# ADDENDUM TO THE BOOK

# INEQUALITY DATA

## Measuring Inequality

Inequality is another word for disparity or unevenness. It is a multidimensional phenomenon. Scholars study inequality of income, wealth, education, health, access, assets, housing, and other variables. Inequality is also conceptualized in several distinct ways in the different disciplines that study the issue. In economics, the focus is often on studying distributions in whole populations; in sociology and anthropology, the focus is on studying groups and their differences; in geography, the focus is on differences between spatial units (like nations, states, cities etc.). We understand inequality by measuring outcomes on dimensions or variables that matter. Inequality measurement is a vibrant and active sub-field in economics, as is, in sociology, the measurement of sociological conceptualizations of inequality (such as segregation, isolation, etc.). There are hundreds of measures of inequality. However, only a handful of measures are used in practice; as a result, the choice is not as difficult as it could be.

Among the multiple dimensions along which inequality is studied—income, wealth, assets, educational attainment, health outcomes (longevity, infant mortality, maternal mortality, etc.)—the primary focus here is on income with a secondary focus on wealth. Income has a direct relationship to welfare and opportunity and as a result it is doubtless the most commonly studied variable among inequality researchers. Wealth is also important, but is generally much more difficult to measure because the wealthy have many ways to hide and obfuscate their

holdings. Some analysts—especially those associated with the Human Development approach—argue that the focus on income takes attention away from other important markers of welfare, such as education and health.[1]

I do not dispute that education and health are very important, but suggest that income is most important because it is the primary determinant of education and health outcomes and it is income inequality (along with government failures to provide adequate public goods) that leads to inequalities in education and health. I do provide some information on education later in this appendix, but as I show there, these figures probably hide as much as they reveal. In fact, the clinching argument in favor of focusing on income is that so little is known about it despite its overwhelming significance. Chapter 7 has been written precisely because so little is known about the different inequalities of income.

Economists tend to analyze the world in terms of the individual (person, firm, or institution), whereas other social scientists, especially from sociology and anthropology, typically think in terms of groups. Mark Granovetter, a prominent sociologist, suggested that economics as a discipline is "undersocialized" whereas sociology is "oversocialized."[2] A nation, in the economic framework, is a collection of individuals, each one serving his own individual interest. Their social identities or spatial locations do not matter in this "abstract" form of

---

**APPENDIX 3**

[1] The human development approach is detailed in the annual Human Development Report, first published in 1990. It came about as a result of dissatisfaction among leading economists from the Indian subcontinent like Mahbub ul Haque and Amartya Sen about the dominating focus on income among inequality researchers and the inadequacies of income to fully explain the opportunities available to individuals. There are now several editions of a separate India Development Report and also multiple development reports at the state level in India.

[2] Granovetter, M. 1985. Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology* 91:481-510.

inequality. But in the other social sciences, the most important unit of analysis is usually not the individual but the group or location. As a result, the social world is understood through the concepts of in-group cooperation and out-group derogation or conflict. Depending on the context, group identity and interest can either be less important or significantly more important than individual identity and interest. Consider, for example, the contrasting self- and group-interests of Wall Street bankers ("greed is good") vs. soldiers ("band of brothers") or Bollywood stars vs. the builders of the sets on which they frolic. Let us think of group identity in terms of social identity. It is fair to say that investigations of social identity and inequality form the core of the field of contemporary sociology.

For example, if a society is composed of two groups—black and white, or Forward caste and Backward caste, or Hindu and Muslim—the only way to understand whether they differ as groups is to measure things that say something about the quality of their lives and see whether there is any difference, and, if there is, how much it is. In other words, the extent of division in any social system is understood by measuring or quantifying the extent of difference or inequality between the divisions. The difference should be over something that matters. To say that black (or Backward caste) has darker skin pigmentation than white (or Forward caste) is beside the point. The question is, whether meaningful outcomes for the group called black (or Backward caste) are measurably different from the group labeled white (or Forward caste) on scales that most reasonable people can agree on? There are intricacies of measurement and making meaning from measurements—and some of those are discussed below—but the basic point must be clear. If social divisions are real, then at some level they are measureable. They will show up as differences in things like income, wealth, assets, longevity, infant mortality, years of education, and so on. The extent of difference is social inequality.

In economics, the primary area of interest is in the distribution of income and the distribution of human capital (simply: education); wealth distribution is also studied, but to a lesser extent, because it is harder to track and crack. Some of the most important contributions to our understanding of income and human capital inequality have come from notable economists like Anthony Atkinson, Gary Becker, Ronald Bénabou, Gary Fields, Branco Milanovic, Thomas Piketty, and Amartya Sen, who have discussed ways of measuring inequalities in income distributions, the ideology and ethics of different distributions and their measurements, and the meanings and consequences of such inequalities for growth and economic development.[3]

By and large, it is possible to associate specific academic disciplines with specific conceptualizations of inequality. To simplify, let us think of three distinct conceptualizations and their associated academic disciplines: income distribution or income inequality in economics, social inequality in sociology and anthropology, and spatial inequality in geography.

## Visualizing inequality

[3] Anthony B. Atkinson, 1970, On the Measurement of Inequality, *Journal of Economic Theory* 2:244-63; and 1983, *The Economics of Inequality*. Second edition. Oxford: Clarendon Press. Gary S., 1962, Investment in Human Capital: A Theoretical Analysis, *Journal of Political Economy* 70:9-49. Ronald Bénabou, 1996, Equity and Efficiency in Human Capital Investment: The Local Connection, *Review of Economic Studies* 63: 237-264. Gary A. Fields, 1980. *Poverty, Inequality, and Development*, Cambridge: Cambridge University Press; and 2001. *Distribution and Development: A New Look at the Developing World*. Cambridge, Mass.: The MIT Press. Branco Milanovic, 1998, *Income, Inequality, and Poverty during the Transition from Planned to Market Economy*. Washington DC: World Bank. Thomas Piketty, 2014, *Capital in the Twenty-first Century*, trans. Arthur Goldhammer, Cambridge, Mass.: Harvard University Press. Amartya K. Sen, 1973, *On Economic Inequality*. Oxford: Clarendon Press; and 1992, *Inequality Reexamined*, New York and Oxford: Russell Sage Foundation and Clarendon Press.

Perhaps the best description of income inequality is the one provided by Jan Pen in his parade of dwarfs and a few giants.[4] Let us say that it was possible to arrange a parade of all income earners in a society where each person's height is proportional to her income; that is, an average income earner would be of average height, say about five and a half feet. If such a parade were to last for one hour, starting with the lowest income earner and ending with the highest, one would "see" the income distribution of a given territorial space in dramatic light. The parade would begin with individuals walking on their hands, representing negative income earners. Using 1978-79 data for the United Kingdom, Anthony Atkinson summarizes the rest of the parade:[5]

> Next come old age pensioners (with) the height of the pensioners not much over a foot. After them come low paid workers, with…the rule of women first for each occupation… The slowness with which the height increases is one of the striking features of the parade… It is only when we pass the average income (with twenty-four minutes to go) that events begin to speed up, but even when we enter the last quarter hour (the top 25 percent), the height of marchers is only some 7'. But then they begin to shoot up. Police superintendents are 11' tall. The average doctor or dentist is some 14'. Around 20' come senior civil servants, admirals and generals. The chairman of a medium sized company may be 35' and for larger companies his height could be 35 yards. Indeed, the highest paid directors are…over 70 yards tall. They are not, however, the last, since the final part of the parade is made up of people of whom Pen says 'their heads disappear into the clouds and probably they themselves do not even know how tall they are'.

---

[4] Jan Pen, 1971, *Income Distribution*, Hammondsworth: Allen Lane.
[5] Anthony B. Atkinson, 1983, *The Economics of Inequality*. Second edition. Oxford: Clarendon Press, p. 14-15.

Keeping that vivid image in mind, consider an illustration of the different conceptualizations of inequality in Figure 1 which combines the "Pen's parade" insight with different ways of organizing information about a society that is divided into two groups. Let us call the groups "grey" and "black." One can imagine these two groups in any way one likes—Forward and Backward caste, Hindu and Muslim, vegetarian and non-vegetarian, etc.. Let us also assume, like Pen and Atkinson, that the height of each individual is proportional to his or her income. Figure 1a shows a random arrangement of 50 individuals—25 each from the groups grey and black. Because they are randomly arranged, it is not possible to say much about the overall distribution other than what is obvious: that both the grey and black groups have some tall (or high income) individuals, some short (low income) individuals, and some individuals of medium height (middle income).

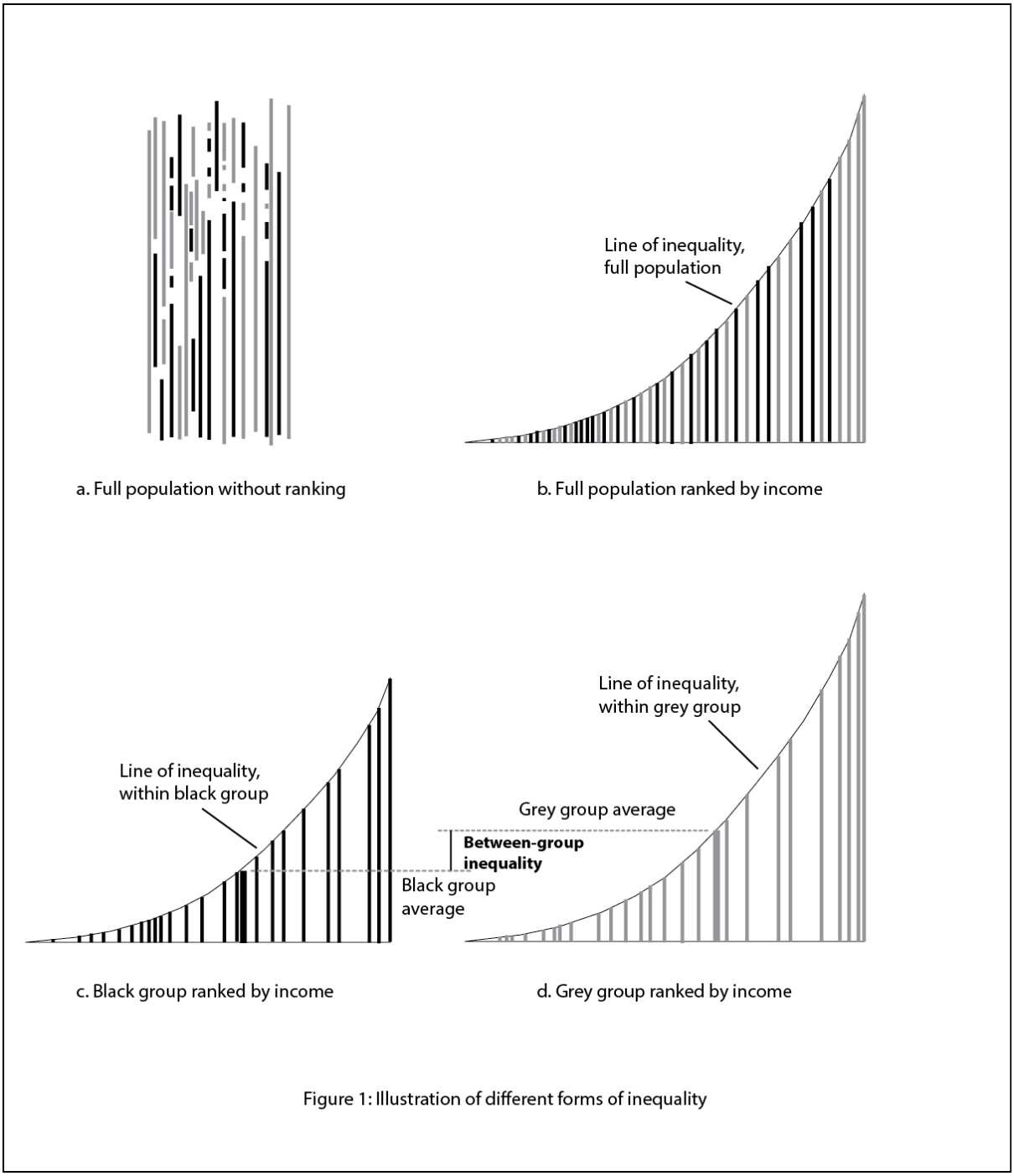



Figure 1: Illustration of different forms of inequality

When we sort these individuals by height and arrange them by rank (in Figure 1b), we are able to see Pen's Parade. This curve represents inequality in this full population. The properties of this curve—such as, how much it sags away from the diagonal—can be estimated (using methods that range from simple to complicated) and summary calculations of inequality derived

from it. This curve is analogous to income inequality in the full population of grey and black individuals in this hypothetical distribution. Economists are primarily interested in this distribution.

Now, the same exercise can be done with the grey and black populations separately. We can sort and rank the black group (Figure 1c) and grey group (Figure 1d) separately and estimate the inequality within these groups by analyzing their separate curves of inequality. These can be thought of as "within-group" inequalities (analogous to inequality within Forward castes and within Backward castes separately). Now, each group (grey and black) has an average height (or income). In this illustration the grey average is higher than the black average. The difference between these averages is analogous to "between-group" inequality; that is, the inequality between Forward and Backward castes (or, as I show below, between Forward and Backward states or districts).

So, the distribution of income can be studied using an abstract method in which everyone in India—from the most destitute to Mukesh Ambani—is ranked without reference to anyone's social identity (this is the common method used by economists). Or, it can be done by grouping society by social identity and looking at the differences within and, in particular, between groups.

This "within" and "between" distinction is important. We know that an average (or mean) is merely one representation of a group. This is illustrated by the "Bill Gates walks into a bar" story: before he enters the bar, the average wealth of its occupants may be USD 100,000; after he enters it could be a billion dollars or more (depending on how many people are in the bar). All groups have internal differences—highs and lows within the groups that are not captured by an average. So, it goes without saying, that all Brahmans do not have a higher

income or a better starting point than all Dalits; conversely, all Dalits do not have a lower income or inferior starting point than all Brahmans.

Economists have developed a number of "decomposable" measures of inequality (such as the Theil Index and the decomposable Gini) which calculate the contribution of between-group and within-group inequality to total inequality. As a general rule, within-groups inequalities contribute more to total inequality than between-group inequalities.[6] But, in India, there is little useful information on within-group inequality: that is, inequality within Dalits or within Muslims, etc.. So, important as it is, we are unable to investigate this in any detail.

There is much interest among economists in creating summary measures—simple measures—that capture in a single number a sense of the inequality in a distribution.[7] Often these calculations are done by grouping the population into equal sizes—for example, broken into 10 segments of 10 percent of the population each (called deciles) or five segments of 20 percent of the population each (called quintiles) ranked by income. How much of the national income does the poorest decile earn? How much does the richest quintile earn? What is the ratio

[6] For instance, in the US, where the white vs. black gaps in wealth, income, and education are quite significant, the contribution of these differences in averages to total inequality is significantly less than the contributions of within-group (within-black and within-white) inequality. See Sanjoy Chakravorty, 1996, Urban Inequality Revisited: The Determinants of Income Distribution in U.S. Metropolitan Areas, *Urban Affairs Review* 31: 759-777. One of the interesting findings of that paper, that is now found more generally, is that inequality within the black population was higher than within the white population. For a recent example see Frédéric Chantreuil and Thérèse Rebière, 2016, *Decomposition of Income Inequality by Attributes: Does the Race Matter in the US*? Mimeo. Available at https://conf-tepp2016.sciencesconf.org/99419/document.

[7] Simple measures of inequality often miss some key features that interest inequality researchers. So, there are also many complex measures of inequality. In fact, a thriving subfield of inequality research is that of research on inequality measurement. Amartya Sen (1973, *On Economic Inequality*), in a single footnote, cited over 100 articles and books on the measurement of inequality in economics (counting nothing from sociology or geography).

of the income share of the richest (decile or quintile) to the poorest? What is the income share of the superrich—the top one percent, or the top one percent of the top one percent?

There are two alternative approaches in comparing different income distributions—whether to include the complete distribution (including all income earners) or simply compare the top and the bottom of the distribution. The former approach accounts for everyone whereas the latter approach is useful for investigating changes at the extremes of a given distribution. Using the latter suggests that the investigator is interested in studying income polarization. If the full distribution is to be used, certain measurement properties are considered desirable. Discussions on these desirable properties are available at many sources, the most accessible of which is on the World Bank's website.[8] In general there are five key axioms or principles that inequality measures should follow: The Pigou-Dalton transfer principle, the axiom of income scale independence, the principle of population, the axiom of anonymity, and the principle of decomposability.

These axioms are not, however, value-free. Consider the second axiom of income scale independence: that if every individual's income increases by the same proportion (say everyone receives a five percent increase in income), a proper inequality measure should not change. However, since the rich will receive higher absolute income increases than the poor, this is at best a *status quo* condition that can also be considered regressive. If we believe the utilitarian argument that each successive marginal income increase produces less utility or welfare (since the first lakh rupee one earns is valued more highly than say the tenth lakh), then an equal proportional increase in all incomes produces less overall utility or welfare than when the same

[8] See http://www.worldbank.org/poverty/inequal/methods/. Also see P. B. Coulter, 1989, *Measuring Inequality: A Methodological Handbook,* Boulder: Westview Press; and Sanjoy Chakravorty, 2007, *Fragments of Inequality*.

total income increase is distributed more heavily among the lower income groups. Partly in response to such normative anomalies in supposedly value-free inequality measures, a group of explicitly normative or welfarist measures have been created. Among inequality scholars, Anthony Atkinson's measure based on explicit choices of "inequality aversion" is well known.[9]

In general, we avoid these complicated measures. Instead, for a summary measure, we use only the Gini Coefficient. It is an useful visual analog of both the Pen's Parade and the distribution of income by groups like deciles or quintiles. There is much information on the Gini Coefficient on the net. The Wikipedia page is as useful as any.[10] The Gini can take a value between 0 and 100. When 0, everyone has the same income (or wealth, or whatever is being measured); when 100, one rich person has everything. Wherever possible, I will use information that is even simpler than the Gini.

To put the magnitude of Gini income inequality in perspective: the lowest Gini indexes for income in the world are in the mid to high 20's. These low inequalities can be found in countries reputed for their high tax and high redistribution regimes (as in Scandinavian countries like Iceland, Finland, Sweden, and Norway) or in post-Soviet societies in central Europe (like Ukraine, Slovenia, Slovakia, the Czech Republic, and Belarus) that have retained some or much of the egalitarian ideology and apparatus of the Soviet years. The highest Gini indexes in the world are in the lows 60's. The most egregiously high levels are in southern Africa (specifically South Africa, Namibia, and Botswana), in regimes that are deeply divided, especially by racial groups or extractive classes where the key is control of gems and precious minerals.[11]

---

[9] Anthony B. Atkinson, 1970, On the Measurement of Inequality.
[10] https://en.wikipedia.org/wiki/Gini_coefficient.
[11] These countries are said to suffer from a "resource curse", which is the counterintuitive effect in which being "blessed" with rich natural resources often turns into a "curse" for the very people in whose lands those resources are located. It is not hard to see how the "resource curse"

## Expenditure, Income, and Wealth Inequality

In this section I discuss the reality of inequality in India using the best available information and data. First I consider economic inequality and the three different ways it is conceptualized: by expenditure (what people spend), by income (what people earn), and by wealth (what people own). Following that, I consider the available information on social inequality; that is, inequality between social groups. The sources of the analyses include official data (produced by the government) and unofficial data (produced by non-government institutions).

Broadly, the story of inequality in India that emerges from the available resources and studies is one of high and growing economic inequality, a story that is at odds with the official narrative on inequality in India—that it is low and unchanging. The argument I make is not an isolated one. It is one that is supported by all serious scholars of inequality in India. Why then is there such a fundamental difference between the official and scholarly conclusions? The simple answer is that the official position in India is based on information on expenditure, whereas the rest of the world studies income (and, increasingly, wealth). There are other, deeper explanations, but we can discuss those only after we have gone over the basics.

Branco Milanovic, one of the leading scholars of inequality in the world, writes: "How unequal is India? The question is simple, the answer is not."[12] That is largely because, in India, we can say nothing about *income* inequality from official data because income has *never* been officially measured. This seems like an outrageous statement, but it is true. This is not because the Indian

---

thesis can easily be applied to the India's poorest regions (southern Bihar, Jharkhand, Chhattisgarh, and Odisha) which also happen to be its most resource-rich.

[12] Branco Milanovic, 2016, The question of India's inequality, http://glineq.blogspot.in/2016/05/the-question-of-indias-inequality.html.

government does not measure social conditions. Quite the contrary. The Indian system for gathering social statistics—led by the National Sample Survey Organization (NSSO)—is considered among the most sophisticated and professional in the developing world.[13] But the NSSO does not estimate income in any of its national surveys. It estimates consumption or expenditure. That is, it estimates what households spend rather than what they earn. As a result, the estimates of inequality in India are for expenditure rather than income.

Expenditure inequality is, however, not considered an adequate measure of inequality of condition. Households at lower income levels tend to spend all they earn; in fact, they often have to borrow to meet unexpected expenditures (like illness), or sell assets (like land and gold, if they have any), or rely on remittances (money sent by close relatives working somewhere else). Higher income households, on the other hand, are able to save; that is, they do not spend all they earn, and instead put the additional money into assets like stocks, gold, and property. Their unspent income is converted into wealth.

As a result, expenditures do not capture the true range of quality of life conditions, and expenditure inequality does not provide a good sense of the true inequality of quality of life (or opportunity or access to value-producing resources). Expenditure, by definition, is narrower in range than income, and, by definition, expenditure inequality is lower than income inequality. Some analysts have estimated the gap between income and expenditure inequality for the Gini

[13] The NSSO was created in 1950 (initially it was called NSS). It has so far conducted 71 rounds of surveys (for which the data are available). Its major surveys in a ten year cycle include: Consumer Expenditure and Employment & Unemployment (twice); Social Consumption (health, education etc.) (twice); Un-organised Manufacturing (twice); Un-organised services (twice); and Land & Livestock holdings. The NSSO also undertakes special surveys, such as the 70th round titled the Situation Assessment Survey of Agricultural Households, All India Debt and Investment & Land and Livestock Holdings (whose findings I will refer to later in this chapter). It also conducts the Annual Survey of Industries.

Coefficient/Index to be around 5-6 points.[14] As we shall see, the gap in India is considerably larger. It is so large that the measurement of expenditure inequality may be meaningless in India.

The origins of this choice (to measure expenditure rather than income) goes back to the early post-independence years when basic decisions were being taken on a number of issues (including this one). The focus then was more on poverty than inequality. In fact, inequality did not become a serious issue to study or fight until after the mid-1970's, after some development economists began to discover that economic growth did not automatically mitigate poverty or improve the lives of populations at the bottom of the income distribution.[15] At very low levels of income (as India had in the post-independence years), expenditure (rather than income) was rightly considered to be the superior measure of poverty. As a result, from its very first surveys in 1951, the NSS (as it was named then) was geared to measuring how much people spend (to understand, among other things, how many calories they intake), in order to understand the depth and breadth of poverty in the country. The expectation was that policies to mitigate poverty would be based on these data. That method (of measuring expenditure rather than income) continues to be used to the present day.[16]

Expenditure

[14] See P. B. Coulter, 1989, *Measuring Inequality: A Methodological Handbook,* Boulder: Westview Press; Hongyi Li, Lyn Squire, and Hengfu Zhou, 1998, Explaining International and Inter-temporal Variation in Income Inequality, *The Economic Journal* 108: 26-43.

[15] Among the most influential studies that returned attention to inequality was Hollis B. Chenery and Moises Syrquin, 1975, *Patterns of Development: 1950-1970*, New York: Oxford University Press.

[16] The use of this method has continued despite a rising awareness of its problems and findings. As Milanovic points out: from the early 1990s, "the survey numbers began to diverge more and more from National Accounts statistics: NSS was showing consistently lower rates of growth, and higher poverty than many people thought it should be given India's fast growth." Moreover, it was increasingly clear that the top earners were not being captured by the surveys, which is probably why the NSS survey averages were low and did not match with the National Accounts statistics.

Table 1 lists the Gini Index estimates of inequality of expenditure or consumption in rural, urban, and all India from three sources (Himanshu, Subramanian and Jayaraj, and NSSO) for the last four decades.[17] The estimates are not identical because different analysts tend to use slightly different assumptions and methods for calculating the Gini Index; but the underlying data for all three sets of estimates are the same: all were collected by the NSSO. Let us not focus on the minor differences between the different estimates (which are meaningless), nor the more important finding that expenditure inequality in urban India is consistently higher than in rural India (it is not particularly meaningful because the phenomenon of higher urban than rural inequality is seen all over the world).

Let us focus instead on the magnitude of inequality and its consistency. The magnitude of Gini inequality in rural India is seen to be in the high 20's and it appears to be more or less unchanged in four decades. The magnitude of Gini inequality is roughly 35-36 for urban and all-India, with a possible small uptick from the low 30's after the early-2000's. If these figures were true, that is, if they represented the reality of distribution, then inequality in India would be among the lowest in the developing world and among the most stable and unchanging.

In international comparisons of inequality, the low official Gini Indexes of the NSSO are usually taken at face value. In the absence of official data on income in India, there is a widespread conflation between income and expenditure inequality. They are assumed to be the same—which leads to the misleading conclusion that India is a low inequality country with a

[17] Himanshu, No date, Inequality in India, available at http://india-seminar.com/2015/672/672_himanshu.htm. Sreenivasan Subramanian and Dhairiyarayar Jayaraj, 2015, The Evolution of Consumption and Wealth Inequality in India; also see S. Subramanian and D. Jayaraj, 2015, *Growth and Inequality in the Distribution of India's Consumption Expenditure, 1983 to 2009-10*, WIDER Working Paper 2015/025. NSSO figures reported in http://planningcommission.nic.in/data/datatable/data_2312/DatabookDec2014%20106.pdf.

stable Gini hovering in the low to mid-thirties for decade after decade. The confusion is evident in many international documents: for example, in the World Development Report of the World Bank which mentions that "India had fairly low income inequality," in the United Nations Development Program which reports that the "income gini coefficient" in India is 33.9, and in policy papers by the International Monetary Fund which use the same figures.[18] Today, in early 2018, the websites of the World Bank and IMF that list inequality for all countries show India's income Gini Index to be 35.1, which we know is India's expenditure (not income) inequality level.

[18] World Bank, 2007, *World Development Report—Agriculture for Development*, The International Bank for Reconstruction and Development/The World Bank, p. 46. The UN data are available at http://hdr.undp.org/en/content/income-gini-coefficient. These calculations based on consumption data continue to be used and conflated with income data on a regular basis. A good example from the IMF is in Rahul Anand, Volodymyr Tulin, and Naresh Kumar, 2014, *India: Defining and Explaining Inclusive Growth and Poverty Reduction*, IMF Working Paper WP/14/63.

**Table 1: Expenditure Inequality in India, 1973-74—2011-12**

| | Source: Himanshu | | | Source: Subramanian & Jayaraj | | Source: NSSO | |
|---|---|---|---|---|---|---|---|
| | Rural Gini | Urban Gini | All-India Gini | Rural Gini | Urban Gini | Rural Gini | Urban Gini |
| 1970-71 | | | | 28.9 | 34.7 | | |
| 1972-73 | | | | 30.7 | 34.5 | | |
| 1973-74 | | | | | | 28.1 | 30.2 |
| 1977-78 | | | | 34.2 | 34.8 | 33.6 | 34.5 |
| 1983 | 27.1 | 31.4 | 29.8 | 31.6 | 33.9 | 29.7 | 32.5 |
| 1987-88 | | | | 30.2 | 35.7 | | |
| 1993-94 | 25.8 | 31.9 | 30 | 28.6 | 34.4 | 28.2 | 34.0 |
| 1999-2000 | | | | 26.3 | 34.7 | 26.0 | 34.2 |
| 2004-05 | 28.1 | 36.4 | 34.7 | 30.5 | 37.6 | 26.6 | 34.8 |
| 2009-10 | 28.4 | 38.1 | 35.8 | 29.9 | 39.3 | 27.6 | 37.1 |
| 2011-12 | 28.7 | 37.7 | 35.9 | | | 28.0 | 36.7 |

Income

But there are few serious analysts of inequality who would consider the NSSO expenditure data and the Gini Indexes calculated from them to represent the reality of inequality in India. Consider what we know from one of the most important alternative sources of large scale survey data—the India Human Development Survey (IHDS)—that is also the one major "unofficial" but reliable source of income inequality data in India. The IHDS is a nationally representative survey of about 41-43K households that has been carried out in two rounds so far: in 2004-5 and 2011-12.[19] The IHDS calculations show that income inequality is considerably higher than expenditure inequality: in the range of Gini 54 in 2004-5 and 2011-2.[20] These results bolster the innovative findings of Luke

[19] The IHDS is a joint undertaking by researchers at the University of Maryland and the National Council of Applied Economic Research (NCAER), New Delhi.

[20] There are several estimates of Gini using the IHDS data and they all vary slightly based on assumptions and adjustments made by the specific analyst. These figures are from Mehtabul Azam, 2016, Income Inequality in India 2004-2012. Also see Reeve Vanneman and Amaresh Dubey, 2013, Horizontal and Vertical Inequalities in India. More papers that use IHDS data are available at https://ihds.umd.edu/. It is important to note that the IHDS data allow both

Chancel and Thomas Piketty who combine household survey data (from NSSO and IHDS), national accounts statistics, and tax data to argue that income inequality in India is very high, perhaps the highest it has ever been, primarily because the share of national income accruing to the top one percent of income earners is 22 percent of the total income, the highest level in a century, far above the 6 percent it was in the early 1980's; a visual representation of Chancel and Piketty's findings is shown later in this chapter in Figure 2.[21]

Along with two colleagues (S. Chandrasekhar and Karthikeya Naraparaju), I have studied some aspects of income distribution over the last decade in rural India.[22] We analysed the Situation Assessment Surveys of Farmers/Agricultural Households undertaken by the NSSO in 2003 and 2013. We found that there was a very large difference between the two measurement concepts—income vs. expenditure inequality—where the Gini Indexes of per capita income and expenditure were around 60 and 30 respectively during this study period. We argue that while our findings are narrow in coverage (being limited to the agricultural sector, that covers roughly half the population) that narrowness itself leads to greater robustness. Therefore, the startling gap of 30 Gini points between expenditure and income inequality should be taken seriously. Added to the findings of IHDS and Piketty and his associates, these findings should conclusively burst the mythical balloon of low inequality in India.

In fact, I argue that the true level of income inequality in India is higher than anything calculated by any analyst so far. There are several reasons for taking this position. First, most

---

expenditure and income inequality to be estimated. The 2005 expenditure Gini estimate from IHDS is around 38, roughly equivalent to the NSSO survey based estimate of 36 for the same time. In short, the IHDS roughly captures the same population that NSSO does and is as reliable (or unreliable) as the latter.

[21] Luke Chancel and Thomas Piketty, 2017, *Indian Income Inequality, 1922-2014.*

[22] Sanjoy Chakravorty, S. Chandrasekhar, and Karthikeya Naraparaju, 2017, *Income Generation and Inequality in India's Agricultural Sector*.

inequality calculations are unlikely to include the very top and bottom ends of the income distribution. For example, our own work on rural India misses the population that has little or no income from agricultural activities; much of this group is likely to be the landless population that may comprise more than 40 percent of rural households.[23]

The far bigger problem is that most income data derived from surveys are likely to miss or have unreliable figures on the very top end of the income distribution. The Indian upper middle class is notoriously difficult to survey. Even if a survey team can make it to their doors (which is very hard to do in the gated housing estates in which the upper middle class tends to live), it is usually refused entry. The upper class is, of course, well and truly beyond questioning by anyone. For example, in the IHDS 2004-5 survey, the individual with the highest income out of the 41,000 plus families surveyed earned less than Rs. 22 lakh per year. It seems obvious that the IHDS survey missed the top one percent of earners. Even more troubling are the NSSO expenditure surveys. For the 2011-2 round, their highest spending group, the top five percent of urban India, averaged expenditures of merely Rs. 1.2 lakh per year. This is roughly what government college professors earn per month. I have no doubt that the NSSO also missed the top one percent (perhaps the top 2-3 percent) of consumers. On top of this is the well-known tendency of the poor to over-report and the rich to under-report their incomes.[24]

These problems with survey-based inequality calculations are beginning to become widely recognized. Laurence Chandy and Brina Seidel write: "Missing top incomes in household surveys is a long established problem in both developed and developing economies…The more new

---

[23] Vikas Rawal, 2008, Ownership Holdings of Land in Rural India: Putting The Record Straight, *Economic and Political Weekly*: 43-47.

[24] Jeffrey B. Nugent, 1983, An Alternative Source of Measurement Error as Explanation for the Inverted Hypothesis, *Economic Development and Cultural Change* 31: 385-396.

information we uncover about top incomes, the less faith we have in traditional survey-based inequality measures, and the less knowledge we can claim to have about the distribution of income across an economy's entire population." They "use the missing income between surveys and national accounts as a proxy for missing top incomes in surveys" following a method suggested by Christoph Lakner and Branco Milanovic.[25] The new calculations of Chandy and Seidel show large increases in Gini Indexes for several countries—the average increase is from 39 to 48. One of the largest increases is for India, where the Gini goes from 36 (calculated from official expenditure data) to 56 for the early 2010's.

That too may be an underestimate. If the Gini Index of agricultural income alone is 60 (as my work with Chandrasekhar and Naraparaju has shown), there is almost no doubt that the Gini index is significantly higher at the national scale. There are two reasons to justify this claim. First, we know that urban inequality is higher than rural inequality by 5-8 Gini points even using the flawed NSSO expenditure data. The gap between urban and rural inequality is likely to be higher with income data. Second, we know that average urban incomes are at least twice as high as average rural incomes for every size subgroup (decile or quintile) of the population.[26] Hence, if we add the two distributions—rural and urban—and it is possible to assess the income of the top one to two percent and bottom decile of households with any reasonable accuracy, a strong argument can be made that income inequality in India is among the most extreme in the world. It would not be a

[25] Laurence Chandy and Brina Seidel, 2017, How Much do we Really Know About Inequality Within Countries Around The World? Christoph Lakner and Branko Milanovic, 2013, Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession, World Bank Policy Research Working Paper, https://doi.org/10.1596/1813-9450-6719.
[26] Sonalde Desai, Amaresh Dubey, Brij L. Joshi, Mitali Sen, Abusaleh Shariff, and Reeve Vanneman, 2010, *Human Development in India*.

surprise if the true level of income inequality in India was in the range of Gini 65, on par with or higher than the highest known level of inequality in South Africa.

Wealth

The recently published figures on wealth inequality in India strongly suggest that the worst-case scenarios may indeed be true. There have been a spate of such publications in recent years, spurred by the annual Global Wealth Reports produced by Credit Suisse beginning in 2010. The tone of the Credit Suisse reports is largely celebratory, but the U.K.-based NGO Oxfam produces an annual Global Inequality report (based on the same wealth data) whose tone is anything but. For example, Oxfam's 2017 report argued that the richest eight billionaires in the world (Bill Gates, Amancio Ortega, Warren Buffett, Carlos Slim Helú, Jeff Bezos, Mark Zuckerberg, Larry Ellison, and Michael Bloomberg) had as much wealth between themselves as the poorest 50 percent of the world's population put together, and that the richest one percent of the world had as much wealth as the remaining 99 percent. The situation was "beyond grotesque," the Oxfam report said. For India, the Credit Suisse report stated that the richest 10 percent possessed 73 percent of the nation's wealth, whereas Oxfam stated that 73 percent of the wealth generated in 2016-7 in India went to just the richest one percent. According to the latest available Credit Suisse report, the Gini Index of wealth inequality in India is 83, among their list of the highest in the world.[27]

It is obvious that neither Credit Suisse nor Oxfam has the resources or ability to actually study wealth in India by themselves…and they do not. The primary data source for both is the decennial All India Debt and Investment Survey (AIDIS) carried out by the NSSO (last undertaken in 2012-3). One should be hesitant to rely on distant sources like Credit Suisse and Oxfam which

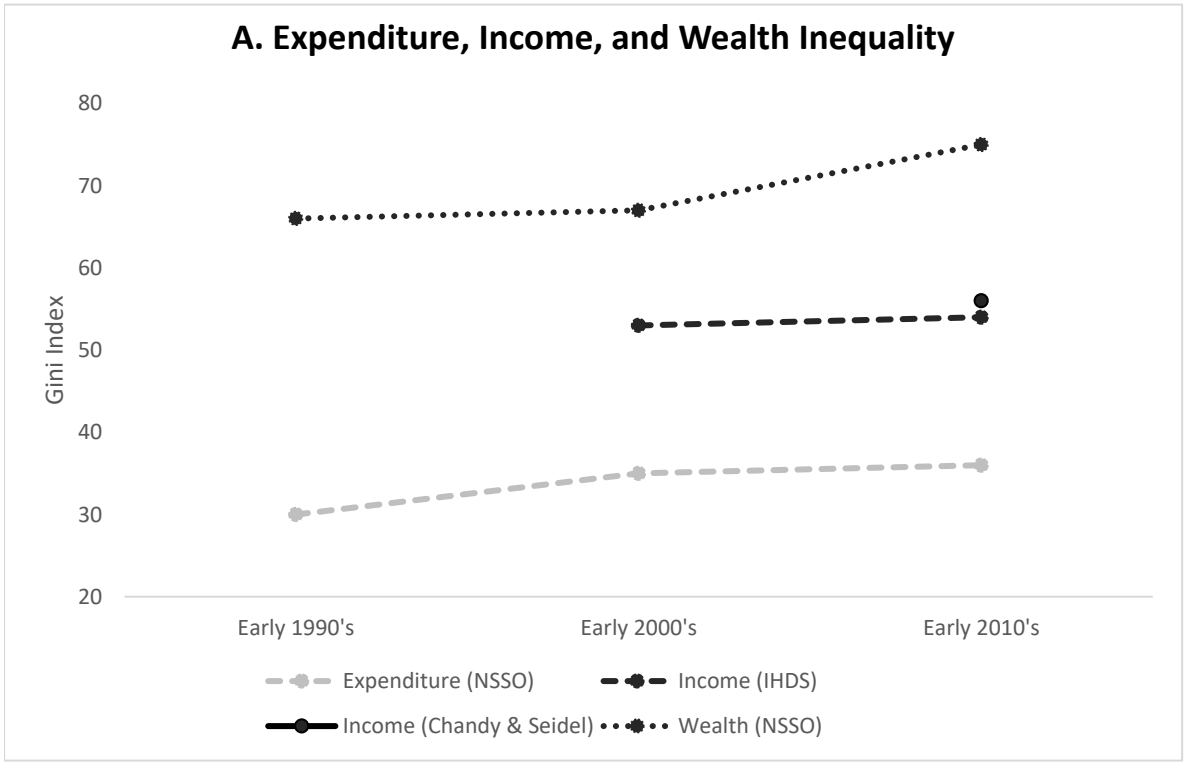
[27] Credit Suisse, 2017, Global Wealth Databook 2017.

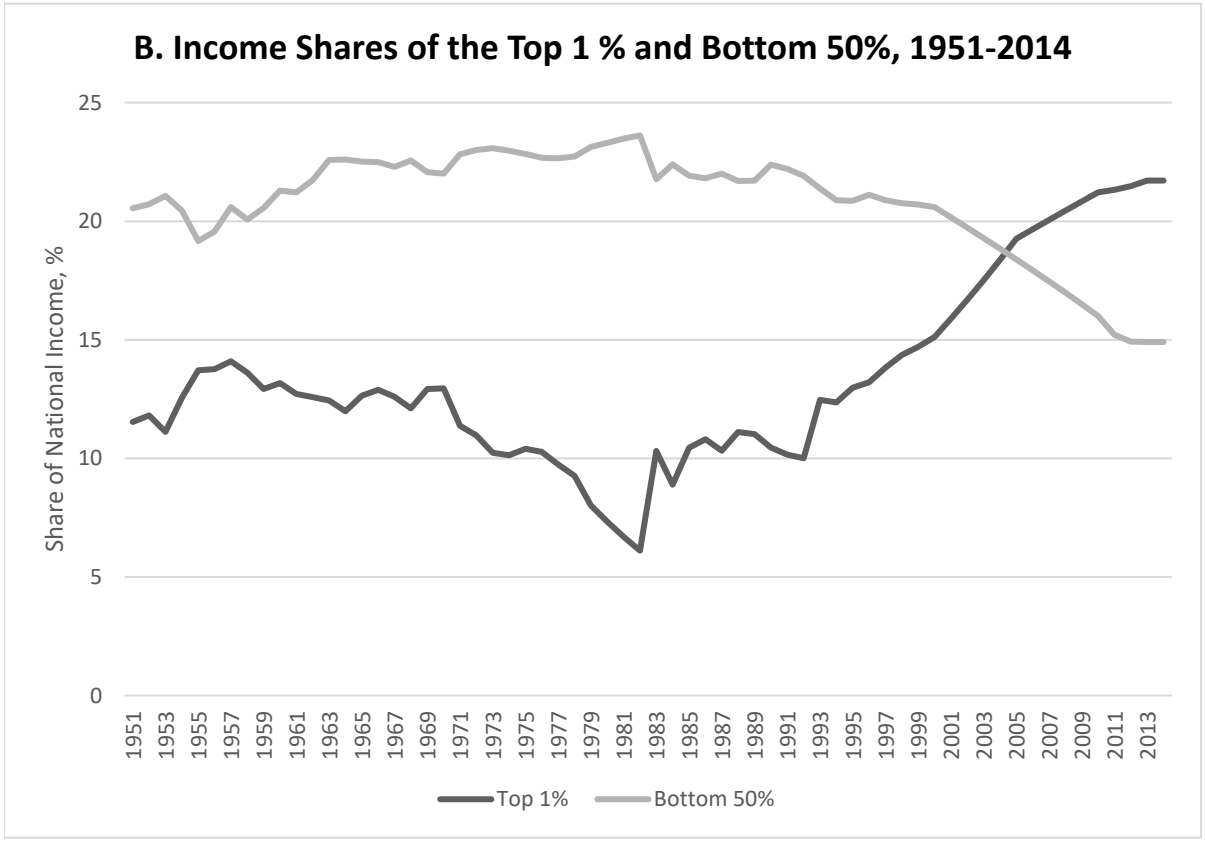
may be tweaking the raw data from NSSO in ways that are not visible to observers (which would seem to be the case if new findings are generated every year though no new AIDIS data are available). Their audience is global whereas we need to stay closer to the ground. Therefore, it may be better to look at the findings of scholars who have looked at the AIDIS data directly and carefully. The most recent of these is a paper by Ishan Anand and Anjana Thampi, in which the Gini Index of assets and net worth are shown to be 74 and 75 respectively in 2012, having risen from 65 and 66 in 1991 (and about the same levels in 2002).[28]

Note that the AIDIS data itself is open to serious criticism. It suffers from some of the main problems of the NSSO expenditure surveys; most notably, the difficulties with getting good data on the top of the distribution. The 2013 data show that in the decade that the Indian stock market boomed (roughly 2002 to 2012), the weight of shares/stocks actually went down to 0.13 percent of total wealth in its survey sample; it was Rs 112 per capita. This is simply not credible. Given that the market capitalization of all stocks was about 70 percent of GDP in 2012, the per capita stock holding should have been about Rs 60,000 per capita. The AIDIS sample clearly missed the *entire* population that owns stock—that is, all of India's upper middle class, and, of course, the entire upper class.

[28] Ishan Anand and Anjana Thampi, 2016, Recent Trends in Wealth Inequality in India. A longer time series (beginning in 1961-2) is available in Subramanian and Jayaraj, The Evolution of Consumption and Wealth Inequality in India, but for rural and urban data separately; no national estimates are presented. Similar analyses are available in Arjun Jayadev, Sripad Motiram and Vamsi Vakulabharanam, 2007, Patterns of Wealth Disparities in India during the Liberalisation Era.

**Figure 2: Change in Expenditure, Income, and Wealth Inequality over Time**




Sources: A. As shown in figure; B. Calculated from data in Luke Chancel and Thomas Piketty, 2017, *Indian Income Inequality, 1922-2014*.

shown to be in land and buildings. About 70 percent of rural wealth and a little under half of urban wealth is shown to be in land alone. As a result, much of the calculations (of wealth and inequality) depend on how accurately land is valued. There is serious case to be made that it is generally undervalued, especially given the five-fold increase in land prices across the country in the period 2000-2013, and is evidenced by the AIDIS calculation that in urban areas the value of land is roughly the same as the value of buildings. That too is simply not credible. Depending on the city and location, the value of land in total property is much above 50 percent, and for the upper class it easily surpasses 95 percent.[29] Therefore, it is very likely that the very high levels of wealth inequality calculated from AIDIS data are nonetheless significant underestimates because the survey was unable to capture the two main sources of wealth for the Indian upper middle and upper classes—stocks and land.

That possibility is highlighted by the findings of Chancel and Piketty in the second part of Figure 2, that show the long-term trajectories of income earned by the top one percent and the bottom 50 percent of families. If correct, this should be a severe indictment of, if nothing else, the absence of a serious discourse on inequality by government after government.

## Social Inequalities

As discussed earlier, social inequality is conceived, in a sociological sense, as the average difference between social groups. In our case, the social groups under consideration are those that have been identified as marginalized from pre-independence India (Scheduled Castes and Scheduled Tribes, to be called Dalit and Adivasi in the remainder of this discussion), new groups that have been brought into consideration for reservation or affirmative action after the Mandal Commission

[29] Sanjoy Chakravorty, *The Price of Land*.

recommendations (Other Backward Classes), and others (who, depending on the data available, may be called "Forward Castes" or "Brahmans" plus "Other Forward Castes"). The explicit assumption of the Indian system of reservations is that there are sizable gaps between the Backward and Forward groups, and the explicit goal of the reservations is to narrow those gaps. So the question before us is: What do we know about: (a) how far apart these groups are from each other, and (b) whether the gaps between them now are narrower or wider than before?

These are questions of fact and can only be answered with data. As we have seen above, the official data-gathering system in India does not collect some critical information for anyone (that is, income) and what it does collect surely does not include households at the top, which are very likely to be dominated by the Forward groups. We do not even know how much of the "top" is missing in surveys; the top one percent almost certainly, and perhaps as much as the top 2-3 percent. As a result, we know little about their income or wealth (that is, their land and stocks). Perhaps just as crucially, we do not know their social identities either (that is, what religion or caste they belong to). So, if our goal is to measure the gap from the "low" groups to the "top" groups, it is necessary to recognize from the very outset that it cannot be done, at least not easily, and not without violating some privacy barriers (such as those that protect the identities of tax payers from public scrutiny).

It is possible to tease out some indicators of what may have been happening to social inequality using the available official data from NSSO and unofficial data from IHDS. Several of the scholars who have been cited above have included sections on caste inequality in their larger studies of inequality. These studies are not uniform because they all use different formulations of social groups: in some studies, Dalits (Scheduled Castes) and Adivasis (Scheduled Tribes) are combined; some studies separate out Other Backward Classes or Forward castes or Brahmans, some do not; some allow the identification of religious identity; most do not. Therefore, these studies are

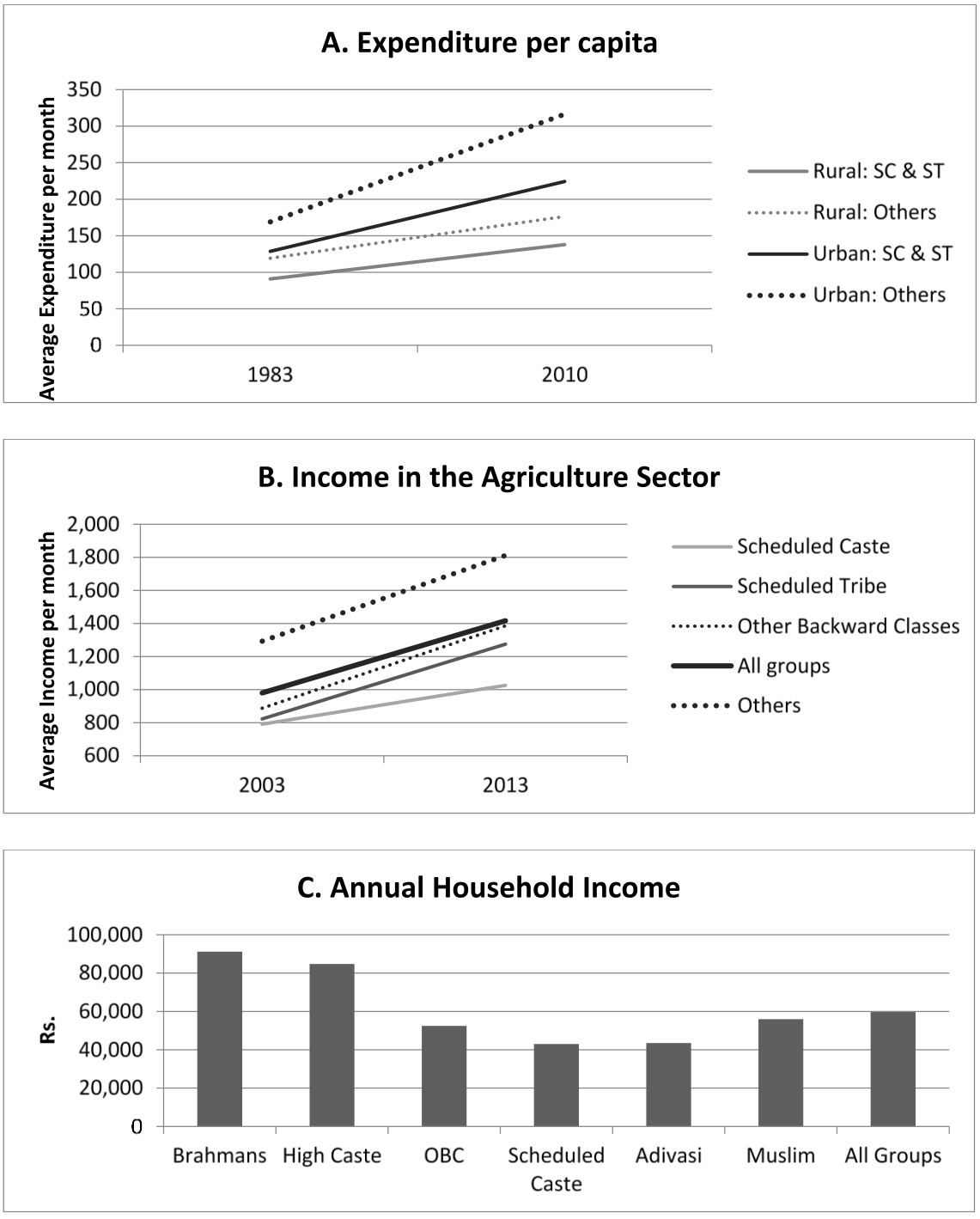
not comparable. They do not have the same definitions of low, middle, and high, do not use the same data over the same time period, and, above all, are flawed for the same reasons that all these studies are flawed—they measure inequality without knowing much about the top of the distribution. It is not the researchers' fault. They have to work with what they have, and what they have is flawed.

These indicative figures on social inequality are combined in a set of graphics in Figure 3 that provide some information on expenditure, income, wealth, and education over some period of time. This allows us to see the extent of the gaps between Forward and Backward groups and the changes in the conditions and their trajectories over recent decades. Note that if we had access to information on the uppermost section, these gaps would likely have been larger, and, crucially, growing over time.

As it is, the data show that the gaps between the averages of the Forward and Backward groups are considerable. Moreover, they have been growing over time for all the variables for which comparable temporal data are available. Consider expenditure (in Figure 3a-A), which we know is the primary welfare information collected by the NSSO and have seen earlier is the least meaningful marker of quality of life as far as inequality is considered. The highest-spending group is, as expected, the urban non-Dalit non-Adivasi population and the lowest-spending is the rural Dalit and Adivasi population. The ratio of their expenditures has increased from 1.9 to 2.3 from 1983 to 2010. That is, the average urban non-Dalit non-Adivasi person spent almost twice as much as the average rural Dalit or Adivasi in 1983; a quarter century later the former spent about 2.3 times as much as the latter. All the other gaps on expenditure widened during the period: between the rural Backward and the rural majority and between the urban Backward and the urban majority.

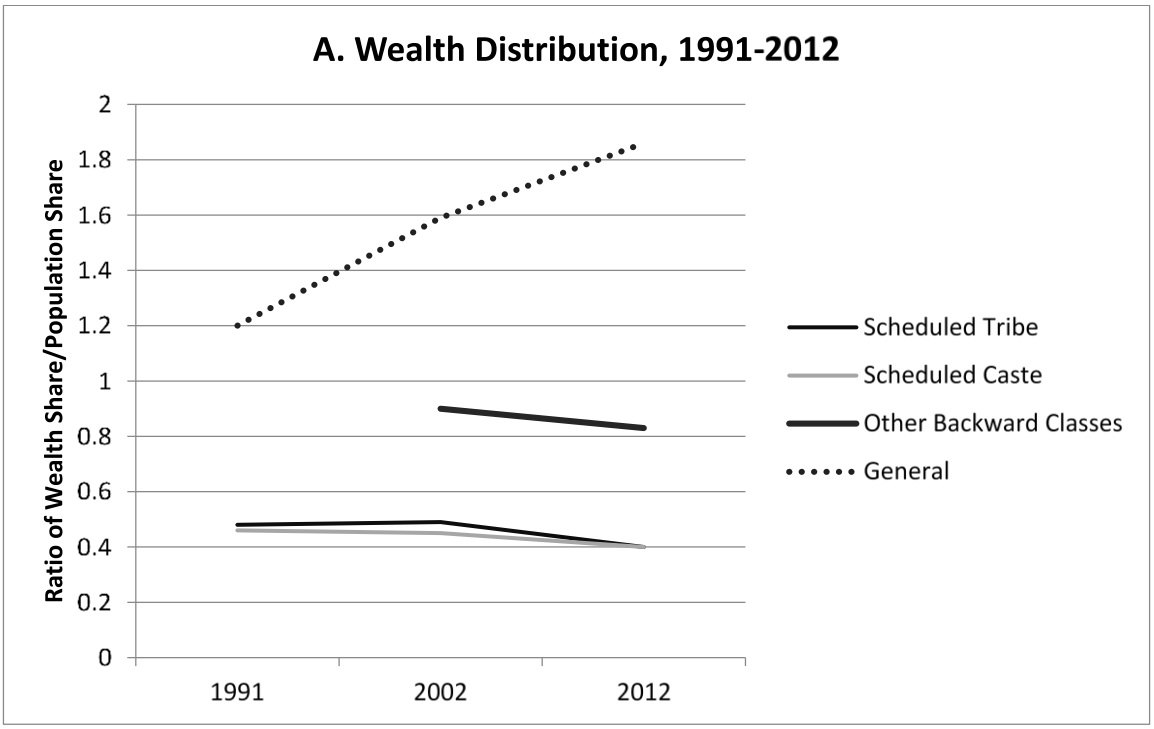
There is unambiguous evidence of a large and growing gap in expenditure between the socially marginalized and the rest of the population.

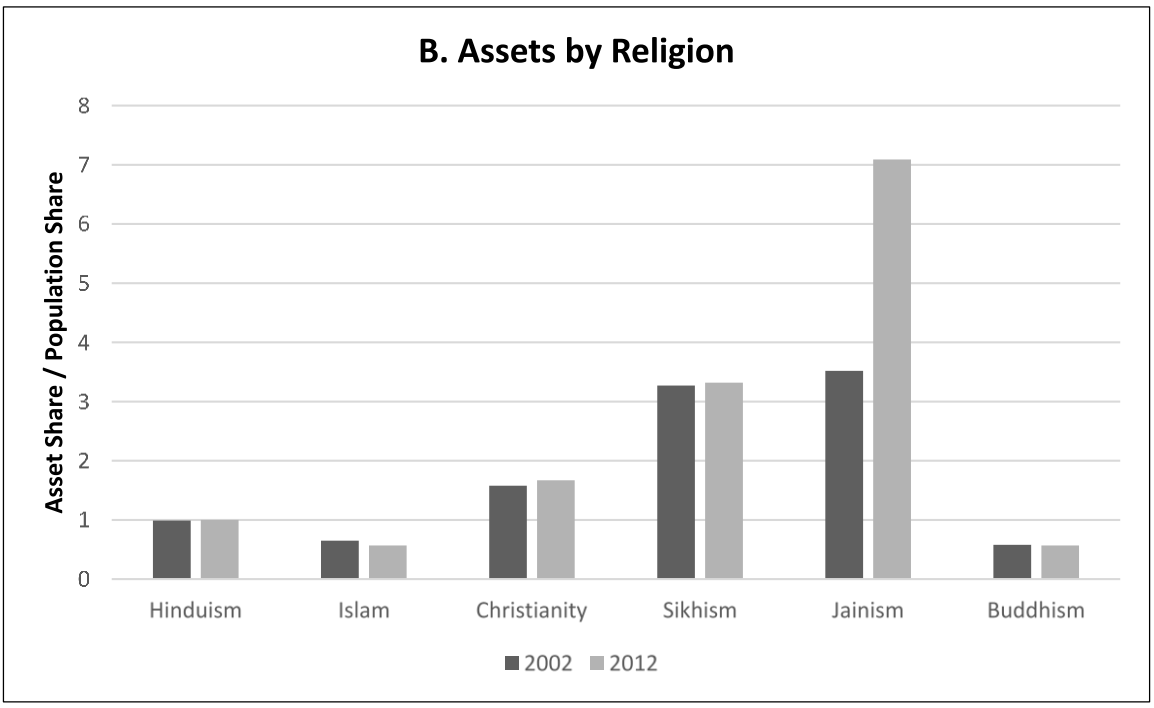
**Figure 3a: Expenditure and Income for Social Groups**




Sources: A. S. Subramanian and D. Jayaraj, 2015, Growth and inequality in the distribution of India's consumption expenditure 1983 to 2009-10; B. Sanjoy Chakravorty, S. Chandrasekhar, and Karthikeya Naraparaju, 2017, Income Generation and Inequality in India's Agricultural Sector: The Consequences of Land Fragmentation; C. https://ihds.umd.edu/.

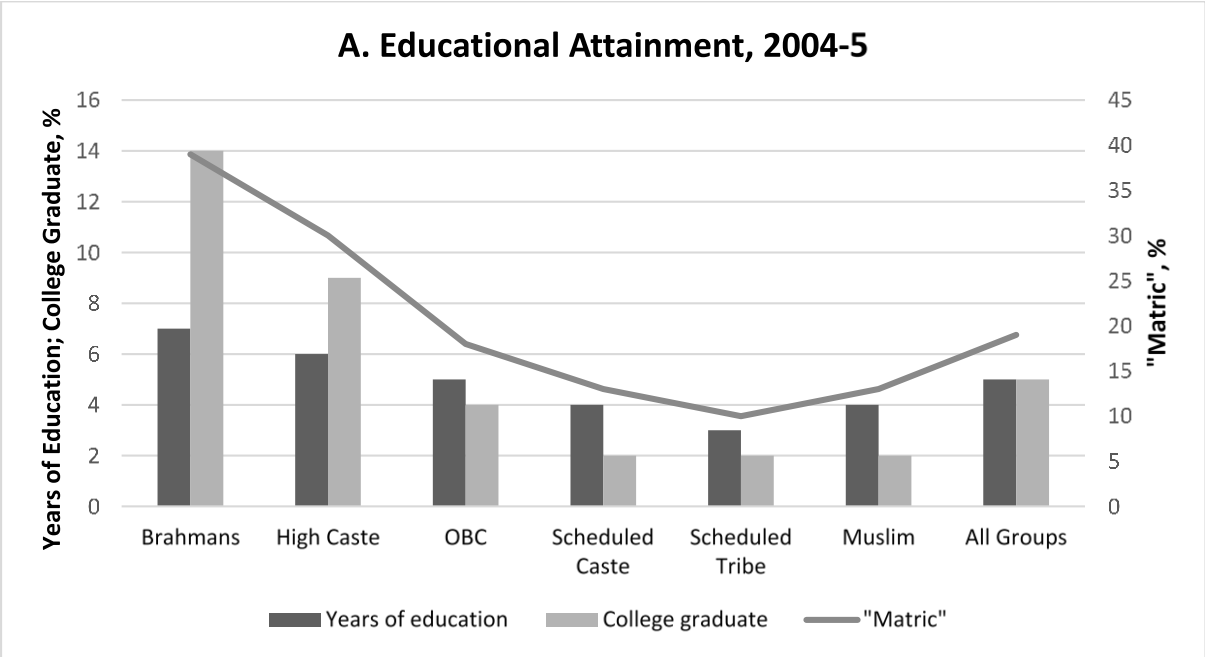
**Figure 3b: Wealth by Social Group**

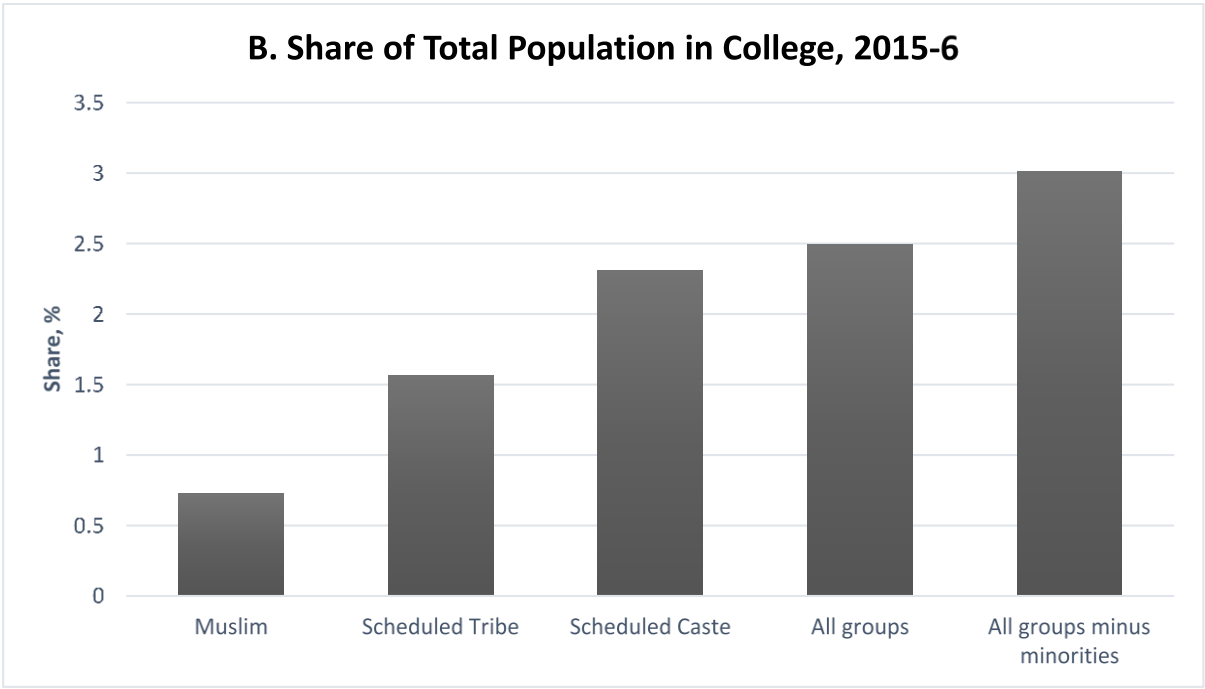



Sources: A. Ishan Anand and Anjana Thampi, 2016, Recent Trends in Wealth Inequality in India; B. Thiagu Ranganathan, Amarnath Tripathi, and Ghanshyam Kumar Pandey, 2017. Income Mobility among Social Groups. Economic & Political Weekly, 52(41), 73-76.

**Figure 3c: Education by Social Group**




Sources: A. IHDS available at https://ihds.umd.edu/; B. Calculated from Government of India, 2016, *All India Survey on Higher Education*, Ministry of Human Resource Development.

The income data from the agriculture sector (also from NSSO surveys) are more fine-grained and show somewhat better outcomes for the marginalized. Here, it is possible to differentiate between Dalit and Adivasi incomes, and between OBC and everyone else. Again we see large gaps between the non-marginalized "Others" and the Backward groups, but the gap is smaller than for expenditure (above). We see a growing gap between Dalit income and "Others", but the gap between "Others" income and both Adivasi income and OBC income, though large, narrowed between 2003 and 2013.

The income data from IHDS shown in Figure 3a-C are available for a single year, and they show, again, large gaps between Forward and Backward group incomes. Brahman average incomes (identifiable only in IHDS data) are twice as large as average Dalit and Adivasi incomes. The average incomes of OBC and Muslim families is about 20 to 30 percent higher than Dalit and Adivasi incomes.

Because the IHDS surveyed the same set of households at two time periods (2004-5 and 2011-2), it has become possible to analyze change—or income mobility—at the household-level (in addition to the usual population-level). The surveys cover a short period (7 years), but that was also a time of great economic change. The findings in a study by Ranganathan, Tripathi, and Pandey are generally negative.[30] Forward castes are of course heavily represented in the top income group (two to three times more heavily than their population weight) and Backward castes least represented. The progress of Forward castes up the income ladder is also the most rapid. There is income growth among the Dalit and Adivasi households too; close to one-third experienced upward mobility. But among all social groups studied in the paper, Dalits had the least upward mobility (30 percent of

[30] Thiagu Ranganathan, Amarnath Tripathi, and Ghanshyam Pandey, 2017, Income Mobility among Social Groups.

families) and most downward mobility (41 percent of families). Using the same IHDS data sets, Iversen, Krishna, and Sen find that there is "higher occupational mobility among forward castes than among SCs and STs…[and] a much higher prevalence of sharp descents among SC and ST sons."[31]

The wealth scenario (in Figure 3b) is even more stark and appears to have deteriorated sharply in the last decade. The Dalit and Adivasi share of national wealth had each been roughly half their population share till the early 2000's but dropped to 40 percent in the last debt and investment survey of the NSSO. The wealth share of the OBC also dropped from 90 to 80 percent of population share in the same time. In contrast, the wealth of the general (non-marginalized) population was 20 percent above its population share in 1991 and almost 90 percent above in 2012. *The general (non-Dalit non-Adivasi) population's wealth per capita in 2012 was almost five-fold higher than that of the Backward population*. The wealth gap between the Backward and non-marginalized populations was large to begin with and had roughly doubled in two decades.

It is important to remember that almost much of this "wealth" is notional rather than real; it is derived from land ownership and the assumed value of land. Hence, it is possible that what these data really reveal are differences between where people live—the marginalized on marginal/remote and less valuable land, the non-marginalized on more urban and generally more valuable land. It is also possible that since the NSSO has seriously undervalued urban land and has almost no account of the stock market, the wealth gap (notional or real) between the marginalized and non-marginalized is considerably higher than five-fold.

The graphic on the distribution of assets by religion shows the affluence of the small minorities (Jains, Sikhs, Christians) and the poverty of the large minority (Muslims). Not only do

[31] Vegard Iversen, Anirudh Krishna, Kunal Sen, 2017, Rags to Riches? Intergenerational Occupational Mobility in India.

the small minorities have significantly greater assets than average, but their shares grew over the decade 2002-12. Jains, already the wealthiest religious group by far, saw their asset share more than double in a decade, during the same time that Muslims saw their asset share shrink measurably.

Finally, we look at some information on educational attainment by social group. The reasoning is simple. Education is "the hard core of the 'hard core' of human capital."[32] It is the key to income generation, intergenerational mobility, and social status, not to mention citizenship and awareness of self and rights. If the educational gaps between social groups do not close, the material gaps between them will not either. Education and educational inequality in India are big subjects. Contributions to it range from articles in well-known journals like *Nature* to technical expert analyses. This minimal discussion here does no more than scratch the surface of a deep problem in which the major issues include access, quality, and cost (by social identity, location, and income class).

There is general agreement that some aspects of educational inequality have improved in the preceding decades. Notably, there have been big gains in literacy and school attendance among the young (including girl children) in all segments of society, and a general surge in college attendance (which nonetheless remains biased toward Forward castes). At the same time, many analysts recognize that the education market has become increasingly segmented, which means there are significant differences in quality (all "literates" are not the same, and neither are all college degrees) and that the Backward continue to fall behind in quality (even if they are catching up in quantity, having started from a very low base).

[32] G. S. Sahota, 1978, Theories of Personal Income Distribution: A Survey, *Journal of Economic Literature* 16:1-55, p. 12.

The graphics in Figure 3c highlight the significant differences in educational attainment by social identity among the adult population in India. There are vast differences between the most educated group (Brahmans) and the least educated (Adivasis): the former have twice as many years of education, are four-fold as likely to matriculate from school, and seven-fold more likely to hold a college degree. Dalits and Muslims are also very far behind Brahmans and other "high caste" groups. A snapshot of current college enrollees shows that, while some gaps may be closing, very large differences remain between social groups. Adivasis are still half as likely to be in college as non-marginalized groups, and the Muslim population is far behind, only one-fourth as likely to be in college as the non-marginalized Hindu groups.

These differences in averages exist all along the economic and social spectrum. A recent article in *The Economist* graphed the gaps between India's social groups on poverty and malnourishment. They called the gaps "unconscionable." In 2010-11, the poverty rates for Forward, OBC, Dalit, and Adivasi groups were respectively 12.5 percent, 20.7 percent, 29.4 percent, and 43 percent. Rural poverty was three- and two-times higher in the Adivasi and Dalit populations respectively compared to non-marginalized groups. Urban poverty was about three-times higher for both.[33] Malnutrition was almost twice as high for Adivasis compared to "upper" castes, and in the 1990's, had declined more slowly; that is, the gap was growing larger. In an innovative new paper,

[33] The Economist, 2018, Unconscionable: Low-caste Indians are better off than ever—but that's not saying much, Jan 25, Asia Edition. Poverty data from Arvind Panagariya and Vishal More, 2013, *Poverty by Social, Religious & Economic Groups in India and Its Largest States, 1993-94 to 2011-12*, Working Paper No. 2013-02, http://indianeconomy.columbia.edu/sites/default/files/working_papers/working_paper_2013-02-final.pdf; also see R, Radhakrishna, 2015, Well-being, Inequality, Poverty and Pathways out of Poverty in India, Economic & Political Weekly, Vol.50, No.41. Additional malnutrition findings from Michele Gragnolati, Meera Shekar, Monica Das Gupta, Caryn Bredenkamp and Yi-Kyoung Lee, 2005, *India's Undernourished Children: A Call for Reform and Action*, HNP Discussion Paper, World Bank.

Diane Coffey, Payal Hathi, Nidhi Khurana, and Amit Thorat document, among other issues, the extent of prejudice against Dalits—more than half their rural survey respondents (in Rajasthan and Uttar Pradesh) practiced untouchability and were in favor of having laws banning inter-caste marriages.[34] The numbers speak for themselves. No editorial commentary is needed.

[34] Diane Coffey, Payal Hathi, Nidhi Khurana, and Amit Thorat, 2018, Explicit Prejudice: Evidence from a New Survey, *Economic & Political Weekly* 53(1): 46-54.