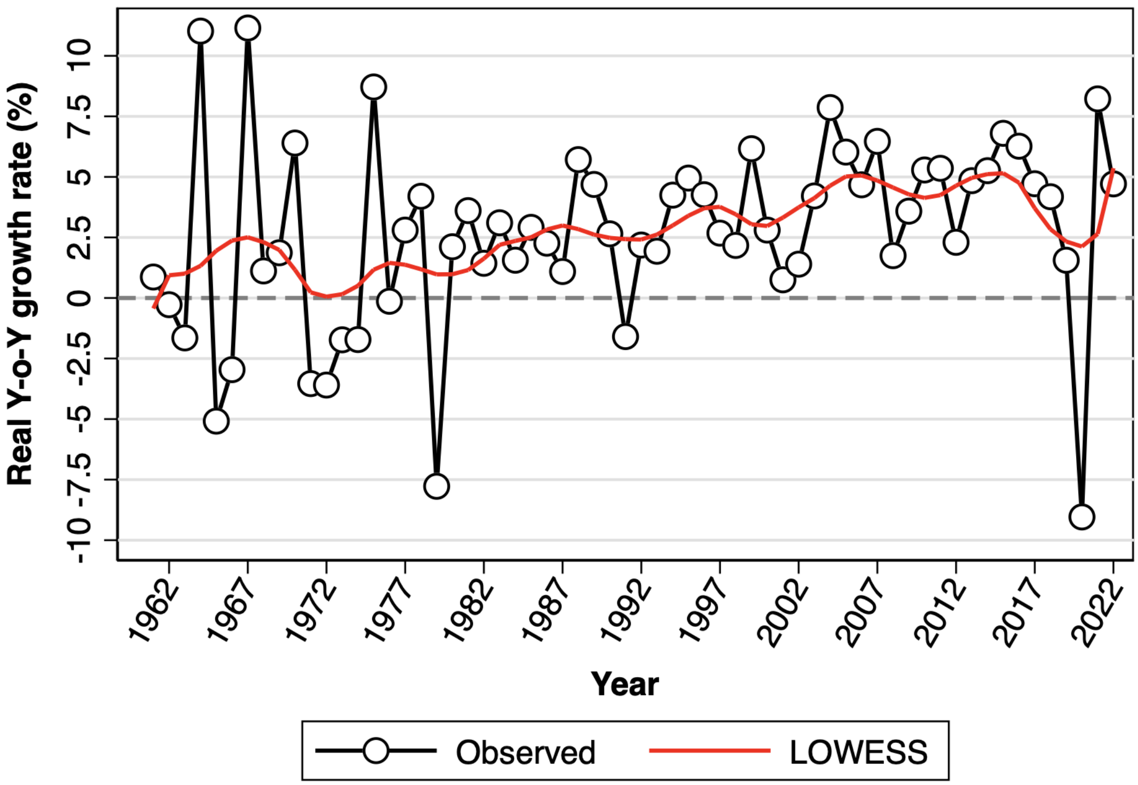


Note: The figure presents the real cumulative growth rate of incomes and wealth at each percentile starting *p*10 for the period 2014-2022. Growth rates were as high as 175% and 100% at the top of the wealth and income distributions compared to around 50% and 25% at their respective medians. With regards to wealth, we believe the declining growth rates between *p*90 and *p*99 are an artefact of the growing non-representativeness of AIDIS, as Anand and Kumar (2023) also highlight. Unfortunately, the rich lists at our disposal cover such few individuals that we prefer to use them only to correct the survey-based distribution for the top 0.1% till 2017 and the top 0.5% 2018 onwards - very conservative choices in our opinion.

Sources: Authors' estimates combining national income and wealth aggregates, surveys (consumption, income, wealth), tax tabulations, and Forbes billionaire rankings.

A Data Appendix - Long run growth rates of Indian incomes

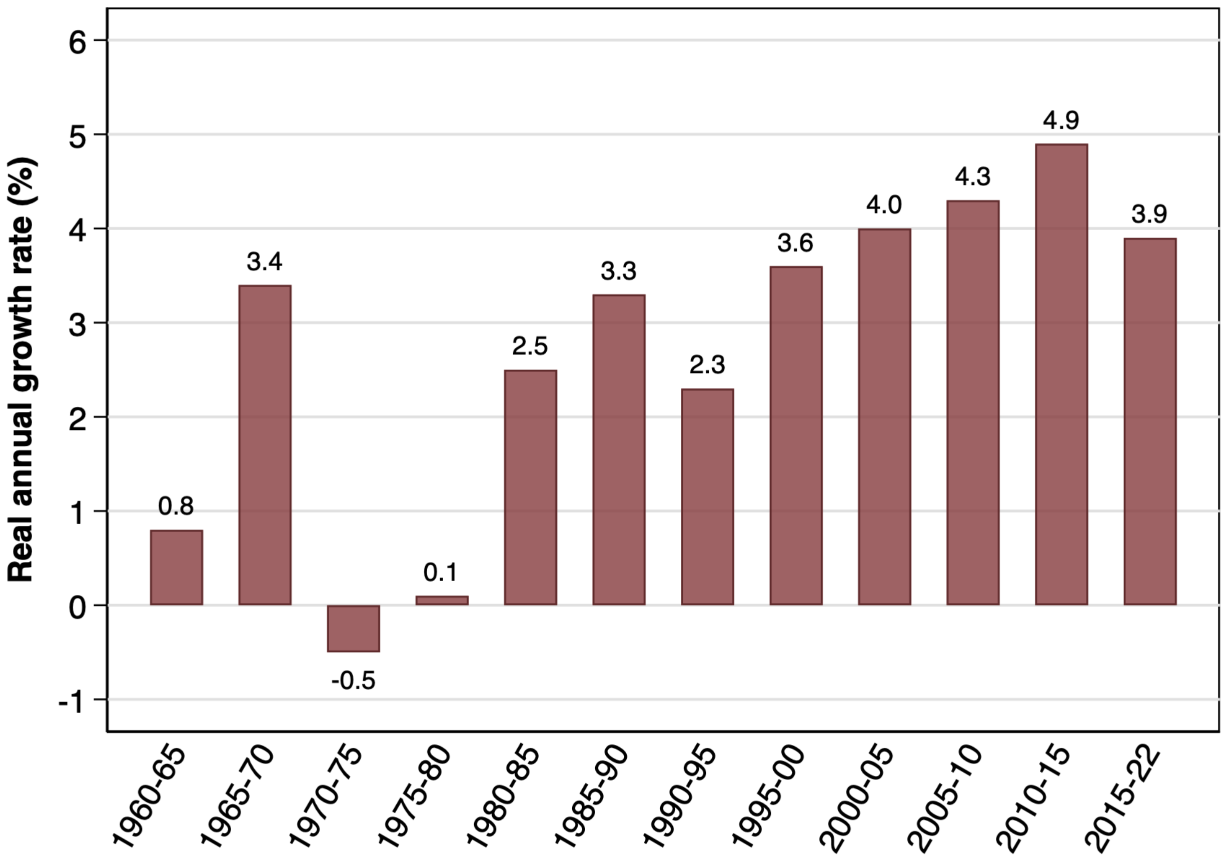
Figure A.1: Annual growth rates, 1960-2022



Note: The figure presents the real year-on-year growth rate of average incomes in India between 1960-2022. Nominal values converted to 2022 prices using the GDP deflator. LOWESS estimated via a locally weighted regression with a bandwidth of 0.2 and tri-cube weights.

Sources: Authors' estimates using national income data from WID for pre-2014 and Table 1.1, Statistical Appendix, Economic Survey 2022-23 for post-2014.

Figure A.2: Compound annual growth rates over 5-year periods, 1960-2022



Note: The figure presents the real compound annual growth rate of average incomes in India for successive 5-year periods between 1960 and 2022, with the last period being 7 years. Nominal values converted to 2022 prices using the GDP deflator sourced from World Bank's World Development Indicator database.

Sources: Authors' estimates using national income data from WID for pre-2014 and Table 1.1, Statistical Appendix, Economic Survey 2022-23 for post-2014.

B Data Appendix - Income series

Table B.1: Per-adult pre-tax national income shares (%), 1951-2022

Year	Bottom 50%	Middle 40%	Top 10%	Top 1%	Top 0.1%
1951	20.6	42.8	36.7	11.5	4.4
1952	20.7	43.2	36.1	11.8	4.7
1953	21.1	42.9	36.0	11.1	4.2
1954	20.4	41.7	37.9	12.6	4.8
1955	19.2	41.3	39.6	13.7	5.3
1956	19.6	41.2	39.2	13.8	5.3
1957	20.6	40.8	38.6	14.1	5.6
1958	20.1	41.2	38.8	13.6	5.2
1959	20.6	41.8	37.6	12.9	4.9
1960	21.3	41.8	36.9	13.2	5.3
1961	21.2	41.6	37.2	12.7	4.9
1962	21.7	41.3	36.9	12.6	4.6
1963	22.9	42.6	34.6	12.2	4.5
1964	22.6	41.9	35.5	12.0	4.1
1965	22.5	41.6	35.9	12.6	4.7
1966	22.5	41.3	36.2	12.9	4.7
1967	22.3	41.9	35.8	12.6	4.5
1968	22.6	42.3	35.2	12.1	4.3
1969	22.1	41.8	36.2	12.9	4.6
1970	22.0	41.5	36.5	13.0	4.5
1971	22.8	42.7	34.4	11.4	3.8
1972	23.0	42.9	34.1	11.0	3.6
1973	23.1	42.9	34.0	10.2	3.3
1974	23.0	43.0	34.0	10.1	3.1
1975	22.8	43.0	34.2	10.4	3.3
1976	22.7	43.0	34.4	10.3	3.1
1977	22.7	43.2	34.2	9.7	2.8
1978	22.7	43.6	33.7	9.3	2.6
1979	23.1	44.6	32.2	8.0	2.3
1980	23.3	45.2	31.5	7.3	2.0
1981	23.5	45.8	30.7	6.7	1.8
1982	23.6	46.3	30.1	6.1	1.7
1983	21.8	43.0	35.3	10.3	2.9
1984	22.4	44.2	33.4	8.9	2.5
1985	21.9	43.3	34.8	10.5	3.2
1986	21.8	43.1	35.1	10.8	3.3
1987	22.0	43.5	34.5	10.3	3.2
1988	21.7	42.9	35.4	11.1	3.5

Per-adult pre-tax national income shares (%), 1951-2022 (cont'd)

Year	Bottom 50%	Middle 40%	Top 10%	Top 1%	Top 0.1%
1989	21.7	42.9	35.4	11.0	3.4
1990	22.4	44.1	33.5	10.5	2.9
1991	22.2	43.7	34.1	10.2	2.7
1992	21.9	43.0	35.1	10.0	2.8
1993	21.4	41.9	36.7	12.5	3.8
1994	20.9	40.9	38.2	12.4	3.8
1995	20.9	40.8	38.3	13.0	4.9
1996	21.1	40.9	37.9	13.2	4.6
1997	20.9	40.4	38.7	13.8	4.6
1998	20.8	40.0	39.2	14.4	4.6
1999	20.7	39.8	39.5	14.7	4.7
2000	20.6	39.5	39.9	15.1	4.9
2001	20.2	38.9	41.0	16.7	5.4
2002	19.7	38.2	42.1	17.5	5.4
2003	19.3	37.5	43.2	18.4	6.1
2004	18.8	36.8	44.3	18.4	6.1
2005	18.4	36.1	45.5	19.3	6.4
2006	17.9	35.3	46.8	19.7	6.7
2007	17.5	34.5	48.1	20.1	7.1
2008	17.0	33.6	49.4	20.4	7.4
2009	16.5	32.7	50.8	20.8	7.8
2010	16.0	31.8	52.2	21.2	8.1
2011	15.3	30.5	54.1	21.1	8.3
2012	15.1	29.9	55.0	21.3	8.2
2013	15.0	29.8	55.2	21.6	8.5
2014	14.7	29.2	56.1	21.3	8.2
2015	14.5	28.7	56.7	21.7	8.2
2016	14.1	27.7	58.2	22.2	8.6
2017	13.9	27.3	58.8	22.3	8.6
2018	14.4	27.7	57.8	22.4	8.8
2019	14.7	27.7	57.6	21.6	8.3
2020	15.5	27.7	56.8	21.5	8.2
2021	15.4	27.6	57.0	21.7	8.2
2022	15.0	27.3	57.7	22.6	9.6

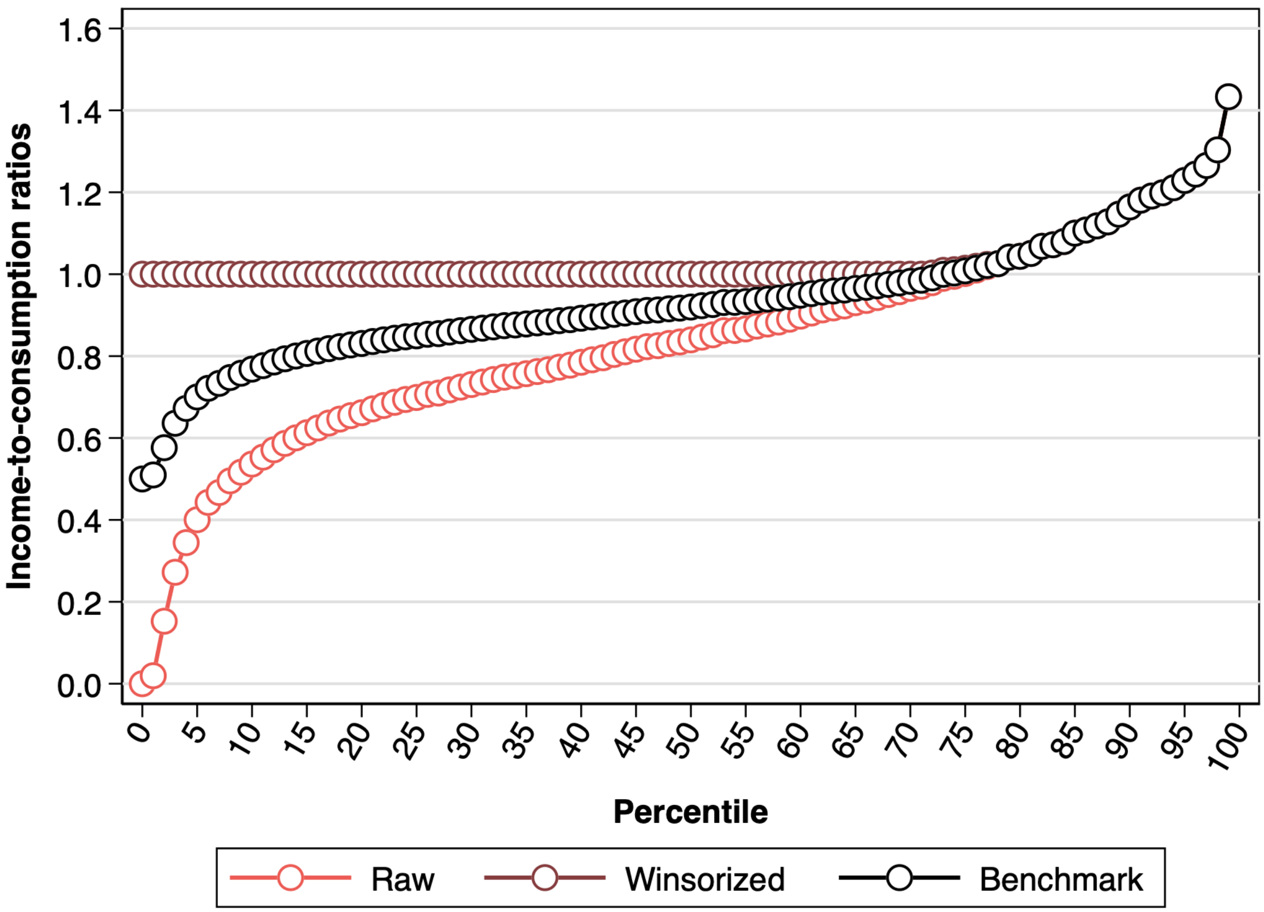
Sources: Authors' estimates combining national income accounts aggregates, tax tabulations, and income and consumption surveys.

Table B.2: Merge point for tax tabulations, 1953-2020

Year	Fractile	Year	Fractile
1953	0.998	1989	0.993
1954	0.998	1990	0.992
1955	0.998	1991	0.991
1956	0.997	1992	0.990
1957	0.996	1993	0.990
1958	0.996	1994	0.988
1959	0.996	1995	0.987
1960	0.996	1996	0.984
1961	0.996	1997	0.981
1962	0.995	1998	0.976
1963	0.994	1999	*
1964	0.994	2000	*
1965	0.994	2001	*
1966	0.994	2002	*
1967	0.993	2003	*
1968	0.993	2004	*
1969	*	2005	*
1970	0.993	2006	*
1971	0.993	2007	*
1972	*	2008	*
1973	0.993	2009	*
1974	0.993	2010	*
1975	0.993	2011	0.946
1976	0.993	2012	0.940
1977	0.995	2013	0.935
1978	0.993	2014	0.931
1979	0.997	2015	0.923
1980	0.997	2016	0.919
1981	0.997	2017	0.909
1982	0.998	2018	0.906
1983	0.996	2019	0.915
1984	0.997	2020	0.912
1985	0.996		
1986	0.995		
1987	0.995		
1988	0.994		

Note: The income series presented in this paper is estimated by combining surveys (income and consumption) with tax tabulations. This table presents the fractile $p \in (0, 1)$ in the distribution from where the tax tabulations kick-in for each year. * denotes years for which tax data is unavailable.

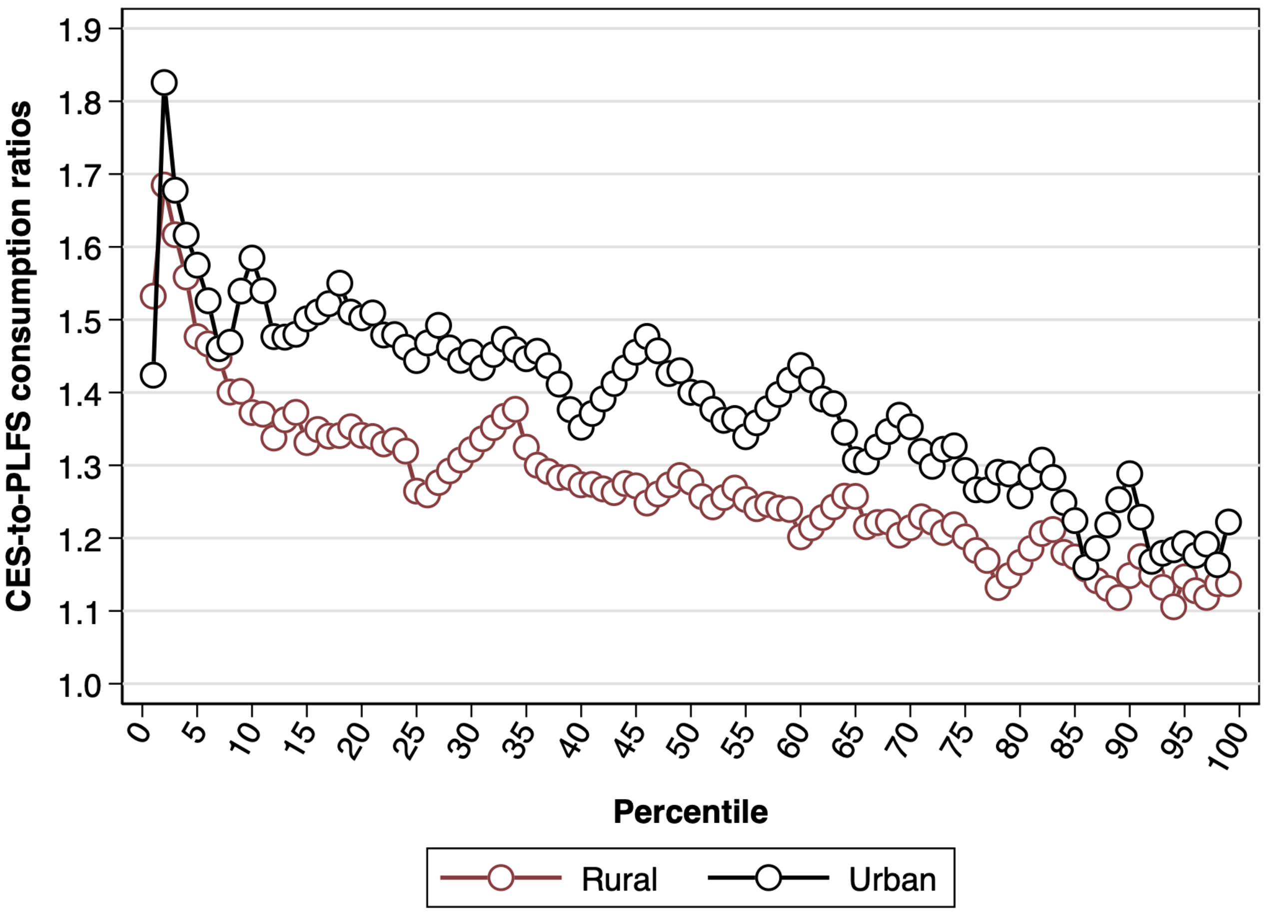
Figure B.1: Income-to-consumption ratios



Note: The figure presents the scaling ratios we use to move from the consumption distribution to an income distribution. These ratios are estimated as $\alpha_p = y_p/c_p$ where $p \in (0, 1)$ denotes percentiles and y and c income and consumption respectively. The 'raw' variant are the ratios as observed in the data, the 'winsorized' variant restricts the observed ratios to be at-least as large as 1, and the 'benchmark' variant (used in our series) is an average of the raw and winsorized variants (see Section 2.2 for details).

Sources: Authors' estimates using unit-level IHDS data.

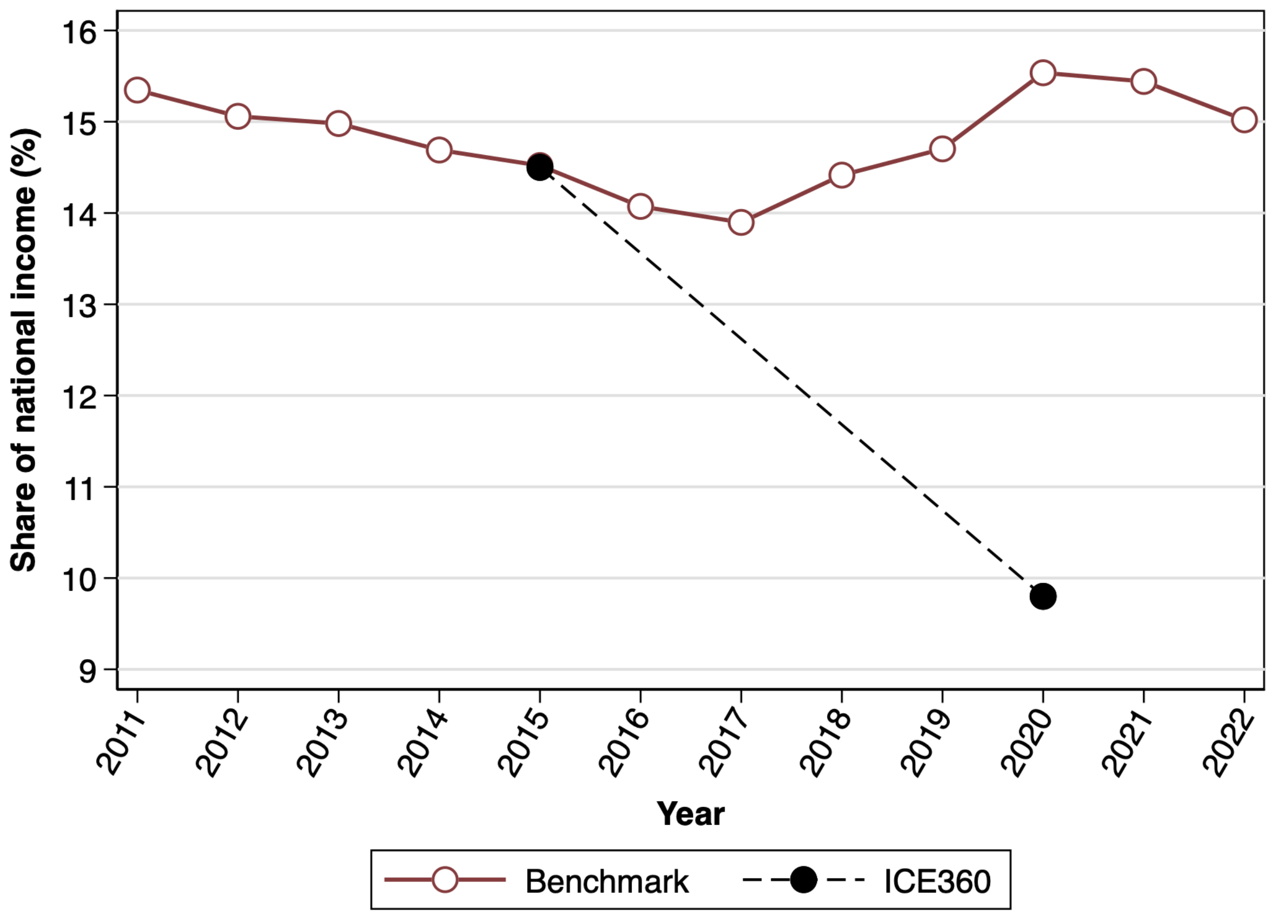
Figure B.2: PLFS-to-CES consumption ratios



Note: The figure presents the scaling ratios used to move from consumption observed in PLFS to a CES-comparable consumption measure. To estimate these, we use generalized Pareto interpolation to extract a full distribution of monthly per-capita consumption expenditure (MPCE) for rural and urban areas available in the tabulated summary of the CES 2017-18. We then estimate the MPCE distribution from the unit-level PLFS 2017-18 data using the 'usual' consumption expenditure variable in the dataset. Then we divide the CES consumption by the PLFS consumption at each percentile.

Sources: Authors' estimates combining unit-level PLFS data and CES tabulations for 2017-18.

Figure B.3: Bottom 50% national income shares, 2011-2022



Note: The figure presents the share of national income going to the bottom 50% as per two different sources for estimating bottom incomes (top incomes estimated from tax tabulations in both series). The first is our benchmark series based on CES and PLFS and the other based on the per-capita income recorded in the ICE360 survey rounds conducted in 2015-16 and 2020-21. The ICE360 survey suggests a sharp drop in bottom 50% shares (and rise in top 10% shares) over the 2015-2020 period with a steeply upward sloping growth incidence curve suggesting our benchmark series is likely to be conservative.

Sources: Authors' estimates combining national income accounts, surveys on income and consumption, and tax tabulations.

C Data Appendix - Wealth series

Table C.1: Per-adult national wealth shares (%), 1961-2023

Year	Bottom 50%	Middle 40%	Top 10%	Top 1%	Top 0.1%
1961	11.4	43.7	44.9	12.9	3.2
1971	11.0	45.0	44.0	12.2	2.8
1981	10.9	44.1	45.0	12.5	2.7
1991	8.8	40.7	50.5	16.1	4.4
2002	6.9	35.9	57.1	25.4	15.1
2003	7.0	36.1	56.9	25.1	14.7
2004	6.8	35.3	57.9	26.8	16.6
2005	6.8	35.0	58.2	27.3	17.2
2006	6.5	33.8	59.7	29.9	20.1
2007	6.1	31.8	62.1	34.0	24.8
2008	6.3	30.5	63.1	31.4	19.1
2009	6.6	31.7	61.7	28.7	16.0
2010	6.3	30.3	63.4	31.8	19.6
2011	6.2	29.7	64.2	33.3	21.4
2012	6.4	30.8	62.8	31.1	18.5
2013	6.4	30.7	63.0	30.4	18.8
2014	6.4	31.0	62.6	33.3	18.0
2015	6.2	29.7	64.1	32.6	21.4
2016	6.2	30.0	63.8	33.6	20.6
2017	6.1	29.5	64.3	32.4	21.8
2018	7.2	31.8	61.0	33.1	23.3
2019	6.8	30.3	62.9	36.5	25.6
2020	6.9	30.9	62.2	35.2	24.2
2021	6.5	29.1	64.4	39.0	28.4
2022	6.4	28.6	65.0	40.1	29.7
2023	6.5	29.0	64.6	39.5	29.0

Note: Estimates based solely on AIDIS till 1991 and combining AIDIS with Forbes rich list 2002 onwards. Estimates for 2023 are tentative as they are based on Hurun's 2023 rich list (truncated at the top 100 individuals) since Forbes data is not yet out. See footnote 31 for details.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys (AIDIS) and Forbes billionaire rankings.

Table C.2: Growth of very high net worth individuals, 1988-2022

Year	Forbes		Hurun	
	Number of individuals	*Net wealth as % of NNI*	*Number of individuals*	*Net wealth as % of NNI*
1988	1	1.7	*	*
1989	1	0.5	*	*
1990	1	0.5	*	*
1991	1	0.6	*	*
1992	1	1.0	*	*
1993	1	0.7	*	*
1994	2	1.3	*	*
1995	2	1.1	*	*
1996	3	1.5	*	*
1997	4	1.6	*	*
1998	2	1.5	*	*
1999	7	2.7	*	*
2000	9	7.0	*	*
2001	4	3.4	*	*
2002	5	3.2	*	*
2003	6	3.1	*	*
2004	8	4.6	*	*
2005	9	5.1	*	*
2006	19	9.3	*	*
2007	33	16.7	*	*
2008	50	27.8	*	*
2009	21	8.3	*	*
2010	47	15.1	*	*
2011	52	14.4	*	*
2012	46	12.0	100	14.4
2013	52	11.8	141	17.4
2014	51	10.9	230	24.3
2015	90	15.5	*	*
2016	85	12.5	338	21.8
2017	101	14.0	*	*
2018	119	18.0	831	25.4
2019	106	16.1	*	*
2020	102	13.5	829	26.5
2021	140	21.5	*	*
2022	162	24.6	1103	27.5

Note: Forbes billionaire rankings track all individuals with net wealth exceeding 1 billion USD MER since 1988. Hurun rich lists track all individuals with net wealth exceeding 1000 crore INR (~ 120 million USD MER on March 1st, 2024) since 2012. Post-2014, Hurun has only published a full list every alternate year. NNI = Net national income.
Sources: Authors' estimates combining national accounts aggregates with Forbes rankings and Hurun rich lists.

Table C.3: Merge point for rich lists, 2002-2022

Year	Fractile
2002	0.999
2003	0.999
2004	0.999
2005	0.999
2006	0.999
2007	0.999
2008	0.999
2009	0.999
2010	0.999
2011	0.999
2012	0.999
2013	0.999
2014	0.999
2015	0.999
2016	0.999
2017	0.999
2018	0.995
2019	0.995
2020	0.995
2021	0.995
2022	0.995
2023	0.995

Note: The wealth series presented in this paper is estimated by combining wealth surveys (AIDIS) with Forbes rich lists. The table presents the fractile $p \in (0, 1)$ in the distribution till where the survey is considered representative, above which a correction is applied based on data from the rich list. From 2002 till 2017, we assume that the survey is non-representative for the top 0.1% and from 2018 onwards for the top 0.5%. The downward adjustment is driven by the growing non-representativeness at the top in AIDIS, especially in the 2018 round (see Section 2.3 for details).

Table C.4: Choice of threshold and top wealth shares, 2018

p_0	Top 10%	Top 5%	Top 1%	Top 0.1%	Top 0.01%	Top 0.001%
90.0	71.0	62.2	45.6	29.0	18.2	11.1
91.0	70.7	62.0	45.6	29.1	18.2	11.1
92.0	70.5	61.8	45.5	29.1	18.3	11.2
93.0	70.4	61.7	45.5	29.1	18.3	11.2
94.0	69.9	61.2	45.3	29.1	18.4	11.4
95.0	69.6	60.8	45.1	29.2	18.5	11.4
96.0	68.9	60.0	44.7	29.1	18.7	11.6
97.5	67.6	58.2	43.5	28.9	18.8	11.9
98.0	66.9	57.3	42.7	28.7	18.9	12.1
98.5	66.0	56.2	41.6	28.3	18.9	12.3
99.0	64.7	54.6	39.7	27.6	18.9	12.5
99.5	63.4	52.8	37.4	26.6	18.7	12.7
99.9	61.0	49.8	33.3	23.1	17.5	12.8

Note: The wealth series presented in this paper is estimated by combining wealth surveys (AIDIS) with Forbes rich lists. The table presents how the choice of the fractile $p_0 \in (0, 1)$ in the distribution till where the survey is considered representative affects top shares for the year 2018 (last AIDIS round). Not surprisingly, we see that by assuming lower p_0, we get higher top shares (except for at the very top - Top 0.001%). The top 10% share declines from 71% with $p_0 = 90.0$ to 61% with $p_0 = 99.9$. We consider our choice of $p0$ for our benchmark series (99.99 till 2017 and 99.95 2018 onwards) as conservative given that the survey is likely to be non-representative even at lower percentiles within the top decile.

Sources: Authors' estimates combining AIDIS 2018 round and Forbes 2018 billionaire rankings using *standard* Pareto interpolation.

Table C.5: Fraction of rich list wealth recovered using two Pareto methods

Year	Forbes		Hurun	
	Constant (%)	*Log-Linear (%)*	*Constant (%)*	*Log-Linear (%)*
2002	2	94	*	*
2003	4	97	*	*
2004	7	103	*	*
2005	5	101	*	*
2006	6	107	*	*
2007	6	109	*	*
2008	54	81	*	*
2009	39	93	*	*
2010	34	89	*	*
2011	31	94	*	*
2012	26	100	160	93
2013	18	106	160	91
2014	17	103	112	102
2015	10	118	34	130
2016	8	123	93	103
2017	9	129	25	131
2018	18	118	170	85
2019	21	114	53	107
2020	11	115	171	79
2021	20	107	20	112
2022	20	103	136	90
2023	*	*	18	109

Note: To top-correct AIDIS, we simulate wealth for the top 0.1% (0.5% 2018 onwards) for which we need the Pareto parameter α treating average wealth at $p99.9$ ($p99.5$ 2018 onwards) in the survey as the wealth-level from where the Pareto law applies. As described in Section 2.3, we start with two approaches to recover α from rich lists: the constant inverted Pareto coefficient method and the log-linear method. We then judge the approaches based on their ability to recover the sum total wealth reported in the rich lists. Let us say the *total* net wealth of the wealthiest N persons covered in the rich list equals $\overline{W}$. This table presents the fraction of $\overline{W}$ we are able to recover with the top N individuals in our simulations based on the two approaches using both Forbes and Hurun rich lists. The constant inverted coefficient method falls far short (for Forbes mainly) whereas the log-linear method is on average quite close to the mark with both Forbes and Hurun. Hence, we pick the log-linear approach for our benchmark series. * denotes rich list data not available for those years.

Table C.6: Decomposition of total wealth (surveys), 1981-2018

Assets	1981	1991	2002	2012	2018
Land	60.1	59.3	56.6	56.4	58.2
Building	27.0	29.5	30.3	32.9	28.5
Livestock	4.3	2.6	3.0	0.8	0.7
Agricultural machinery etc.	4.4	4.4	5.0	3.0	3.1
Financial assets	4.2	4.2	5.2	7.0	9.6
Loans receivable	0.2	0.2	0.2	0.2	0.2

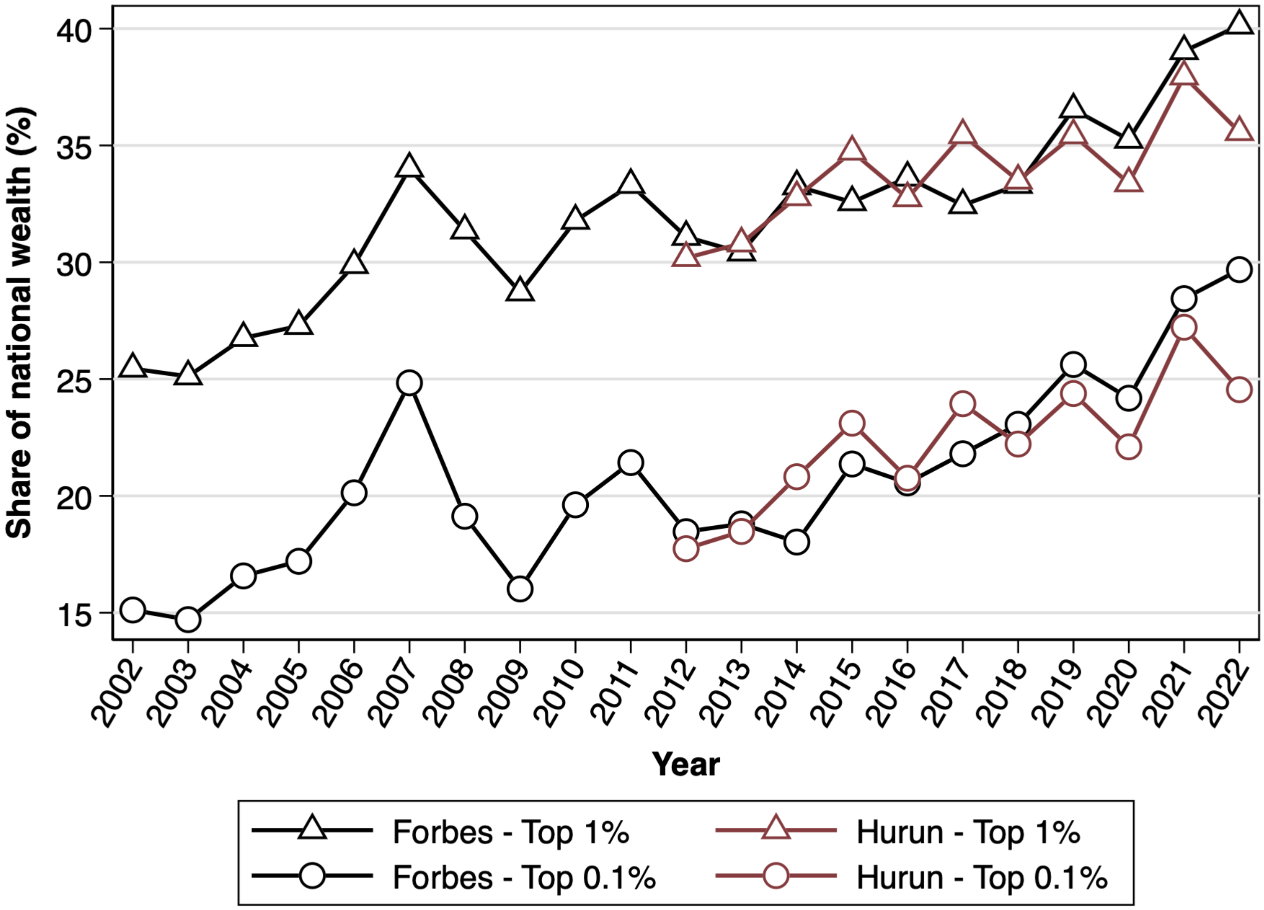
Sources: Authors' estimates using AIDIS micro-data (1991-2018) and published report (1981).

Table C.7: Coverage of assets & liabilities in wealth surveys, 1961-2018

Type of assets / liabilities	1961	1971	1981	1991	2002	2012	2018
Physical Assets							
Land	✓	✓	✓	✓	✓	✓	✓
Building	✓	✓	✓	✓	✓	✓	✓
Livestock	✓	✓	✓	✓	✓	✓	✓
Agricultural implements, machinery, etc.	✓	✓	✓	✓	✓	✓	✓
Non-farm business equipment	✓	✓	✓	✓	✓	✓	✓
Transport equipment	✓	✓	✓	✓	✓	✓	✓
Durable assets	✓	✓	✓	✓	✓	×	×
Financial Assets							
Shares, debentures, mutual funds, etc.	✓	✓	✓	✓	✓	✓	✓
Deposits (company, bank, post office, etc.)	✓	✓	✓	✓	✓	✓	✓
Dues receivable – Cash	✓	✓	✓	✓	✓	✓	✓
Dues receivable – Kind	✓	✓	✓	✓	✓	✓	✓
Liabilities							
Dues payable – Cash	✓	✓	✓	✓	✓	✓	✓
Dues payable – Kind	✓	✓	✓	✓	✓	✓	✓

Sources: Authors' compilation from documentation of successive AIDIS rounds.

Figure C.1: Top wealth shares - Forbes vs. Hurun, 2002-2022



Note: The figure presents the share of national wealth going to the top 1% and top 0.1% based on two different rich lists used to top-correct the AIDIS surveys. The Forbes series is our benchmark while we present Hurun as a robustness check. As we see, both levels and trends of top shares track each other closely across both sources.

Sources: Authors' estimates combining national wealth aggregates, wealth surveys (AIDIS) and Forbes and Hurun rich lists.

D Technical Appendix - Generalized Pareto interpolation

To extract a distribution of top incomes from the tabulated tax data, we rely on generalized Pareto interpolation methods developed by Blanchet et al. (2022). Going back to Vilfredo Pareto's empirical observation based on income data from Swiss cantons that top incomes are well approximated by a power law (Pareto, 1896), various interpolation methods relying on the property of 'scale in-variance' of the Pareto distribution have been deployed over the years. Consider a Pareto distributed variable with a probability density function (PDF) and complementary cumulative distribution function (CCDF) given by:

$$\text{PDF:}\quad \mathbb{P}(Y \in [y, y+dy]) = f_Y(y) = \left(\frac{\alpha}{y_0}\right)\left(\frac{y_0}{y}\right)^{\alpha+1} \quad ; \quad \text{CCDF:}\quad \mathbb{P}(Y > y) = 1 - F_Y(y) = \left(\frac{y_0}{y}\right)^{\alpha}$$

where Y is the continuously distributed random variable denoting incomes (or wealth), y a realization of it, $y_0 > 0$ some minimum income level from which the distributional form applies, and $\alpha > 0$ the tail parameter which describes the thickness of the tail of the distribution. Lower values of α imply the probability of observing extreme realizations decays slower to 0, leading to fatter tails.[50] Two things to note. First, taking logs on the CCDF we get:

$$\log\left(1 - F_Y(y)\right) = c - \alpha \log\left(\ y\ \right)$$

where $c = \alpha \log(y_0)$ is some constant given that y_0 is a constant. We then can recover the tail parameter α by regressing the log of normalized rank on the log of income (or wealth). This is what we refer to as the "log-linear" approach in section 2.3. Second, based on the PDF and CCDF above, it is easily shown that the ratio of average incomes above a threshold divided by the threshold itself is given by:

$$\frac{\mathbb{E}(y \mid y > y^*)}{y^*} = \frac{\alpha}{\alpha - 1} = \beta = \text{“Inverted Pareto coefficient”}$$

That this ratio does not depend on the threshold itself (hence scale invariant) means once β is identified (from the tax data or rich lists), average incomes above any threshold $y^* > y_0$ can be interpolated as $\mathbb{E}(y \mid y > y^*) = \beta y^*$, which in turn allows estimating income shares at any threshold above y_0. In this context, β has a more intuitive interpretation than the tail parameter α and has come to be known as the "inverted Pareto coefficient" in the literature (Atkinson et al., 2011). Higher values of β imply more concentration of incomes, i.e. fatter tails. Building on this framework, "standard" Pareto interpolation techniques have typically relied on the strict Paretian assumption that an exact power law with a constant tail parameter β holds within each income bracket to interpolate the distribution between two bracket thresholds from tabulated tax data (Pareto, 1896; Kuznets, 1953;

[50] With $2 < \alpha$, both the mean and variance are finite; with $1 < \alpha \leq 2$, the mean is finite but not the variance; with $\alpha \leq 1$, both the mean and variance are infinite.

Piketty and Saez, 2003; Banerjee and Piketty, 2005). However, these methods do not make use of all the information in the tax data, and they do not always work well, even for describing top incomes. Relaxing the strict Paretian assumption that β remains constant across the distribution, instead allowing it to vary by rank, Blanchet et al. (2022) develop *generalized* Pareto curves. Letting $p \in (0, 1)$ denote rank in the distribution and $Q(p)$ its associated quantile function, the generalized Pareto coefficient is given by:

$$\beta(p) = \frac{\mathbb{E}\big(y \mid y > Q(p)\big)}{Q(p)} = \frac{1}{(1-p)Q(p)} \times \int_p^1 Q(s)\, ds$$

These yield generalized Pareto curves (typically U-shaped) summarizing the concentration of income across the distribution. Applying quintic spline interpolation to the income thresholds and averages in the tabulated tax data, the generalized Pareto interpolation algorithm is able to extract a continuous and smooth Pareto curve which is used to interpolate a full distribution from the tabulated data. As Blanchet et al. (2022) demonstrate using data from countries where tax micro-files are available, the algorithm outperforms other interpolation techniques.

E Appendix - Call for democratic access to data

Table E.1: Example of improved tax tabulations

Gross income (INR)	Total returns		Total incomes by source (INR)								
	All	*Women*	*Gross*	*Salary*	*Property*	*Business*	*LTCG*	*STCG*	*Interest*	*Loss set-off*	*Others*
< 1.5 lakh											
1.5 lakh - 2.0 lakh											
2.0 lakh - 2.5 lakh											
2.5 lakh - 3.0 lakh											
3.0 lakh - 3.5 lakh											
3.5 lakh - 4.0 lakh											
4.0 lakh - 4.5 lakh											
4.5 lakh - 5.0 lakh											
5.0 lakh - 5.5 lakh											
5.5 lakh - 9.5 lakh											
9.5 lakh - 10 lakh											
10 lakh - 15 lakh											
15 lakh - 20 lakh											
20 lakh - 25 lakh											
25 lakh - 50 lakh											
50 lakh - 1 crore											
1 crore - 5 crore											
5 crore - 10 crore											
10 crore - 25 crore											
25 crore - 50 crore											
50 crore - 100 crore											
100 crore - 500 crore											
> 500 crore											

Note: The table presents an example of what an improved tax tabulation could look like based entirely on information already collected by the tax authorities. The brackets of gross income correspond exactly to those currently published in the tax tabulations. The key improvement we propose is the linking of information on incomes from different sources into a single table, as well as the inclusion of information on the number of returns filed by women in each bracket of gross incomes. (LTCG = Long-term capital gains; STCG = Short-term capital gains; 1 lakh = 0.1 million; 1 crore = 10 million)