

## INDIAN INCOME INEQUALITY, 1922-2015: FROM BRITISH RAJ TO BILLIONAIRE RAJ?

BY LUCAS CHANCEL\*

*Paris School of Economics  
World Inequality Lab  
AND*

THOMAS PIKETTY

*Paris School of Economics  
Ecole des Hautes Etudes en Sciences Sociales  
World Inequality Lab*

We combine household surveys and national accounts, as well as recently released tax data to track the dynamics of Indian income inequality from 1922 to 2015. According to our benchmark estimates, the top 1 percent of earners captured less than 21 percent of total income in the late 1930s, before dropping to 6 percent in the early 1980s and rising to 22 percent in the recent period. Our results appear to be robust to a range of alternative assumptions seeking to address numerous data limitations. These findings suggest that much more can be done to promote inclusive growth in India. We also stress the need for more transparency on income and wealth statistics, which is key to allow an informed democratic debate on inequality.

**JEL Codes:** D31, N35, O15

**Keywords:** inequality, India, top incomes

### 1. INTRODUCTION

India introduced an individual income tax with the Income Tax Act of 1922, under the British colonial administration. From this date, up to the turn of the 20th century, the Indian Income Tax Department produced income tax tabulations, making it possible to track the long-run evolution of top incomes in a systematic manner. Using this data, Banerjee and Piketty (2005) showed that the share of fiscal income accruing to the top 1 percent earners shrank substantially from the mid-1950s to the mid-1980s, from about 13 percent of fiscal income, to less than 5 percent in the early 1980s. The trend was reversed in the mid-1980s, when pro-business, market deregulation policies were implemented. The share of fiscal held of the top 1 percent doubled from approximately 5 percent to 10 percent in 2000.

*Note:* We thank Nitin Bharti and Amory Gethin for outstanding research assistance. NSSO micro data for years 1983 and after obtained through CEPREMAP and PjSE, which we gratefully acknowledge for their support.

\*Correspondence to: L. Chancel, Paris School of Economics 48 Bd Jourdan, 75014, Paris, France (lucas.chancel@psemail.eu)

© 2019 International Association for Research in Income and Wealth

[Correction added on 13 April 2020, after first online publication: Captions of Figures 6 to 12 have been corrected, and the missing sources and notes for Figures 1 to 5 and Figure 13 have been added.]

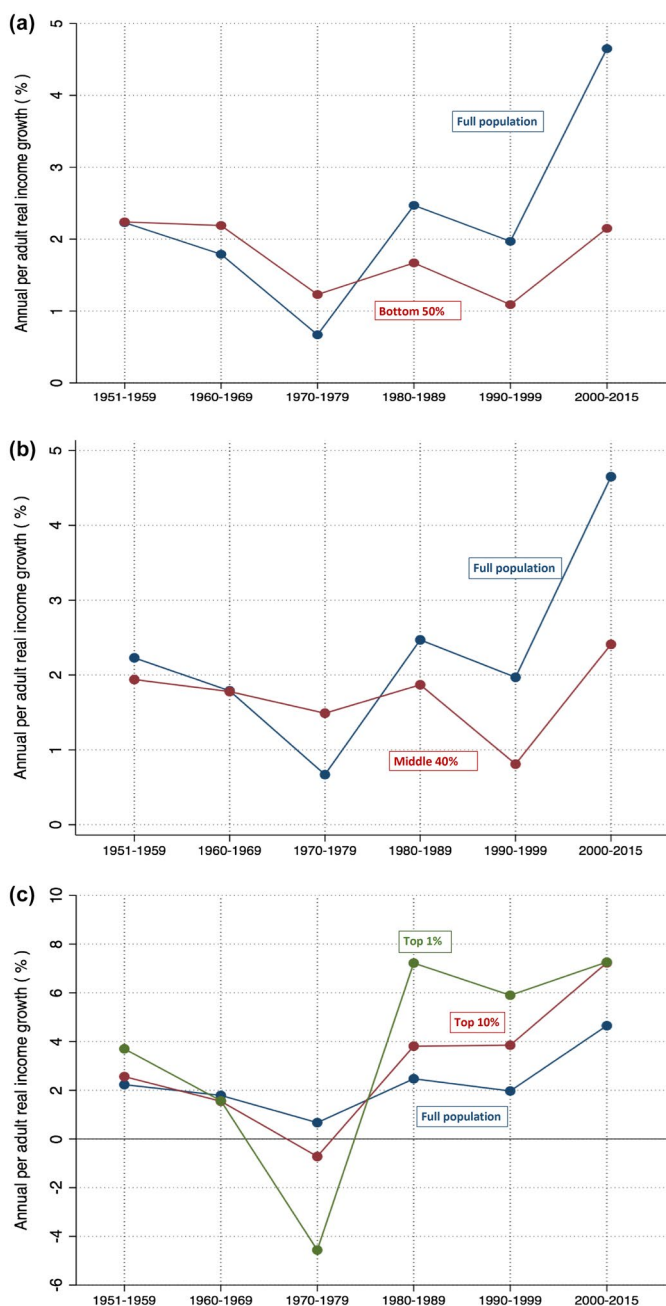


Figure 1. (a) National Income Growth in India: Full Population vs. Bottom 50% Income Group, 1951–2015. (b) National Income Growth in India: Full Population vs. Middle 40% Income Group, 1951–2015. (c) National Income Growth in India: Full Population vs. Top 1% and Top 10% Income Groups, 1951–2015

*Source:* Authors' estimates combining survey, fiscal and national accounts data. Average annual per adult real income growth rate from 1970 to 1979 was 0.67%. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

According to National Accounts estimates, post-2000 income growth has been substantially higher than in the previous decades. Average annual real income growth was below 2 percent in the 1960 and 1970s, it reached 2.5 percent in the 1980s and 2 percent in the 1990s.<sup>1</sup> Since 2000s it is of 4.7 percent on average (Figure 1). Little is known however on the distributional impacts of economic policies in India after 2000 in part because the Income Tax Department stopped publishing income tax statistics in 2000, and also because self-reported survey data does not provide adequate information concerning the top of the distribution (fiscal data is not perfect either, but it delivers higher and more plausible income levels for the top). In 2016, the Income Tax Department released tax tabulations for recent years (2011–12, 2012–13 and 2013–14), making it possible to revise and update previously published top income estimates and better inform public debates on growth and income inequality. We find that the bottom 50 percent group grew at a substantially lower rate than average growth (Figure 1a) since the 1980s. Middle 40 percent grew at a slower rate than the average (Figure 1b). On the contrary, top 10 percent and top 1 percent grew substantially faster than the average since 1980 (Figure 1c).

The first objective of this paper is to mobilize this newly released set of tax data in order to track the evolution of income inequality from 1922 to 2015. The second objective is to go beyond top income shares and produce estimates of income dynamics throughout the entire distribution using concepts that are consistent with National Accounts (following, as much as possible, the Distributional National Accounts Methodology, see Alvaredo *et al.*, 2016).

To do so, we combine in a systematic manner household survey, fiscal and national accounts data. Such an exercise is fraught with methodological and conceptual difficulties given the lack of consistent historical income inequality data in India. Indeed, the tax data available only covers the very top of the distribution of Indian earners (around 7% of total population in fiscal year 2014–15). In addition, the National Sample Survey Organization (NSSO) household surveys measure consumption rather than income. We repeatedly stress that there are strong limitations to available data sources, and that more democratic transparency on income and wealth statistics is highly needed in India. That said, we find that our key results are robust to a large set of alternative assumptions made to address data gaps. The present paper should be viewed as an exercise in transparency: we propose a method to combine the different available sources (in particular national accounts, tax and survey data) in the most possible transparent way, and we very much hope that new data sources will become available in the future so that more refined estimates can be constructed. All our computer codes are available on-line so that everybody can use them and contribute to improve the methods.

The rest of this paper is organized as follows. Section 1 discusses the Indian income inequality data gap of the past two decades, Section 2 describes our data sources and methodology, Section 3 presents our key findings, Section 4 briefly discusses their policy relevance and Section 5 concludes.

<sup>1</sup>Appendix 1 presents real per adult annual growth rates using GDP from United Nations National Accounts Database (used in this paper) and the World Bank Database.

## 2. ENTERING THE DIGITAL AGE WITHOUT INEQUALITY DATA

### 2.1. *Economic Policy Shifts Since the 1980s*

Over the past 30 years, the Indian economy went through profound evolutions. In the late seventies, India was recognized as a highly regulated economy with socialist planning. From the 1980s onwards, a large set of liberalization and deregulation reforms were implemented. In this context, it is unfortunate that Indian authorities stopped in 2000 publishing income tax tabulations, which represent a key source of data to track consistently the evolution of top incomes.

Under Prime Minister Jawaharlal Nehru (in power from 1947 to 1964), India was a statist, centrally directed and regulated economy. Transport, agriculture and construction sectors were owned and administered by the Central Government, commodity prices were regulated and the country had important trade barriers. Nehru's followers, including Indira Gandhi (1966–77 and 1980–84), prolonged these policies and implemented a highly progressive tax system. In the early 1970s, the top marginal income tax rate reached record high levels (up to 97.5 percent).

From the mid 1980s onwards, liberalization and trade openness became recurrent themes among Indian policymakers. The Seventh Plan (1985–90), led by Rajiv Gandhi (1984–89), promoted the relaxation of market regulation, with increased external borrowing and increased imports. The tax system was also gradually transformed, with top marginal income tax rates falling to 50 percent in the mid-1980s. In the late 1980s, when India faced a balance of payment crisis, it called for International Monetary Fund assistance. Financial support was conditioned to structural reforms which pushed forward the deregulation and liberalization agenda.

What came to be known as the first set of economic reforms (1991–2000) placed the promotion of the private sector at the heart of economic policies, via denationalizations, disinvestment of the public sector, deregulation (dereservation and delicensing of public companies and industries).<sup>2</sup> These reforms were implemented both by the Congress government of P.V.N. Rao (1991–96) and its successors, including the conservative Janata Party government of A. Vajpayee (1998–2004). The reforms were prolonged after 2000, under the 10th and subsequent five-year plans. These plans ended government fixation of petrol, sugar or fertilizer prices and led to further privatizations, in the agricultural sector in particular.

The impact of these reforms in terms of growth has been praised by public authorities. Real per adult national income growth, which has more sense from the point of view of individual incomes than commonly used GDP,<sup>3</sup> significantly increased after the reforms. It was 0.7 percent in the 1970s, 2.5 percent in the 1980s, 2.0 percent in the 1990s and 4.7 percent since 2000 (Figure 1). However, little is known on the distributional characteristics of post-2000 growth.

<sup>2</sup>Economic policies also sought to rationalize the public sector, its branches now had to pursue the objectives of profitability and efficiency. The opening of imports, exchange rate floating regime and banking, capital market opening were also implemented.

<sup>3</sup>Net national income is equal to GDP minus depreciation of fixed capital plus net foreign incomes.

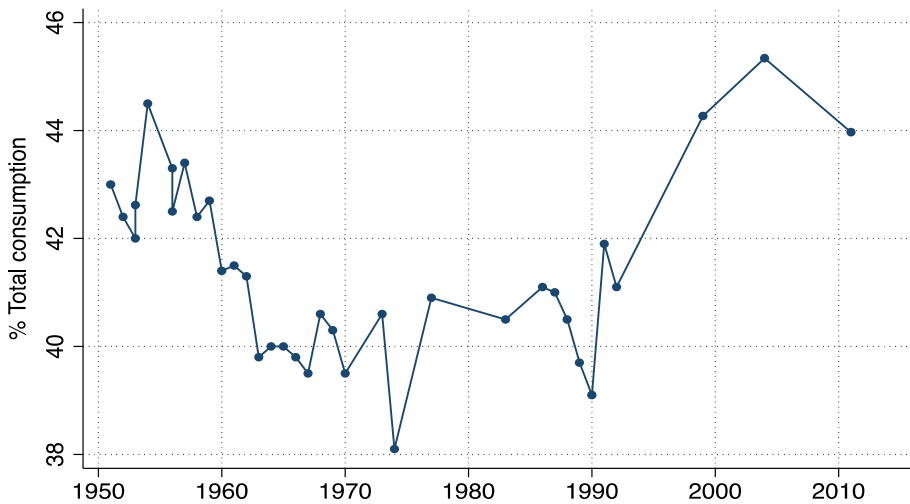


Figure 2. Top 20% Consumption Share from NSSO Surveys

Source: Authors' computations using data from United Nations WIDER Income Inequality Database and World Bank India Database (based upon NSSO surveys). [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

## 2.2. The Income Inequality Data Gap

Public debate over liberalization policies largely focused on their macroeconomic impacts (Kotwal *et al.*, 2011) and on the impacts on poverty, with a substantial reduction in poverty rates<sup>4</sup> (Deaton and Dreze, 2002; Deaton and Kozel, 2005; World Bank, 2017). How the Indian economy fared in terms of inequality has been arguably less discussed. This can partly be explained by a lack of consistent data on the distribution of incomes or wealth for the recent period.

Some evidence suggesting a rise in income inequality in India after the turn of the century can however be found in NSSO surveys and in openly-available data sources. Figure 2 presents the share of total consumption attributable to the top 20 percent of consumers, available online from the World Bank and United Nations WIDER World Income Inequality Database (UN-WIDER WIID). The data shows a decrease in top quintile consumption share from the fifties to the seventies from around 43 percent to 40 percent and an increase thereafter (in line with Banerjee and Piketty findings) to close to 44 percent. There are important irregularities with the data, but the overall “U-shape” trend seems relatively consistent.<sup>5</sup>

The shortcomings of household survey data in monitoring the evolution of inequality are well known; because of underreporting and undersampling issues, surveys fail to properly capture inequality dynamics at the top of the distribution (Atkinson and Piketty, 2007, 2010). What is more, NSSO surveys only focus

<sup>4</sup>The share of Indians under the \$1.9 poverty line went from 45.9% in 1993 to 21.2% in 2011 PovcalNet, World Bank (2017).

<sup>5</sup>As discussed below, income surveys sources are available for 2005 and later years; in particular data from the National Council for Applied Economic Research (NCAER) and from the Inter University Consortium for Applied Political and Social Sciences Research (ICPSR). These data sources however do not enable comparison before and after 2000.

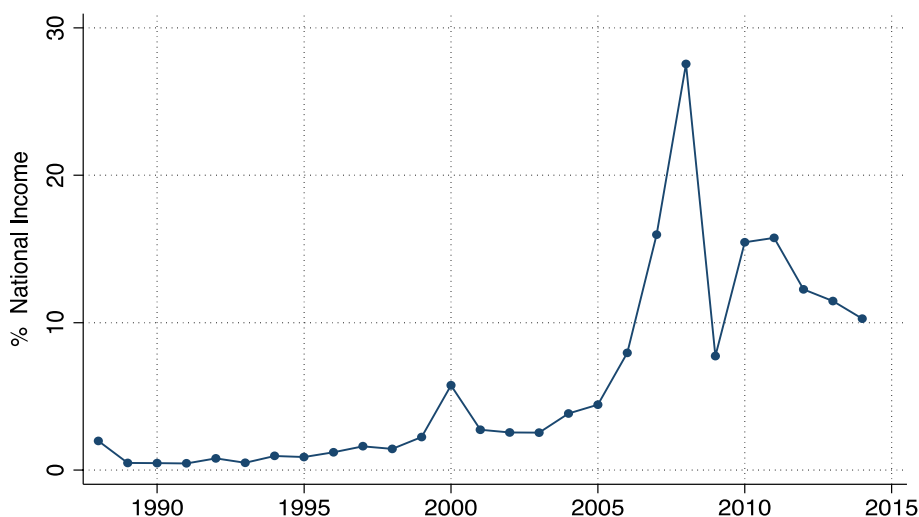


Figure 3. Wealth of Richest Indians in Forbes' Rich List, 1988–2015

*Source:* Authors' computations based upon Forbes billionaire rankings and WID.world national income data. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

on consumption rather than income and the distributional dynamics of these two concepts can differ notably. In addition, the relatively limited magnitude of the changes observed in NSSO data calls for care in the interpretation of such results. Consumption data available through surveys constitutes part of the evidence, but are not sufficient to inform debates on Indian inequality.

Other data sources, such as Forbes' Indian Rich lists, suggest an important increase in the wealth of the richest Indians after 2000 (see Figure 3). The wealth of the richest Indians reported in Forbes' India Rich List, amounted to less than 2 percent of National income in the 1990s, but increased substantially throughout the 2000s, reaching 10 percent in 2015 and with a peak of 27 percent before the 2008–9 financial crisis. Such data suggests a rise in wealth inequality levels throughout the post-2000 period, but does not enable a consistent analysis of income inequality over the long run. This is confirmed by simple simulations using a fixed normalized wealth distribution and taking into account rising average nominal wealth over the period (unfortunately Indian wealth data is very limited so it is difficult to go further).

The recent release of income tax tabulations by the Indian Income Tax Department for the post 2011 period does, however, allow for a more consistent analysis of the dynamics of income in India since the turn of the century.

### 3. DATA SOURCES AND METHODOLOGY

We present the data used to produce series on the evolution of income for the entire distribution from 1951–52 to 2014–15 (period covered by both household surveys and tax data, as well as national accounts) and for the evolution of incomes of the top 1 percent share and above from 1922–23 to 2014–15 (period covered by tax data and national accounts only, with no survey data prior to 1951). For pre-1922

Indian income inequality series, see Alvaredo, Bergeron, and Cassan (2017). The authors mobilize income tax data to study the concentration of income in British India, from 1885 to 1946.<sup>6</sup>

### 3.1. *Description of the Different Data Sources*

#### Tax Data

Standard measures of inequality in India mostly rely on self-reported survey data, which frequently underestimate top income or consumption levels and trends (see Atkinson and Piketty, 2007, 2010; Alvaredo *et al.*, 2016). Tax data, despite its limitations (low population coverage in emerging countries and potential tax fraud and evasion), provide a much more realistic picture of *top* income levels and trends, and can thus be used to complement distributional information obtained from household surveys in a country like India.

The Indian Income Tax Department released tax tabulations for the fiscal years 1922–23 to 1998–99, and interrupted the publication in 2000. After several public calls for more democratic transparency over Indian inequality data, the ITA released tax tabulation for fiscal years 2011–12 to 2014–15. All these tabulations report the number of taxpayers and the gross and returned income for a large number of income brackets.<sup>7</sup> Gross income corresponds to pre-tax income before certain deductions are applied to compute returned income.<sup>8</sup> Tax units are defined as individuals or Hindu Undivided Families (HUF, family clusters allowed to file their income jointly). The number of HUF represented roughly 20 percent of tax returns in the interwar period, 5 percent in 1990 and less than 2.5 percent in 2011.<sup>9</sup>

The exact reason why the Income Tax Department stopped publishing data in 1999–2000 remains unknown. One potential explanation for this is the change in

<sup>6</sup>See for instance <http://www.bbc.com/news/world-asia-india-36186116>

<sup>7</sup>According to the Income Tax Department, a number of tax payers paid their taxes but did not file returns in fiscal years 2011–12 to 2014–15. These represent an additional 25 percent taxpayers. In order to take into account these “non-filers” taxpayers, we tested alternative assumptions: i) non-filers are scattered across all brackets, in the same way as filers, ii) non-filers fall in the lowest taxable bracket, iii) non-filers fall in the four lowest income brackets. We find that these alternative assumptions have very limited impact on our final results. Minor corrections were done to raw tax data and mainly pertain to the clubbing of brackets in some years as the average income was incompatible with the bracket they were categorized. In such rare cases, we club erroneous brackets in the lower bracket. Year 1997 was removed altogether, as reported data for top income stands out as an extreme outlier.

<sup>8</sup>Deductions are defined at chapter VI of the Income Tax Act. They include premiums of annuity plans, equity fund investments, medical or health insurance, certain forms of donations, etc. Focusing on gross income is more accurate in terms of pre-tax income and is also less impacted by changes in the definitions of deductions. Income losses (such as business income losses) have to be adjusted while computing Gross Total Income as per Income Tax law. Note that imputed rent for owner occupied dwellings were included in Income tax computations before 1986 and removed afterwards. More precisely, post 1986 tax data excludes imputed rent for first residence, but not for secondary residences.

<sup>9</sup>One should note that the Indian income tax data is entirely based upon individual income. This corresponds to equal-split income (i.e. income shared among spouses) only if we assume that tax-payers are either single or married to other tax-payers falling in the same bracket, which strictly speaking cannot be true. This implies that our estimates tend to over-estimate inequality as compared to the equal-split benchmark. The equal-split benchmark however tends to under-estimate inequality as compared to an individualistic benchmark (a benchmark in which one assumes no sharing of income among spouses). If and when we access to micro-level Indian tax data, we will be able to refine this analysis and compute separate equal-split and individualistic series.



the sampling method employed in the late 1990s, with a resulting loss in the precision of estimates. Indeed, official tax tabulations were based on the entire population until the early 1990s—or based on stratified samples with sampling rates close to 100 percent for top incomes as is the case in most OECD countries, but seem to be based on uniform samples of all tax returns after this period and up to 2000 (Banerjee and Piketty, 2005). The latter method led to less precise results. Another potential explanation for the halt in tax reporting could just be the lack of interest in income statistics and inequality.

Interestingly enough, the number of income tax payers in India has increased substantially over the past decades. Less than 0.5 percent of the population filed tax returns up to the 1950s, between 0.2 and 1 percent over the period between 1960 and 1990, before a substantial increase thereafter; from 1 percent to close to 3 percent in the late 1990s and 7 percent in the latest period (Figure 4). This increase over twenty years is impressive. Yet, comparatively, the current figure is similar to the levels observed in France and in the USA in the late 1910s, and much lower than the levels observed in the interwar period (about 10–15 percent) and in the decades following World War 2 (50 percent or more) in these two countries (Piketty, 2001; Piketty and Saez, 2003). With revenues from income tax equivalent to approximately 2 percent of GDP, India receives more revenue than China (1 percent), but significantly less than other emerging countries such as Brazil and Russia (4 percent), and South Africa and the OECD countries (9 percent) (OECD, 2017).

#### NSSO Consumption Data

The NSSO, led by the Ministry of Statistics and Program Implementation started an all-India consumer household expenditure survey (AIHS) after its independence in 1947. The first round of the AIHS was carried out in 1951 and

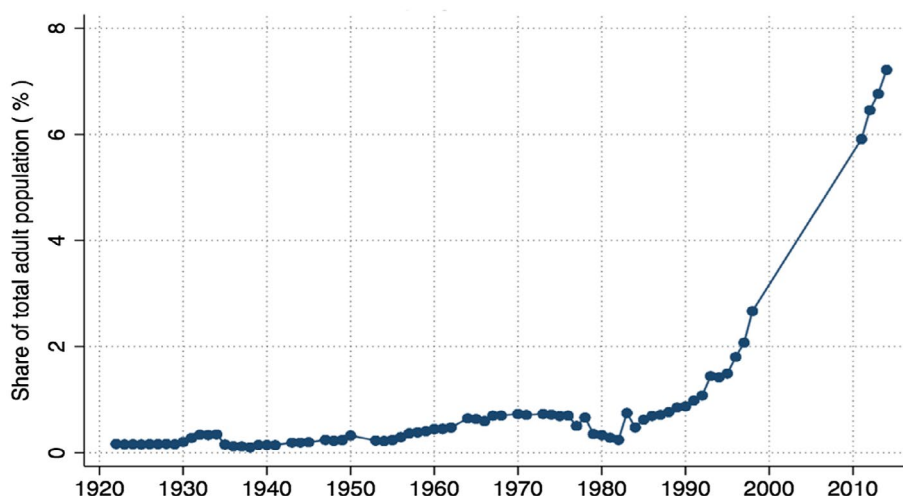


Figure 4. Proportion of Income-Tax Taxpayers in India, 1922–2015

Source: Authors' computations based on Indian Tax Administration statistics, United Nations Population Database and Banerjee and Piketty (2005). Income tax payers refer to individuals and Hindu Undivided Families. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



surveys were then conducted on an annual basis. The size of rounds varies since the quinquennial AIHS has a larger sampling of about 120 000 households and five times less for smaller other rounds. The reach of the quinquennial survey is extensive in terms of consumption items (ranging from daily used food, clothing to durable goods and services such as construction, education and healthcare). NSSO surveys however do not measure individual or household incomes,<sup>10</sup> in part because agricultural and business incomes are judged to be volatile and assumed to be much less reliably measured than consumption.

Since the first survey rounds, NSSO produced 30 days reference period estimates. This period is known as the Universal Reference Period. Post-1990, concerns were raised about the sensitivity of the reference period on the estimates and NSSO started publishing alternative reference periods (7 days and 365 days). As Deaton and Kozel (2005) note, shorter recall periods tend to lead to higher consumption estimates. However, experiments carried out with different reference periods by the NSSO working group concerned concluded that there is no clear superiority of a period over another. We thus use the Universal Reference Period. This choice is also motivated by the fact that the 30 days period is the only one that is consistent throughout the entire period of analysis (1951–2015).

For recent years (1983 to 2010) we use quinquennial rounds 38 (1983), 43 (1987–88), 50 (1993–94), 55 (1999–2000), 61 (2004–5), 66 (2009–10). Micro data at the household level was obtained from the NSSO. For earlier rounds (rounds 3 to 32), for which we could not access micro data files, we use the Poverty and Growth in India Database of the World Bank (Ozler *et al.*, 1996) which provides rural and urban per capita consumption tabulations for a dozen quantile groups for years 1951 to 1978. All rounds and corresponding years used are summarized in Appendix 3, along with the summary statistics of each round. We describe in Section 3.2 the procedure used to infer the full distribution of income from these surveys and how we interpolate missing years.

### National Accounts Data

From 1950 to the present day, we use GDP data from WID.world, based on National Accounts Statistics (NAS) from 1971 to 2013, on World Bank (after 2013) and on Maddison (2012) from 1950 to 1970.<sup>11</sup> WID.world then performs its own computations to infer Net Foreign Income and Consumption of Fixed Capital (Blanchet and Chancel, 2016). Before 1950, we use historical National Income growth rates from Sivasubramonian (2001).

A well known puzzle in Indian statistics (Deaton and Kozel, 2005; CSO, 2008) pertains to the difference in survey consumption growth rates and national accounts growth rates, particularly during the recent period. Figure 5 shows the total growth rate of Net National Income and Household Final Consumption Expenditure

<sup>10</sup>The Employment Unemployment Surveys report wages for the working-age population, but other sources of income are not covered.

<sup>11</sup>In the 1990s we observe noticeable differences between real GDP growth estimates obtained from UN SNA and those reported by the World Bank (see Appendix 1).

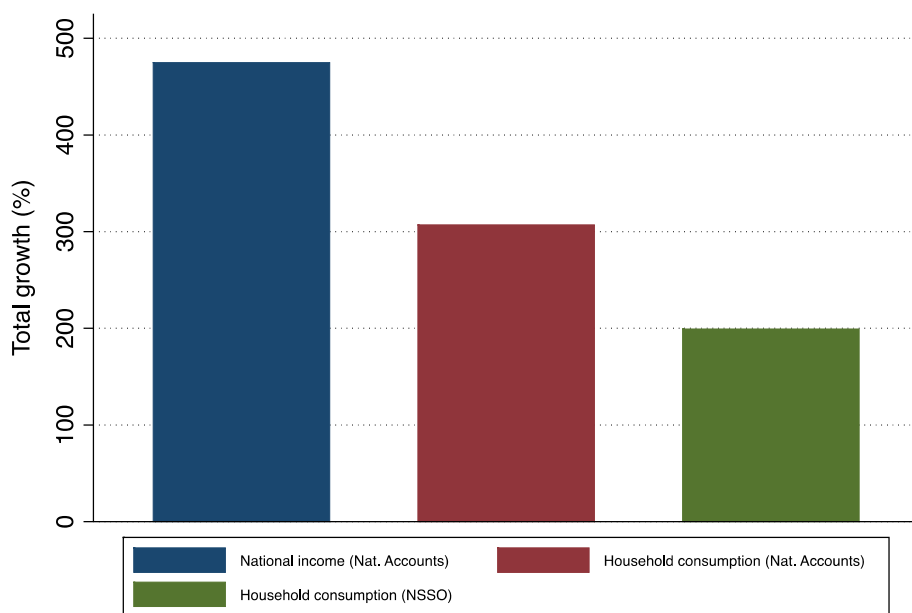


Figure 5. Income and Consumption Growth Rates in India, 1988–2011

Source: Authors' computations using National Accounts and NSSO data. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

from NAS and personal consumption from NSSO, from 1983 to 2011. According to NAS, national income grew at 475 percent and household consumption grew at slightly more than 300 percent, while NSSO data indicates that household consumption grew at 200 percent.

Several reasons have been put forward to explain this gap, including (i) population coverage (it is different between NSSO and NAS, since Non Profit Institutions Serving Households and homeless individuals are not covered by NSSO surveys); (ii) valuation and integration of certain types of services in survey questionnaires (it was argued that the treatment of cooked meals served by employers to employees leads to underestimation of the total value of services consumed by households in the NSSO surveys (CSO, 2008) while other services such as financial intermediation that are particularly important among top earners, are not included in survey estimates (Sundaram and Tendulkar, 2005); (iii) imputed rents (while the NAS incorporates imputed rents, NSSO surveys do not<sup>12</sup>); (iv) consistency of National Accounts estimates (Kulshreshtha and Kar, 2005); (v) under-reporting and under-sampling of top incomes in survey data (Banerjee and Piketty, 2005). We should stress from the outset that we do not pretend to solve this complex issue. The divergence probably involves several, if not all of the factors above cited. What we seek here is to better estimate the fraction of the difference that can be explained by the absence of top earners in survey

<sup>12</sup>When correcting for imputed rents the Central Statistical Organization (2008) finds a large and growing share of total consumption remains unexplained.

data. We do not think that this factor alone can explain the entire gap, as it has been suggested (Lakner and Milanovic, 2015).

### IHDS Income and Consumption Survey

The Inter University Consortium for Applied Political and Social Sciences Research (ICPSR), based at the University of Michigan, provides access to the India Human Development Survey (IHDS), conducted in 2004–05 and 2011–12 among more than 40 000 households from rural and urban areas. The survey provides information at the household level on both income and consumption. Consumption related questions were designed so as to match the NSSO questionnaire, using similar item categories and similar referencing periods. The definition of income in the IHDS survey includes all sources of income: labor income (wages and pensions), capital income (rents, interests, dividends, capital gains) as well as mixed (or business) incomes. Government benefits, reported in the survey, are excluded from the analysis for consistency with tax tabulations; our focus is pre-tax income.

The IHDS is one of the very few surveys estimating both consumption and income in India. This is particularly useful as it enables a tentative reconstruction of NSSO unobserved income levels, using IHDS information. We describe this methodology in the next section. IHDS micro data is also openly available via the ICPSR website, which makes it particularly convenient.<sup>13</sup>

### UN Statistics Population Data

We define the theoretical population of tax payers as the total number of adult individuals in India. We use adult population data from UN Population Prospects (United Nations, 2015) from 1950 to today. UN Population prospects provide 5-year age range annual population tables, based on national census and their own estimation procedures. The adult population is defined as the number of individuals over age 20. Before 1950, we use total population estimates from Sivasubramonian (2001) and reconstruct the adult population using total population growth rates given by the same author.

### 3.2. Methodology

#### Estimation of Top Fiscal Incomes

Following Banerjee and Piketty (2005), we first reconstruct top income thresholds and levels, using generalized Pareto interpolation techniques. The main methodological difference with Banerjee and Piketty lies in the use of generalized Pareto interpolation techniques (Blanchet *et al.*, 2017) rather than standard Pareto distributions. Generalized Pareto interpolation<sup>14</sup> allows for the recovery of the distribution based on tax tabulations without the need for parametric approximations. This method has demonstrated its ability to produce very precise results and also has the advantage of generating smooth estimates of the

<sup>13</sup>We were not able to access the micro files of the National Council for Applied Economic Research's National Survey on Household Income and Expenditure, carried out in 2004-5 and 2011.

<sup>14</sup>Available online at [www.wid.world/gpinter](http://www.wid.world/gpinter).

distribution, i.e. generating a differentiable quantile function and a continuous density, while other methods introduce kinks around the thresholds used as inputs for the tabulation.

The generalized Pareto interpolation procedure generates 127 generalized percentiles, namely  $p_{0p1}$ ,  $p_{1p2}$ , ...,  $p_{99p100}$ , corresponding to 100 fractiles of the distribution. The top fractile is split into 10 deciles ( $p_{99.0}$   $p_{99.1}$ ,  $p_{99.1}$   $p_{99.2}$ , ...,  $p_{99.9p100}$ ), its top decile itself split in ten deciles ( $p_{99.90}$   $p_{99.91}$ ,  $p_{99.91}$   $p_{99.92}$ , ...,  $p_{99.99}$   $p_{100}$ ), the tenth decile again split in ten deciles ( $p_{99.990}$   $p_{99.991}$ ,  $p_{99.991}$   $p_{99.992}$ , ...,  $p_{99.999}$   $p_{100}$ ). The top generalized percentile thus corresponds to the top 0.001 percent of the population. As shown in Figure 4, tax data in India is only reliable above the  $p_{94}$  threshold for the recent period and above the  $p_{99.9}$  threshold when we go backwards in time.

### Estimation of Bottom Survey Incomes

One of the main difficulties of our exercise is related to the fact that NSSO does not include questions on individual and/or household income. Our strategy consists of using observed income-consumption profiles in IHDS data to reconstruct income profiles from NSSO consumption data. We first estimate income and consumption levels for each generalized percentile of the distribution of income and consumption given by IHDS data. For each survey and each percentile of the distribution, we construct observed income-consumption ratios  $\alpha_{1p} = y_p/c_p$ , with  $y_p$  and  $c_p$  respectively representing mean income and consumption within quantile  $p$ . We call this strategy A1. To obtain a theoretical income-consumption profile over percentiles, we take average of years 2004–5 and 2011–12. In practice, the two profiles differ only marginally. We then construct two alternative ratios,  $\alpha_{2p}$  and  $\alpha_{0p}$ , referred to as strategies A2 and A0 respectively. In strategy A2, we assume that  $\alpha_{2p} = 1$  for  $\alpha_{1p} \leq 1$  and  $\alpha_{2p} = \alpha_{1p}$  otherwise. This second strategy is equal to assuming no negative savings rates among the poor. In strategy A0, we define  $\alpha_{0p} = (\alpha_{1p} + \alpha_{2p})/2$  for  $\alpha_{1p} \leq 1$ . This strategy assumes that there can be negative savings rates, remittances or household transfers, but that the true  $\alpha_p$  value lies between strategy A1 and strategy A2. Income consumption ratios for the different strategies are presented in Appendix 4. We find that these different strategies have no effect on the broad trends we observe and have a limited impact on top share estimates, as we show in Appendix 14–15. These strategies have a relatively marked impact on bottom 50 percent income shares, as expected (see Appendix Figures 14–15), but not on the trends observed, particularly after the 1980s.

The choice of these different strategies indeed impacts on the estimated share of total savings in the economy. In strategy A1 total savings are close to 0, which seems too low compared to the current rate of savings in India (about 30 percent). This figure is close to 5 percent in strategy A0 and approximately 10 percent in strategy A2. These values are more or less constant throughout the entire period covered whereas in National accounts they move from about 10 percent in the 1960s to 30 percent today. However, using strategy A0 and factoring in top incomes in the analysis allows us to find an aggregate savings rate of the same order of magnitude as those observed today (see Appendix 5).

## Interpolating Survey and Tax Data for Missing Years

Our objective is to produce yearly estimate for the full distribution from 1951 to 2015. Given that survey or tax data is not available for all years, it is necessary to interpolate tax and/or survey data for a certain number of years. In order to do so, we interpolate missing years using a constant growth rate between known intervals  $t$  and  $t + N$ .<sup>15</sup>

As described in Section 3.1, two surveys can be used for the estimation of survey income for the years 2004–5 and 2010–11, NSSO and IHDS. However, the trends observed in the surveys are somehow divergent. The ratio of reconstructed NSSO total income to total personal income from national accounts decreases, while the ratio of IHDS total income to total personal income from national accounts is stable. The choice of one or the other source of data has implications on our final inequality statistics: using IHDS income group averages for the estimation of the bottom of the distribution (strategy B1) yields a lower rise in top income shares than when using the NSSO survey (strategy B2). In fact, using NSSO totals mechanically accentuates the rise in top shares over the period and the strategy B1 is therefore used as our benchmark, as it represents the conservative approach. That said, we cannot rule out strategy B2, if we believe NSSO surveys are consistent throughout the entire period covered. We provide results for strategy B2 in the data appendix.

Between 2000 and 2011, we do not observe any tax statistics, but we do observe survey data in 2004–5 and in 2010–11. Survey data is not satisfactory to track the dynamics of top incomes, but it is better than no data at all. We thus estimate the growth rates of each percentile between 1999 and 2005 on the basis of their evolution observed in the survey distribution. The resulting estimates show the top 10 percent share evolving in the same direction between 2005 and 2011 in our final results as in the survey. We see this strategy as the best we can have with the available data at hand for this specific sub-period.

## Combination of Tax and Survey Data

Several strategies can be used to correct for missing top incomes in survey data. These include the modification of the weights assigned to top earners in household surveys, the addition of extra observations of top earners or the multiplication of income levels at the top (Burkhauser *et al*, 2016), and each has its own strengths and weaknesses. We think that an acceptable method should be consistent, in producing distributions with plausible statistics, in particular, the shape of inverted Pareto beta coefficients curves should be relatively smooth. The method followed should also be transparent, in so-much as it should provide a statistical outcome that could be anticipated from an economic perspective; survey inequality should in principle increase when we factor in top fiscal incomes. Furthermore, a simple strategy would also be better than a complex one.

<sup>15</sup>In practice, for each average income at percentile  $p$  of the survey (or tax) distribution, we define  $y_{pt+1} = y_{pt} \times g$  where  $g = (y_{pt+N}/y_{pt})^{1/N}$ , with  $g$  the growth rate,  $y_{pt+1}$  the average income at percentile  $p$  and year  $t + 1$ .

Our preferred strategy is to assume that surveys are reliable from the bottom of the distribution up to a certain percentile and that tax data is reliable after another (in line with Piketty *et al.*, 2019). In practice, this amounts to multiplying income of the top percentiles in the survey by a certain factor, given by tax data. More precisely: we suppose that survey data is reliable from  $p_0$  to  $p_1$  - this means that between  $p_0$  and  $p_1$ , averages and thresholds are given by the distribution of interpolated (estimated) survey income. In our benchmark scenario, which we refer to as strategy C1,  $p_1 = p_{90}$ . We also test alternative ranges: (i)  $p_1 = p_{95}$ , which we refer to as strategy C2 and (ii)  $p_1 = p_{80}$ , referred to as strategy C3. As shown in Section 3.5, these different strategies have no impact on the recent and long-term income trends observed in India and have only a moderate impact on income concentration levels.

We then suppose that tax data is reliable from a certain percentile,  $p_2$ , up to the top of the distribution.  $p_2$  is given by the population share lying under the first taxable bracket observed in the tax data. This value varies from  $p_2 = 99.9$  in the 1950s to  $p_2 = 93$  in the 2010s. Therefore, our strategy implies that averages and thresholds for all percentiles above  $p_2$  are given by the distribution interpolated from observed tax data. Appendix 5 gives the precise value of  $p_2$  for each year.

Between  $p_1$  and  $p_2$ , we test several strategies for the progression of income levels and thresholds at a given point of time. We define a convex junction profile (strategy D1), a linear profile (strategy D2) and a concave profile (strategy D3). We adopt D1 (convex profile) as our benchmark strategy as it corresponds to the profile observed for recent years, for which we have more observed fiscal data at the top; more than 6% of the population against 0.1 percent for the earlier period (see Appendix 6). We find that these different strategies have negligible impacts on top share results. In fact, the bulk of the correction we apply to survey incomes occurs above  $p_2$ , not between  $p_1$  and  $p_2$ .

#### From Total Fiscal Income to National Income

Total fiscal income is the total personal income that would be reported by individuals or tax units, if all of them reported their revenues to the tax administration. In the case of India, we do not observe this value because of the limited tax base. One way to recover it, following Atkinson (2007), is to start from the sum of primary incomes obtained by households reported in national accounts and operate a series of deductions and additions towards a definition closer to taxable income. This is the approach followed by Banerjee and Piketty (2005) and appears appropriate given that their focus was restricted to top incomes only. By construction, total fiscal income evolves at the same rate as pre-tax national income under this approach.

The other approach consists of reconstructing total fiscal income via the combination of top fiscal incomes and observed (or estimated) survey income, as we detailed in the previous section. This is equivalent to assuming that tax data give true fiscal incomes for individuals over  $p_2$  and that estimated survey data gives the true fiscal incomes for individuals below  $p_1$ . In this approach, reconstructed fiscal income and total national income can evolve at a different pace. Over the years, we observe a growing gap between reconstructed total income from surveys and total national income (see Appendix 7). This divergence is the repercussion of the gap between household consumption surveys and national accounts discussed in



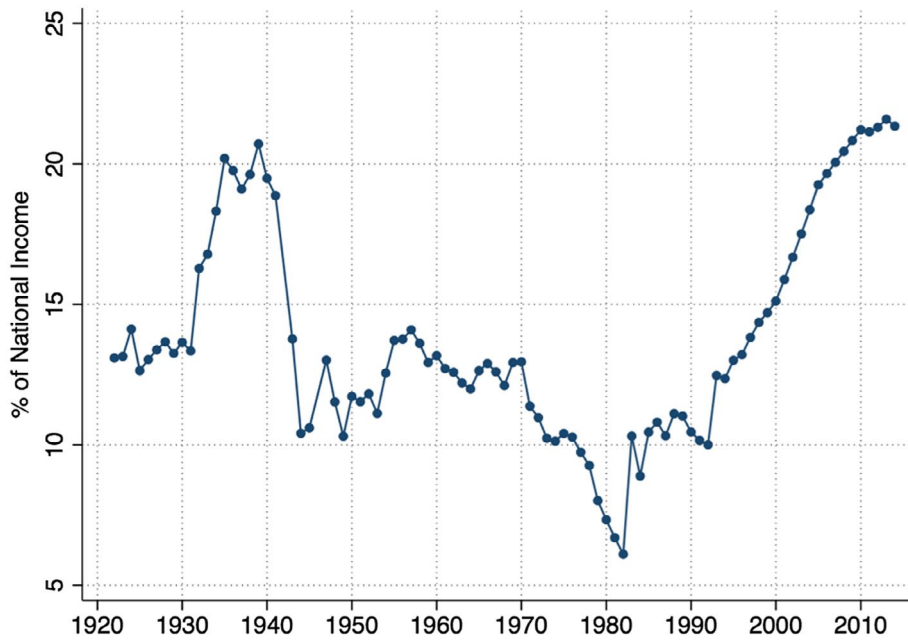


Figure 6. Top 1% National Income Share in India, 1922-2015

Source: Authors' estimates combining survey, fiscal and national accounts data.

Note: Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1).  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Section 3.1. We show in Figure 6 that we can account for a non-negligible share of this gap after the combination of survey and tax data, but that a large part of the difference remains unexplained.

In order to produce income estimates comparable to other countries, we chose to rescale our fiscal income estimates to match total pre-tax national income from national accounts. In practice, we preserve the distribution obtained from the combination of tax and survey data and simply rescale average and threshold levels of all percentile groups by a yearly factor so that we match total national income.

In further work, we intend to distribute retained earnings to the top of the distribution following the DINA guidelines (Alvaredo *et al.*, 2016). This would most likely increase the level of inequality in the recent period, since the growth of retained earnings is likely to be concentrated among top earners. The amount by which our results would vary presumably remains limited though.

#### Definition of A Benchmark Scenario

The combination of our different strategies defines 54 scenarios (3 A scenarios x 2 B scenarios x 3 C scenarios x 3 D scenarios). We stress that most of the combinations of scenarios among these 54 possibilities can be *a priori* justified, and as such, we provide results for all corresponding series in our data appendix. We see our benchmark scenario (A0B1C1D1) as being at the same time plausible



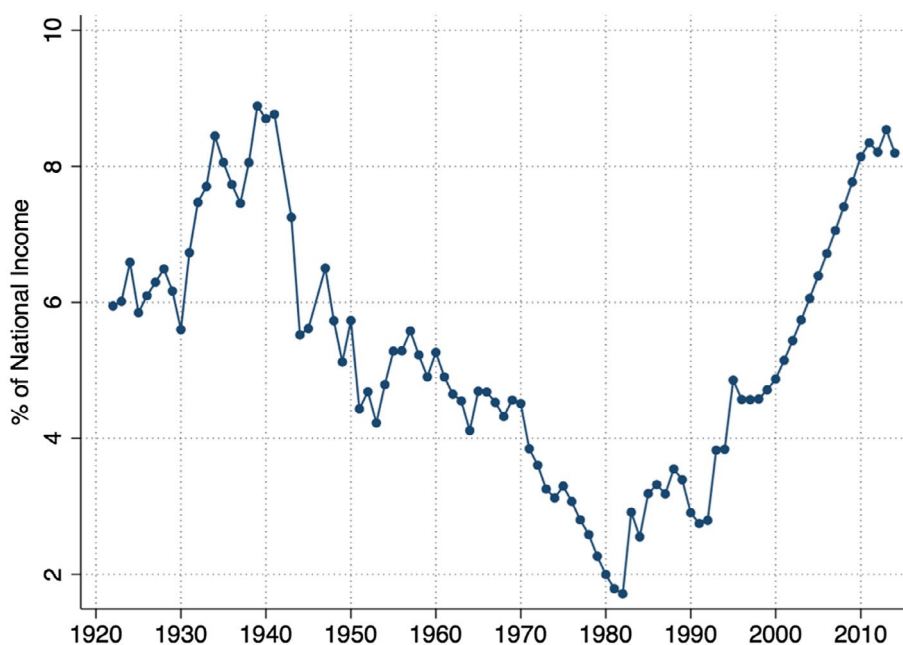


Figure 7. Top 0.1% National Income Share in India, 1922–2015

Source: Authors' estimates combining survey, fiscal and national accounts data.

Note: Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1).  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

and conservative compared to most of the scenarios tested, as top income shares increase at a slower rate over the recent decades than in most scenarios. Robustness tests are presented in Section 4.5.

## 4. RESULTS

### 4.1. Sharp Rise in Top Income Shares since the mid-1980s

Our results exhibit a strong rise in top income shares since the mid-1980s. In our benchmark estimation scenario, the share of national income attributable to the top 1 percent reached 21.3 percent of national income in 2014–15, up from 6.2 percent in 1982–83 (see Figure 7). The top 1 percent share of national income was at 13 percent of national income in 1922–23 and increased to 20.7 percent in 1939–40, at the dawn of World War II. It then dramatically decreased to 10.3 percent in 1949–50 and further decreased from the late 1960s to the early 1980s.

As expected, the top 0.1 percent income share dynamics exhibit a similar pattern in our benchmark scenario (see Figure 8). Top 0.1 percent earners captured 8.2 percent of total income in 2014–15. This is only slightly below its pre-independence peak of 1939–40 (8.9 percent). The top 0.1 percent then saw a strong drop during World War II (down to 5.5 percent in 1944–45), followed by a continued reduction up to 1982–83 (when it reached 1.7 percent). From 1983–84 onwards, the share of national income accruing to the top 0.1 percent rose almost continuously.

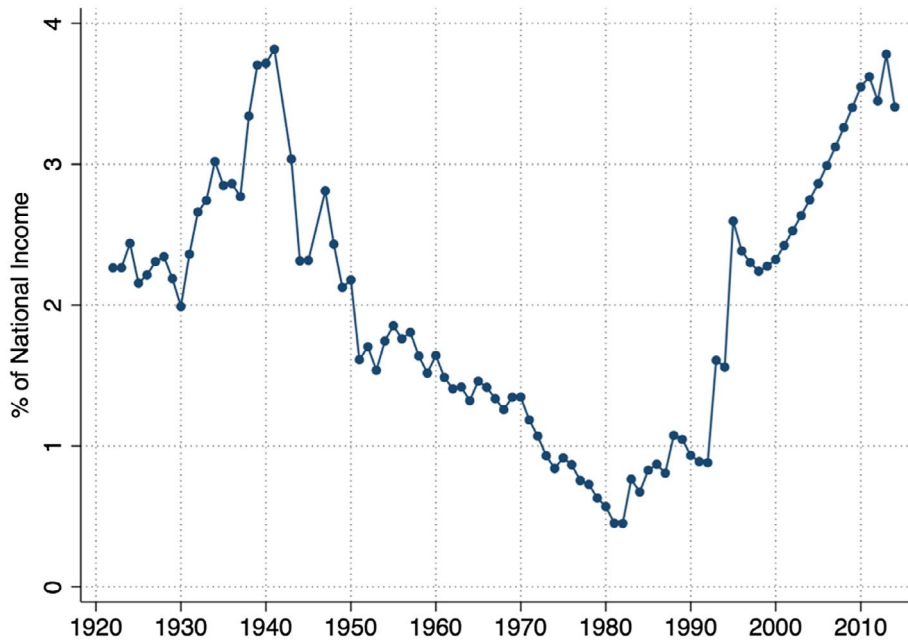


Figure 8. Top 0.01% National Income Share in India, 1922-2015

*Source:* Authors' estimates combining survey, fiscal and national accounts data.

*Note:* Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1).  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

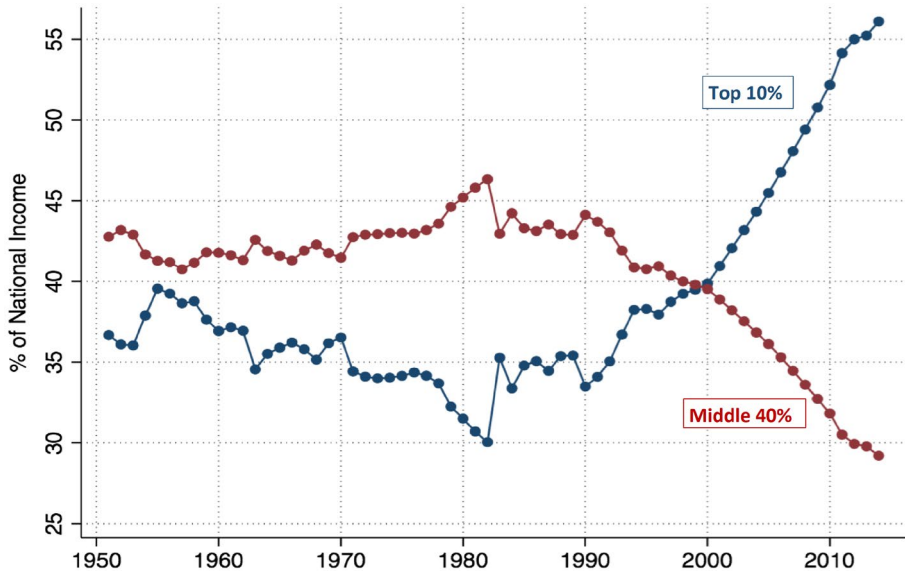


Figure 9. Top 10% vs. Middle 40% National Income Shares in India, 1951-2015

*Source:* Authors' estimates combining survey, fiscal and national accounts data.

*Note:* Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1).  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

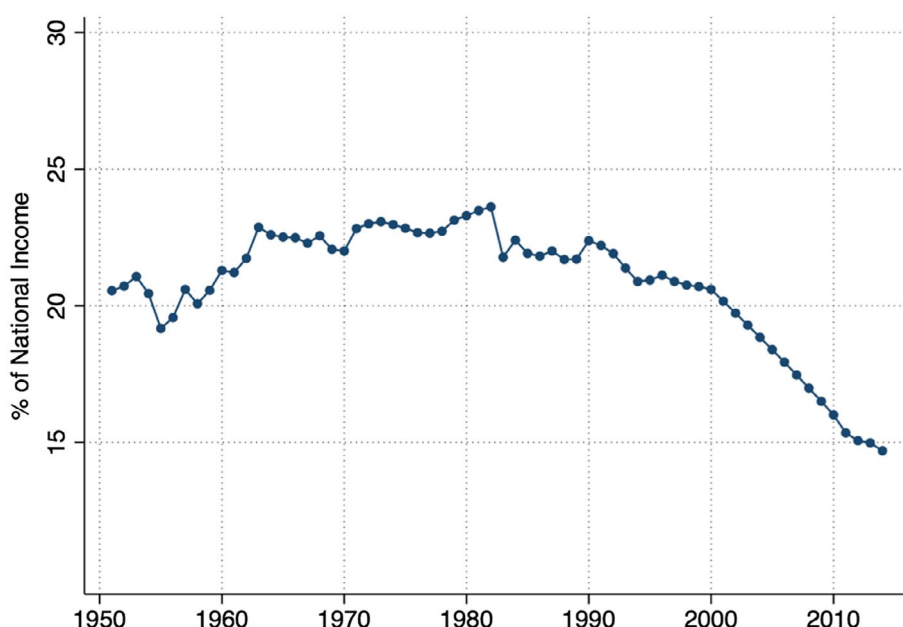


Figure 10. Bottom 50% National Income Share in India, 1951-2015

Source: Authors' estimates combining survey, fiscal and national accounts data.

Note: Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1).  
[Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Looking at the 0.01 percent earners (Figure 9), we also observe a strong increase in their share of national income since the mid 1980s, reaching 3.4 percent in 2014–15, up from 0.4 percent in 1982–83. In 1941–42, the top 0.01 percent earned 3.8 percent of total income.

#### 4.2. Fall in Middle 40 percent and Bottom 50 percent Shares

We now turn to post-1951 results, which we have for the entire distribution of income. Figure 10 shows the mirror evolution of top 10 percent share in total income and middle 40 percent share (i.e. individuals above the bottom 50 percent earners and below the top 10 percent). In the mid-fifties, the top 10 percent and the middle 40 percent held about 40 percent of total income each, the share of the middle 40 percent progressively increased from the mid-fifties to 1982–83, reaching 46 percent of total income. It then decreased afterwards. At the turn of the Millennium, the top 10 percent and the middle 40 percent groups captured exactly the same amount, 40 percent. However, by 2014–15, the middle 40 percent share had fallen to a historically low level of 29.2 percent.

The income dynamics of the poorest half of the income distribution exhibit a similar pattern to that of the middle 40 percent (Figure 11). Bottom 50 percent share of national income increases from 19 percent in 1955–56 to 23.6 percent in 1982–83, but then decreases sharply and almost continuously thereafter (20.6 percent in 2000–01 and 14.9 percent in 2014–15).

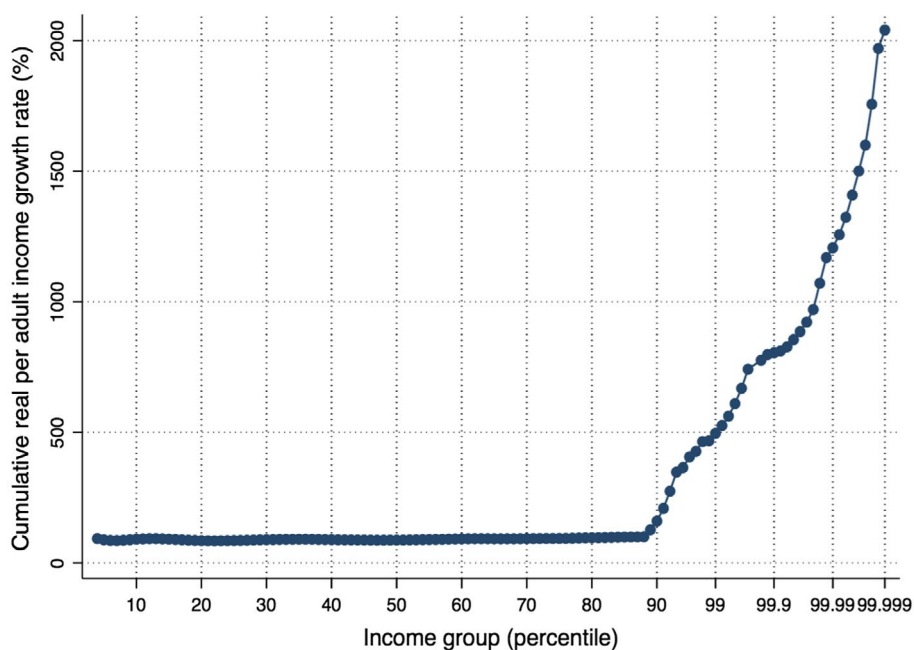


Figure 11. Income Growth by Percentile in India, 1980-2015: The “Cobra Curve” of Inequality and Growth.

*Source:* Authors’ estimates combining survey, fiscal and national accounts data.

*Notes:* Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1). The Figure shows that the average per adult real income growth rate between 1980 and 2015 of the top 0.001% income group was 2040%. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

TABLE 1  
TOTAL GROWTH RATES BY INCOME GROUP IN INDIA, 1980-2015

Income Group (distribution of per-adult pre-tax national income)	Total Cumulated Per Adult Real Growth (1980-2015)			
	India (%)	China (%)	USA (%)	Western Europe (%)
Full population	201	776	74	44
Bottom 50%	90	386	10	34
Middle 40%	94	733	54	36
Top 10%	435	1232	139	62
incl. Top 1%	775	1800	230	74
incl. Top 0.1%	1134	2271	355	79
incl. Top 0.01%	1699	2921	499	90
incl. Top 0.001%	2040	3524	698	124

*Notes:* Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1). Estimates for China, USA, Western Europe are based on WID.world and the World Inequality Report (wir2018.wid.world). Growth rates are net of inflation.

*Source:* Authors’ estimates combining survey, fiscal and national accounts data.

Appendix 13 displays income share estimates and Gini coefficients over the 1951–2015 period. The pretax income Gini coefficient was 46.3 percent in 1951, it decreased to 39.6 percent in 1982 and rose to 60.3 percent in 2014–15.

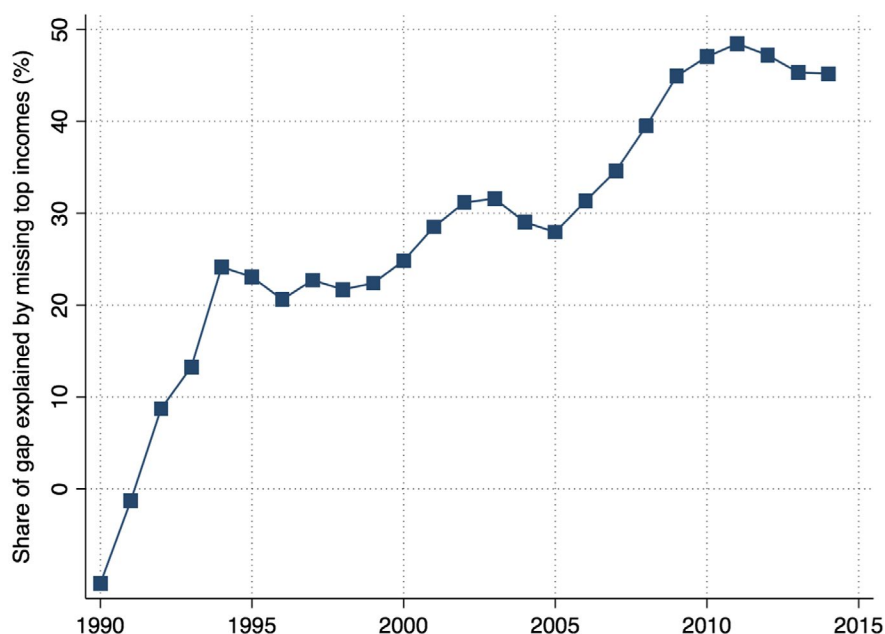


Figure 12. Importance of Missing Top Incomes in India since 1990: Share of Gap Between Survey Income and National Accounts Explained by Missing Top Incomes

*Source:* Authors' estimates combining survey, fiscal and national accounts data.

*Notes:* This graph shows the share of the gap between (reconstructed) survey income and income from the National accounts, which can be explained by the absence of top incomes in survey data. In practice, we compare the share of the gap between fiscal income from the national accounts (assumed to be 70% of national income), reconstructed survey incomes and reconstructed survey incomes, corrected at the top with tax data. Section 2 provides a description of these concepts. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

#### 4.3. Total Growth Rates by Income Group

We now measure total growth rates across the full distribution of incomes over the 1980–2015 period and compare these results to other countries available in the WID.world database, namely China, France and the USA. We also provide global growth estimates for the corresponding global groups.

Table 1 and Figure 12 show that income growth rates in India over the 1980–2015 period substantially increase as we progress upwards through the distribution of income. The bottom 50 percent of earners experiences a growth rate of 90 percent over the period, while the top 10 percent saw a 435 percent increase in their incomes. The equivalent figures for the top 0.01 percent and top 0.001 percent were 1699 percent and 2040 percent, respectively. Appendix 10 shows the same results on an annual growth rate basis.

Unequal growth dynamics over the period are not specific to India. Income growth rises the higher up the income distribution one proceeds in China, in the USA and in France as well. India's dynamics are, however, striking: it is the country with the highest gap between the growth of the top 1 percent and growth of the full population (near factor 4 difference in growth rates between these groups). It is also interesting to note that bottom 50 percent of earners grew 4 times more

TABLE 2  
SHARE OF TOTAL GROWTH CAPTURED BY INCOME GROUPS, 1980-2015: INDIA, CHINA, THE USA, WESTERN EUROPE

Income Group (distribution of per-adult pre-tax national income)	India (%)	China (%)	USA (%)	Western Europe (%)
<i>Total</i>	<i>100</i>	<i>100</i>	<i>100</i>	<i>100</i>
Bottom 50%	10.4	13.3	2.9	17.4
Middle 40%	21.2	43.4	33.1	36.6
Next 9%	40.0	28.4	31.2	29.3
Top 1%	28.3	14.9	33.	16.8
Top 0.1%	11.3	6.8	17.1	6.5
Top 0.01%	4.8	3.5	8.5	2.8
Top 0.001%	2.0	1.5	3.9	1.3

*Notes:* This table shows the share of national income growth captured by different income groups between 1980 and 2015. Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1). Estimates for China, USA, Western Europe are based on WID.world and the World Inequality Report (wir2018.wid.world).

*Source:* Authors' estimates combining survey, fiscal and national accounts data.

slowly in India than in China, whereas the middle 40 percent Indians grew nearly 8 times more slowly than their Chinese counterparts. Differences between the two countries among top groups are much less pronounced.

While Table 1 is particularly meaningful from the perspective of individual growth dynamics (what individuals observe), it is also useful to balance this with information on the share of total growth captured by different income groups. Indeed, high income growth at the individual does not necessarily translate into a high share of total growth captured at the macro level. Table 2 shows that the top 0.1 percent earners captured more total growth than the bottom 50 percent (11 percent vs. 10 percent of total growth) over the period. The top 0.1 percent of earners represented less than 800 000 individuals in 2014–15, this is equivalent to a population similar to Delhi's IT suburb, Gurgaon. It is a sharp contrast with the 397 million individuals that made up the bottom half of the adult population in 2014–15. At the opposite end of the distribution, the top 1 percent of Indian earners captured 28 percent of total growth, as much as the bottom 83 percent of the population. The comparison of these figures with China and other countries is particularly noteworthy. Out of the four countries, India is the country where the middle 40 percent benefitted the least from total growth over the period. We discuss this “missing middle class” issue in the next sections of the paper. The bottom 50 percent however captured a similar share of total growth in India and in China (respectively 10 percent and 13 percent).

Table 3 shows income levels and income thresholds for different groups and corresponding adult population size in 2014–15. Top 1 percent earners earn on average INR 2.9 million (21 times national average) versus INR 40,700 (0.3 times national average) for the bottom 50 percent and INR 101,100 (0.6 times national average) for the middle 40 percent.

Table 4 shows the growth rate over different income groups in India for the 1951–80 period. The situation is reversed as compared to the 1980–2015 period: the higher the group in the distribution of income, the lower the growth rate over the period. Real per adult income of the bottom 50 percent and middle 40 percent groups grew substantially faster (respectively 87 percent and 74 percent) than average income (65 percent). On the contrary, top 0.1 percent, top 0.01 percent

TABLE 3  
INCOME INEQUALITY IN INDIA, 2014-15

Income Group (distribution of per-adult pre-tax national income)	Number of Adults	Income Share (%)	Income Threshold	Average Income	Comparison to Average (ratio)
Average	794 305 664	100	0	138 426 INR	1
Bottom 50%	397 152 832	14.7	0	40 671 INR	.3
Middle 40%	317 722 266	29.2	63 728 INR	101 084 INR	.7
Top 10%	79 430 566	56.1	195 445 INR	776 567 INR	6
incl. Top 1%	7 943 057	21.3	1 303 946 INR	2 954 386 INR	21
incl. Top 0.1%	794 306	8.2	4 459 114 INR	11 346 371 INR	82
incl. Top 0.01%	79 431	3.4	18 260 916 INR	47 154 896 INR	341
incl. Top 0.001%	7 943	1.4	77 801 552 INR	188 558 192 INR	1362

Notes: Distribution of pre-tax per adult national income, benchmark scenario (A0B1CID1). Population estimates for 2014.

Source: Authors' estimates combining survey, fiscal and national accounts data.



TABLE 4  
TOTAL GROWTH RATES BY PERCENTILE IN INDIA, 1951–80

Income Group (distribution of per-adult pre-tax national income)	Total Real Per Adult Income Growth (1951–80) (%)
Full population	65
Bottom 50%	87
Middle 40%	74
Top 10%	42
<i>incl. Top 1%</i>	5
<i>incl. Top 0.1%</i>	–26
<i>incl. Top 0.01%</i>	–42
<i>incl. Top 0.001%</i>	–45

*Note:* Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1).

*Source:* Authors' estimates combining survey, fiscal and national accounts data.

TABLE 5  
SHARE OF TOTAL GROWTH CAPTURED BY PERCENTILE GROUPS IN INDIA, 1951–80

Income Group (distribution of per-adult pre-tax national income)	Share of Income Growth Captured (1951–80) (%)
Full population	100
Bottom 50%	28
Middle 40%	49
Top 10%	24
<i>incl. Top 1%</i>	0.9
<i>incl. Top 0.1%</i>	–1.8
<i>incl. Top 0.01%</i>	–1.0
<i>incl. Top 0.001%</i>	–0.4

*Notes:* Distribution of pre-tax per adult national income, benchmark scenario (A0B1C1D1). Growth rates are net of inflation.

*Source:* Authors' estimates combining survey, fiscal and national accounts data.

and top 0.001 percent income groups experienced a severe decrease in their real incomes (–26 percent, –42 percent and –45 percent, respectively). Appendix presents the same data with annualized growth rates.

Table 5 reveals that bottom 50 percent group captured 28 percent of total growth over the 1951–80 period, vs. 49 percent for the middle 40 percent and 24 percent for the top 10 percent.

#### 4.4. Growing Share of Income Gap Explained by Top Incomes

We compare the theoretical fiscal income obtained from national accounts<sup>16</sup> to our reconstructed fiscal income and the total income estimated from household surveys. This comparison reveals the share of survey and national accounts discrepancy discussed in Section 3.1, that can be attributed to the absence of top earners in survey data. We find that our reconstructed fiscal income bridges a growing and non-negligible gap between national accounts surveys data. The share of the gap explained by our reconstructed fiscal income rises from about 0 percent in 1990 to more than 28 percent in 2014.

<sup>16</sup>Supposed to be 70 percent of net national income, following Banerjee and Piketty (2005).

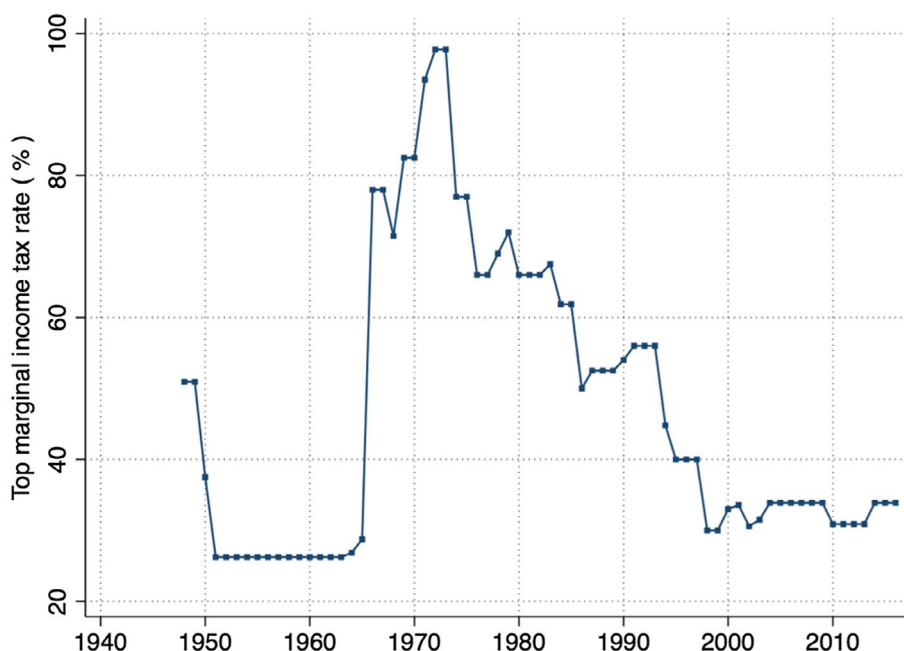


Figure 13. Top Marginal Income Tax Rate in India, 1948–2016

Source: Authors' estimates based on ITD and Union Budget Speeches.

Note: Figures include super tax on top incomes. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

#### 4.5. Measurement Issues and Robustness Tests

One of the main assumptions underlying our results is that tax data measures the actual income shares of the richest. There are a number of reasons why this may not entirely be true. A potential issue with tax data is that the surge in top incomes may reflect improvements in the Income Tax Department's ability to measure and tax the incomes of the richest. The tax cuts in the early 1990s might have reduced the incentives among the wealthy for evading the income tax. Indeed, there were a number of innovations in tax collection in the 1990s, such as the 1998 introduction of the "one in six rule" that required everyone who satisfied at least one of six criteria (such as owning a car and travel abroad) to file a tax return. We note however that the decline in the top marginal rate was quite moderate during the late 1980 to 2000 period: the top marginal tax rate dropped from 50 percent in 1987–88 to less than 40 percent in 1999–2000 (and only minor evolutions after, see Figure 13). By comparison, the increase in the share of the top 0.01 percent was huge: it went up from 0.7 percent in 1987–88 to more than 2 percent in 1999–2000. If this entire change is to be explained by a shift in tax rates, the implied elasticity would have to be enormous. Another key limitation of the Indian tax series is the ten-year break from 2000 to 2010. We did not find evidence of significant changes in the tax legislation, that could explain the rise in top shares post-2000. We also note that the post-2000 rise does not mark a discontinuity in the series, but comes more as the prolongation of rising top shares trend observed in the 1990s. The trend is also in line with the rise of inequality

observed in consumption surveys, in wealth rich lists and recent wealth inequality series (Anand and Thampi, 2016). The release of tax tabulations for the years 2000 to 2010 would allow us to better analyze year-on-year evolutions for this crucial period.

In order to test the robustness of our results to data limitations (including the tax data gap of the 2000s and the growing gap between national accounts and consumption surveys), we present our results along the 54 estimation strategies described in Section 2.2.6. These 54 scenarios reflect a wide range of alternative assumptions to make up for the lack of consistent data for the entire distribution of income. We find that our main results are robust to all the strategies tested.

Appendix 13a-c show the evolution of the top 1 percent, 0.1 percent and 0.01 percent shares from 1922 to 2015 across the 54 scenarios, along with our benchmark series (thick red line). The results only differ slightly between the different scenarios before 2005. In 1982–83, the top 1 percent share indicates 5.5 percent in the lower case scenario vs. 6.6 percent in the upper case. After 2005, the spread between scenarios is higher: top 1 percent income shares indicate 20.3 percent in the lower case scenario and 27.7 percent in the upper case scenario in 2014–15. The higher spread after 2005 is essentially due to strategy B assumptions (i.e. whether NSSO consumption surveys in 2005 and 2010 are rescaled upwards). Our benchmark strategy consists in rescaling the income levels estimated from NSS upwards—on the basis of IHDS data—to temper the rise in top shares at the end of the period. Considering these assumptions, the trends are remarkably similar across all scenarios, but the true top share values could be higher than what we obtain in our benchmark results.

Results for the middle 40 percent and the bottom 50 percent groups are relatively more sensitive to our sets of scenario assumptions, as Appendix 13 show. We find a 2.5 p.p. spread on lower case and upper case scenarios for middle 40 percent shares on average and an average 8 p.p. spread for bottom 50 percent income shares. This spread is essentially due to assumptions on the savings profiles of lower consumption groups (strategies A0, A1, A2). The A0 scenario reflects a mid-range position between the 0 negative assumption (scenario A2) and the profiles obtained from the IHDS dataset, with arguably excessive negative savings rates.<sup>17</sup> Long run results for bottom 40 percent and middle 50 percent groups are consistent across all scenarios: a slight increase from 1951–52 to 1983–84 and a significant decrease afterwards.

To sum up, we see our set of alternative scenario assumptions as a way to shed light on many of the gaps in our current knowledge about Indian income inequality. Our results are robust to a wide range of alternative assumptions but we do not pretend that these new series are definitive.<sup>18</sup> More modestly, we hope they can

<sup>17</sup>We note particular divergences around between 1978 and 1983 for both middle 40 percent and bottom 50 percent shares. This is explained by the fact that from 1978 to 1983, as shown in Appendix 14c, we do not have survey distributional data and we interpolate them on the basis of 1978 and 1983 information. The combination of interpolated survey income levels for these specific years and certain of our strategies—in particular strategy C3 ( $p_{x1} = 80$ ) and D3 (concave junction profile), tend to reduce “next 9 percent” income levels (i.e. individuals above the bottom 90 percent but below the top 1 percent) and relatively increase levels of the bottom 90 percent. These “extreme” scenarios are the less plausible of the set of assumptions in our view.

<sup>18</sup>Note that we do not apply differential price indexes to urban or rural income groups. A comprehensive analysis of inequality in India would ideally include series corrected for variations between urban and rural price index dynamics. Tax tabulations provided by ITA are however only available at the national level for the recent period and cannot be corrected for urban vs. rural price variations. We plan to better take into account these dimensions in further work on regional inequality in India.

encourage the publication of full series from 2000 to 2010. All computer codes are provided in the data appendix of the paper and can be used to produce alternative strategies, particularly if novel data addressing current gaps were to be released.

## 5. DISCUSSION

### 5.1. *The mid-1980s Turnaround*

Our findings confirm and amplify the conclusions of Banerjee and Piketty (2005) on Indian inequality in the long run, namely i) a marked decrease in inequality in the early forties ii) an even stronger reduction in top income shares in the 1950–70s and iii) a significant increase from the mid-eighties onwards. Current income inequality in India is higher than during pre-independence period. This holds true from the creation of the Income Tax in 1922 to independence in 1947 when comparing the top 1% share of national income, but also for the pre-1922 period. Before 1922, the best available estimates show that the top 0.1% income share varied between 5 and 7 percent of national income vs. more than 8 percent today in our benchmark, conservative scenario.

We note that the reduction in top income shares was smaller during the inter-war period than the reduction which occurred throughout the 1950–70s. This seems consistent with the interpretation posited for industrialized countries' (Piketty, 2001; Piketty and Saez, 2003). The shock induced by the Great Depression of the 1930s and the War had relatively lesser impacts in India than in the USA and Europe. In India, strong government control along with an explicit goal to limit the power of the elite<sup>19</sup> seems to have played a key role in reducing top income inequality after independence in 1947. The set of "socialist" policies implemented up to the 1970s included nationalizations, strong market regulation and high tax progressivity.

Railways were nationalized in 1951, air transport in 1953, banking in 1955, 1969<sup>20</sup> and 1980, oil industry in 1974 and 1976 to cite but a few. Along with the transfer of private to public wealth and reduction of capital incomes they implied, nationalizations came along with government setting over pay scales. In the private sector, incomes were constrained by extremely high tax rates: between 1965 and 1973, top marginal tax rates rose from 27 percent to 97.5 percent.<sup>21</sup> Such evolutions may have reduced rent-seeking behavior at the top of the distribution via a process of discouragement, which in presence of excessive bargaining power and rent-seeking is the efficient thing to do (Piketty *et al.*, 2014).

As discussed in Section 1.1, from the early 1980s onwards, the Indian economy underwent reverse transformations. The turnaround of income inequality (in 1983–84, see Figures 7–11) seems consistent with the implementation of a new economic policy agenda to disengage the public sector and to encourage entrepreneurship as well as foreign investments. The start of the process has been associated with the nomination of Rajiv Gandhi as Prime Minister in 1984.

<sup>19</sup>An anecdote may reflect this view on fairness which prevailed in Nehruvian politics: when industrialist Tata asked then Prime Minister J. Nerhu about allowing profits in State-owned industries, J. Nehru answered, "*Never talk to me about profit, [...] it is a dirty word*" (Shah, 2018).

<sup>20</sup>14 banks were nationalized, representing 70% of the sector.

<sup>21</sup>These figures include the "super tax" on top incomes.

In terms of tax progressivity, however, the downwards trend in fact started earlier—in the mid-1970s (Figure 13). That said, marginal income tax rate remained at fairly high levels until 1984–85 when Rajiv Gandhi’s government reduced the rates from 62 percent to 50 percent. Why year 1983–84 marks so abruptly the turning point of our inequality series over the recent period remains a topic of enquiry. Several factors can be at play: anticipations in the 1984–85 change in the top marginal tax rate, and anticipations of a more pro-business environment, could have had a positive impact on top incomes, in line with the rent-seeking theory posited by Piketty *et al.* (2014). Other factors could include the combination of a strong recession in the agricultural sector the previous year (–5 percent agricultural production due to severe droughts in 1982–83), which impacted income groups at the bottom. A surge in top earners filing tax returns, because of less stringent tax policies, is not to be excluded and could explain why the change is so abrupt this year. However, the fact that the rise in inequality is prolonged throughout the 1990s and in the recent period shows that this factor is very unlikely to play decisive role in the observed trends.

Available macro series also show that the wage share in the private corporate sector has been declining in India since the early to mid-1980s (in contrast to the 1970s, when the profit share was declining; see Nagaraj (2000) and Tendulkar (2003), which is consistent with the time for the turnaround proposed here.

Our results are also consistent with the evolution of Indian wealth inequality according to All-India Debt and Investment survey data (Anand and Thampi, 2016). Recently released wealth inequality estimates indeed show a sharp increase in wealth concentration from 1991 to 2012, particularly after 2002. The increase in wealth inequality at the top of the distribution is a logical outcome of the highly unequal income growth we report in this paper over the recent period.

### 5.2. *Shining India for the Rich Mostly?*

The sharp rise in inequality experienced by India since the mid-1980s complements, rather than contradicts, the Indian poverty literature, which documents a large decrease in absolute poverty since the early 1990s. Indeed, the data we use to track growth at the bottom of the distribution mobilizes NSS surveys, ie. the source used to measure poverty headcount ratios in India.<sup>22</sup> This study however replaces such evolutions in the context of the overall distribution of income growth and suggests that poverty reduction could have been faster should macroeconomic growth had been more evenly distributed.

We also shed light on a particularly striking characteristic of Indian growth over the past three decades, which deserves attention: the rise of the “middle class” was only very relative. We show that incomes of the middle 40 percent grew at 102 percent over the 1980–2015 period. Compared to industrialized countries’ growth rates for this group, the figure is impressive. In the Indian context however, the middle 40 percent were notably below average growth (102 percent vs. 187 percent). Since 1980, the middle 40 percent group in India captured a much smaller share of total growth (25 percent) in than its counterparts did in China or Europe (more than 40 percent) or even the USA (33 percent).

<sup>22</sup>The World Bank documents a division by 2 of the number of absolute poor in India since the early 1990s (World Bank, 2017). See Section 1.2, note 4.

This result should help us better characterize what has been termed as “the rise of India’s middle class”. From the perspective of our newly income inequality dataset, “Shining India” corresponds to the top 10 percent of the population (approximately 80 million adult individuals in 2014) rather than the middle 40 percent. Relatively speaking, the shining decades for the middle 40 percent group corresponded to the 1951–80 period, when this group captured a much higher share of total growth (49 percent) than it did over the past forty years. It is also important to stress that, since the early 1980s, growth has been highly unevenly distributed within the top 10 percent group. This further reveals the unequal nature of liberalization and deregulation processes. India in fact comes out as a country with one of the highest increase in top 1 percent income share concentration over the past thirty years.

Under which conditions could growth have been more equally distributed in India over the past decades? A discussion of the factors explaining the persistence of strongly unequal growth in India since the 1980s goes beyond the scope of this paper, which primarily seeks to provide more systematic income inequality. We note however that income inequality was at similar levels in China and India in the early 1980s, it grew at a similar pace till the mid-2000s and was stabilized in China thereafter (see Piketty *et al.*, 2019), while the rise continued in India. What explains such a divergence between the two Asian giants? Differences in national policies, rather than mechanical forces are likely to account for an important part of income inequality variations observed among countries (Alvaredo, Chancel, *et al.*, 2018).<sup>23</sup>

## 6. CONCLUSION

We combine historical and novel tax data with household surveys and national accounts data in order to produce novel estimates of the full distribution of adult pre-tax income in India, from 1951 to 2015 and for the top 1 percent of the distribution from 1922 to 2014.

We document a large increase in the level of inequality in India over the recent period and a large increase in the current level as compared to survey-based statistics generally used in public debates. We find that our results are robust to a large set of alternative estimation strategies addressing important data gaps. According to our benchmark estimates, the top 1 percent income share is at its highest level (22 percent) since the create of the Income Tax during the British Raj, in 1922. Top income shares and top income levels were sharply reduced in the 1950s to the 1970s at a time when strong market regulations and high fiscal progressivity are implemented. During this period, bottom 50 percent and middle 40 percent incomes grew faster than average. The trend reverted in the mid 1980s with the development of pro-business policies.

We certainly do not have the capacity to put an end to debates over the impact of economic reforms on inequality or poverty India. Our contribution is in fact relatively modest; better data series on the distribution of income inequality can

<sup>23</sup>On this point, relatively low investments in infrastructure or health for the bottom 50 percent and middle 40 percent in India may have played an important role in the persistence of a growing income gap between these groups and upper income classes in India since the mid-2000s, while inequality was contained in China over the past decade.



and should lead to better informed democratic conversation on the state of the Indian economy. While being more satisfactory than currently available income inequality estimates, these new series are in no way definitive. We stress the need for more research dedicated to reconcile micro and macro estimates of income and consumption inequality in India, both at the regional and national level. Efforts following the Distributional National Accounts Guidelines (Alvaredo *et al.*, 2016), published on the WID.world database, seek to go in this direction. Ultimately, meeting this objective will not be possible without the participation and expertise of official statistical agencies, in India and elsewhere.

## REFERENCES

- Alvaredo, F., A. B. Atkinson, L. Chancel, T. Piketty, E. Saez, and G. Zucman, "Distributional National Accounts Guidelines: Concepts and Methods used on WID.world," WID.world Working Paper 2016/1, 2016.
- Alvaredo, F., A. Bergeron, and G. Cassan, "Income Concentration in British India, 1885–1946," *Journal of Development Economics*, 127, 459–69, 2017.
- Alvaredo, F., L. Chancel, T. Piketty, E. Saez, and G. Zucman, *World Inequality Report 2018*. Harvard University Press, 2018. Available on [www.wid.world](http://www.wid.world).
- Anand, I. and A. Thampi, "Recent Trends in Wealth Inequality in India," *Economic and Political Weekly*, 51, 59–67, 2016.
- Atkinson, A. B., "Measuring Top Incomes: Methodological Issues," in Atkinson and Piketty (eds), *Top Incomes Over the 20th Century: A Contrast Between Continental European and English-Speaking Countries*, Oxford University Press, 2007.
- Atkinson, A. B. and T. Piketty (eds), *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*. Oxford University Press, Oxford, 2007.
- Atkinson, A. B. and T. Piketty (eds.), *Top Incomes: A Global Perspective*. Oxford University Press, Oxford, 2010.
- Banerjee, A., and T. Piketty, "Top Indian Incomes, 1922–2000," *The World Bank Economic Review*, 19, 1–20, 2005.
- Blanchet, T., and L. Chancel, "World National Accounts Series Methodology," WID.world Working Paper Series 2016/2, 2016.
- Blanchet, T., J. Fournier, and T. Piketty, "Generalized Pareto Curves: Theory and Applications," WID.world Working Paper. 2017/3, 2017.
- Burkhauser, R. V., N. Hérault, S. P. Jenkins, and R. Wilkins, *What has been Happening to UK Income Inequality Since the mid-1990s? Answers from Reconciled and Combined Household Survey and Tax Return Data* (No. w21991). National Bureau of Economic Research, 2016.
- CSO, *Report of the Group for Examining Discrepancy In PFCE Estimates From NSSO Consumer Expenditure Data And Estimates Compiled By National Accounts*. Division Central Statistical Organisation, Ministry of Statistics & Programme Implementation, Government of India, New Delhi, 2008.
- Deaton, A. and J. Dreze, "Poverty and Inequality in India: A Re-Examination," *Economic and Political Weekly*, 37, 3729–48, 2002.
- Deaton, A. and V. Kozel, "Data and Dogma: The Great Indian Poverty Debate," *The World Bank Research Observer*, 20, 177–99, 2005.
- Kotwal, A., B. Ramaswami, and W. Wadhwa, "Economic Liberalization and Indian Economic Growth: What's the Evidence?," *Journal of Economic Literature*, 49, 1152–99, 2011.
- Kulshreshtha, A. C. and A. Kar, "Estimates of Food Consumption Expenditure from Household Surveys and National Accounts," in A. Deaton and V. et Kozel (eds), *The Great Indian Poverty Debate*, Macmillan, New Delhi, 2005.
- Lakner, C. and B. Milanovic, "Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession," *World Bank Economic Review*, 30, 203–32, 2015.
- Maddison, A. (2012) *Maddison Project Database*, Available on [www.ggcc.net/maddison](http://www.ggcc.net/maddison).
- Nagaraj, R., "Indian Economy since 1980? Virtuous Growth or Polarization?," *Economic and Political Weekly*, August 5, 2000.
- OECD, *2017 OECD Economic Survey of India*, OECD Publications, Paris, 2017.
- Ozler, B., G. Datt, and M. Ravallion, *A Database on Poverty and Growth in India*, World Bank Development Research Group, Washington, DC, 1996.



- Piketty, T., *Les hauts revenus en France au XXème siècle*, Grasset, Paris, 2001.
- Piketty, T. and E. Saez, "Income Inequality in the United States, 1913-98," *The Quarterly Journal of Economics*, 118, 1-41, 2003.
- Piketty, T., E. Saez, and S. Stantcheva, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 6, 230-71, 2014.
- Piketty, T., L. Yang, and G. Zucman, "Capital Accumulation, Private Property and Rising Inequality in China, 1978-2015," *American Economic Review*, 109, 2469-96, 2019.
- Shah, S. *The Tata Group: From Torchbearers to Trailblazers*, Penguin Portfolio, 2018.
- Sivasubramanian, S. (ed.), *The National Income of India in the Twentieth Century*, Oxford University Press, Oxford, 2001.
- Sundaram, K. and S. D. Tendulkar, "Poverty Outcomes in India in the 1990s," in A. Deaton and V. Kozal (eds), *The Great Indian Poverty Debate*, Macmillan India Ltd, New Delhi, 2005.
- Tendulkar, S. D., *Organised Labour Market in India—Pre and Post Reform*, University of Delhi, Delhi School of Economics, New Delhi, 2003.
- World Bank, *PovcalNet Database*, 2017. Available on <http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx>. Accessed June 15, 2017.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site:

- Appendix 1:** GDP growth in India, 1960–2015
- Appendix 2:** List of corrections done to raw tax files
- Appendix 3:** List of NSS consumption surveys and summary statistics
- Appendix 4:** Income-consumption profiles by percentile
- Appendix 5:** Aggregate savings in India, 1983–2015
- Appendix 6:** Junction percentile, fiscal years 1922–23 to 2014–15
- Appendix 7:** Income levels in India, 2011: survey vs. tax data
- Appendix 8:** Total survey income with and without tax corrections in India, 1990–2015
- Appendix 9:** Average annual per adult income growth by income group in India, 1980–2015
- Appendix 10:** Average annual per adult income growth by income group in India, 1951–80
- Appendix 11:** Share of growth captured by income group in India, 1951–80
- Appendix 12:** Income growth by percentile in India, 1980–2015
- Appendix 13a:** Income shares and Gini coefficient, 1951–2015
- Appendix 13b:** Top 1 percent national income share in India, 1922–2015: results from 54 alternative scenarios
- Appendix 13c:** Top 0.1 percent national income share in India, 1922–2015: results from 54 alternative scenarios
- Appendix 13d:** Top 0.01 percent national income share in India, 1922–2015: results from 54 alternative scenarios
- Appendix 14a:** Top 10 percent national income share in India, 1951–2015: results from 54 alternative scenarios
- Appendix 14b:** Bottom 50 percent national income share in India, 1951–2015: results from 54 alternative scenarios
- Appendix 14c:** Middle 40 percent national income share in India, 1951–2015: results from 54 alternative scenarios