



PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

Wealth Inequality, Class and Caste in India, 1951-2012

Nitin Kumar Bharti

Masters Thesis Report

Supervisor: Professor Thomas Piketty (PSE)

Referee : Professor Abhijit Banerjee (MIT)

June 28, 2018

Abstract

This research makes two main contributions. First, I combine data from wealth surveys (NSS-AIDIS) and millionaire lists to produce wealth inequality series for India over the 1961-2012 period. I find a strong rise in wealth concentration in recent decades, in line with recent research using income data. E.g. the top 10% wealth share rose from 45% in year 1981 to 68% in 2012, while the top 1% share rose from 27% to 41%. Next, I gather information from censuses and surveys (NSS AIDIS and consumption, IHDS, NFHS) in order to explore the changing relationship between class and caste in India and the mechanisms behind rising inequality. Assortative mating appears to be very high in India, both at the caste level and at the education level (though not hugely larger than than in Western countries at the education level). I stress the limits of our knowledge and indicate possible lines of future research, particularly regarding the interplay between assortative mating and inequality dynamics.

JEL Classification: J00 D63 N30

Keywords: Wealth; Inequality; India; Caste; Assortative Mating; Marriage

Contents

1	Introduction	5
2	Literature	6
3	Data	7
3.1	NSS- All India Debt and Investment Survey (AIDIS)	8
3.2	NSS- Consumption Surveys	10
3.3	IHDS- Indian Human Development Survey	10
3.4	NFHS- National Family and Health Survey	12
4	Demographic Profile	12
4.1	Population Share- Rural Urban divide	12
4.2	Population Share by Religion	13
4.3	Population Share by Caste	13
4.4	Economic Characteristics of different Castes	17
4.4.1	Economic Aspects	17
4.4.2	Education	17
5	Wealth Inequality Series	18
5.1	Concentration of Total Wealth	21
5.1.1	Two Contrasting Ends	21
5.1.2	Weak Base	22
5.1.3	Rich Topping	22
5.2	Representation of Different Castes by Decile	24
5.2.1	Rural	25
5.2.2	Urban	26
5.2.3	Discussion	27
5.3	Total Wealth Inequality within Caste	28
5.3.1	Top Deciles	28
5.3.2	Lower Deciles	29
5.3.3	Discussion	29
5.4	Total Land Inequality	30
5.4.1	Concentration of Land Within Land-Ownning Population	31
5.4.2	Discussion	32
6	Consumption Inequality Series	35
6.1	Concentration of Consumption	36
6.1.1	Two Contrasting Ends	36
6.1.2	Weak Base	37
6.1.3	Rich Topping	37
6.2	Representation of Different Castes-Decile wise	38
6.2.1	All India	38
6.2.2	Rural Urban Comparison	40
6.3	Within Caste Inequality	42
6.3.1	Top Deciles	42
6.3.2	Lower Deciles	43
7	Inter Caste Marriage	43
7.1	Assortative Matching: IHDS 2011 dataset	45
7.1.1	Educational Characteristics of Partners	46
7.1.2	Employment Characteristics of Partners	46
7.1.3	Correlation Coefficients	48
7.1.4	India France Comparison	50
8	Conclusion	50

9	Appendix	54
9.1	A.1 Figures	54
9.2	A.2 Tables	57
9.3	A.3 Methodologies	68
9.3.1	Stratification in NSS AIDIS	68
9.3.2	Converting from Household Level to Individual Level	69
9.3.3	Sampling Methodology in Surveys	69
9.3.4	Imputation of Wage Earnings	70
9.3.5	Correlation Coefficients	72

ACKNOWLEDGEMENT

I would like to thank my supervisor Professor Thomas Piketty for being very supportive and open to ideas. He was a constant source of motivation for whole year. Next I would like to thank Professor Abhijit Banerjee for agreeing to be referee. His suggestions were very useful, some of which I was able to incorporate. I would like to thank Lucas Chancel for helping me at crucial junctures by gaining access to datasets without which this data intensive thesis would not have been possible. I extend many thanks to Thomas Blanchet for helping me in technical parts of estimation. I thank Professor Pulak Ghosh for the important datasets which will certainly have a bigger role in my future research.

I thank a lot to Michelle Cherian, Kentaro Asai and Matthias Endres and for their help in proof reading my initial drafts of this thesis which played big role in finishing this report on time. Also I thank all my friends and family members for their moral support.

1 Introduction

Economic Inequality has become a major issue in the world. With more and more research in the area using different datasets, the gloomy picture of an extremely unequal society is becoming explicit, across the world. Almost all the countries have an increasing graph of income and wealth inequality during the recent few decades. It will be too optimistic to shrug it off as a co-incidence. Certainly there are some common factors around the world driving the level of haves and have-nots scenario in society. In this context it is very important to understand different phenomena which could potentially be the reason behind this. It is further interesting to look at the macro picture to understand why different countries, despite being in different development phases, all face an increasing inequality trend.

In the context of global economic inequality, I present the case of wealth inequality in India. I keep the domestic lens of caste active, to not miss the social inequality which might be propelling the uneven distribution of prosperity. I further divide my work here into two sub-topics. First is preparing the long term evolution of wealth inequality series for India. One of the challenges I faced, as any other researcher generally encounters, is the availability of data. A data-driven research carries more weight due to obvious reasons. Fortunately India has a long tradition of wealth surveys. Since 1951, every ten years, all-India surveys are conducted to know the wealth and debt situation in India. These surveys encompass different physical and financial form of assets. There are some concerns over the quality of data and inter-year comparability- which I have tried to overcome in this paper.

India is a very peculiar country with a complex and regressive caste system. In the study of wealth inequality one can't ignore this societal peculiarity which is historically older than wealth inequality. Thus, the second sub-topic is understanding economic inequality within the framework of caste. This paper is limited to presenting some simple ranking orders of average wealth and consumption, representing inequality by deciles within the caste-based framework (i.e. segregation of castes based on wealth, with higher castes concentrated more in top wealth deciles and lower castes concentrated more in lower deciles). Some may wonder if caste has any relevance today. Unfortunately past unequal distribution of wealth along caste lines has never been corrected. A very easy depiction comes from land distribution in India. Land, even today, forms a major part of wealth in India (Rural India- 65% and Urban India 45%). In the past land distribution was entirely based on caste where the upper-castes of society possessed almost all of the land and lower castes predominantly formed the working class. Colonial period concretized the possessions of land by distributing land titles and land ownership. Indeed, post independence India adopted land reforms aiming for the distribution of land ownership to the real farmers who were mainly from the lower castes. However, even going by the reports of the government this measure has met with partial success and the distribution did not go well beyond certain ranks of the castes. Early adoption of an egalitarian meritocratic system was accompanied by a support system in the form of positive affirmation (reservation/quotas) for the downtrodden castes. The first result of the caste-based inequality was that the situation of all the castes has improved over time, but there is no convergence. The rate of growth of forward/upper castes in terms of acquiring wealth or consumption is higher than the lower castes. In a situation of positive affirmation policies of the government, one would have expected the opposite. It hints towards the harsh reality of ongoing caste-based discrimination in the society.

The second extension of the inequality study is to understand the behaviour of consumption inequality. Consumption is sometimes considered as a correct measure for inequality as it includes the borrowings and transfers. Similar to wealth dataset, India has since 1951, consumption surveys which provides detailed consumption data. I combine the consumption, income and wealth inequality series and intend to do a detailed study in future. There is existing research on comparing the consumption inequality and income inequality trend. This extension is important as it has policy implications. Different trends in consumption, income and wealth inequality describes different nature of problems and hence requires different solutions.

The third important strand of study is estimation of Assortative mating in India and understanding its impact on inequality. Assortative mating implies marrying with a person of similar characteristics. A high level of education, employment and economic assortative mating can potentially impact current household level economic inequality and inter-generational inequality. The dynamics is still not fully deciphered. I fill an important gap in literature by estimating the level of assortative mating in India. The level is quite comparable to developed country like France even when India has a very high level of caste endogamy. Probably since the marriages are fixed by parents and families, there is a consistent effort to find similar characteristics between couples. The correlation between couple's own education has increased in the last 40 years- which might come from different reasons like change in preferences of people to marry with people from same educational background, or due to changing sex ratio and inequality. There is a possibility of reverse causality, i.e. income/wealth inequality might lead to assortative mating.

The paper is structured as follows. Chapter 2 connects this work to the existing literature on wealth inequality, assortative mating and their interactions. Chapter 3 outlines the various data sets used for the analysis. Chapter 4 presents the evolution of demographics of different social groups of India since 1901. Chapter 5 presents the evolution

of wealth distribution in India since 1961. Chapter 6 presents the evolution of consumption distribution in India since 1951. Chapter 7 elaborates on the caste and education level of assortative mating in India.

2 Literature

This paper produces long term wealth and consumption inequality series of India using survey data and correcting the top wealth distribution using the Forbes millionaires data. It complements the income inequality series produced by Banerjee (2005) and Chancel and Piketty (2017) on India. The basic approach is similar in terms of utilizing all the available resources to estimate the top wealth and consumption shares. There can be mainly 5 ways to estimate the distribution of wealth as outlined in Alvaredo, Atkinson, and Morelli (2016). First are Household Surveys of personal wealth. Second, the administrative data on individual estates at death. Kumar (2016) uses inheritance tax returns with mortality tables data with an estate multiplier technique to produce top .01% wealth share during 1961-85. Inheritance tax was active during 1953-85. This shows the limitation of this data to study recent wealth inequality. Third is the administrative data on wealth of the living derived from annual wealth taxes. Wealth tax prevailed in India during 1957-2016. The productive assets like shares, mutual funds and securities were exempted from wealth tax. Low collection of taxes and higher costs was one of the reason for its abolition.¹ Fourth is the administrative data on investment income, capitalized to get estimates of underlying wealth. Fifth is the list of rich individuals produced by different magazines like, Forbes Billionaire's list. This paper uses Forbes to correct the top shares from the survey.

Household surveys are useful to provide information for the bulk of the population, but suffer from the limitation of not capturing the complete top distribution- either due to non-responses, under-reporting or worse exclusion of the rich (Deaton, 2005; Jayadev, Motiram, and Vakulabharanam, 2007). To emphasize on the issue, I found the combined wealth of top 46 Forbes millionaires in 2012 is 2.5% of the total wealth from nationally representative wealth surveys. Hence I utilize the available Forbes list to correct the top series. In terms of methodology I use the advanced version of standard Pareto interpolation namely Generalized Pareto Interpolation techniques which has been developed in Blanchet, Fournier, and Piketty (2017) and has been used in Piketty, Yang, and Zucman (2017). I contribute in producing the wealth and consumption inequality series since 1960's. The combined income-consumption-wealth series on Indian inequality will help in analysing the economic inequality from the three prime economic indicators.

The income share of top 10% population shows an increasing trend since 1980 to reach 55% in 2013. Comparing this to wealth share, I find the top 10% share increasing from 45% in 1981 to 58% (pre-correction) and upto 68% post-correction in 2012. The consumption share of top 10% in this time period increased from 28% to 32%.

This paper supplements the existing literature Anand and Thampi (2016), Jayadev, Motiram, and Vakulabharanam (2007), and Zacharias and Vakulabharanam (2011) on wealth inequality in India and in other developing countries in general. The methodology adopted here is new and I compare the results from different methodologies in detail later. The second important point of demarcation is the comprehensive approach undertaken here by visualising wealth inequality not only from economic point of view, but also from social point of view. Gender, Caste and Place of residence are some other dimensions along which inequality has been scrutinized. The main motivation for this comprehensive approach is the peculiar form of Indian inequality where social and economic inequality are weaved together with the thread of caste.

Through this paper, I also add to the existing literature on inequality studies making caste as central social stratification of the society. Deshpande (2001) argued to make caste as an essential ingredient in the study of stratification patterns in India's population and later works Borooah (2005), Thorat (2002), and Zacharias and Vakulabharanam (2011) have highlighted the economic differences among castes which exactly matches with the caste hierarchy present in the society. One extension is through an attempt to disentangle the heterogeneous group of Forward caste² into finer categories, Brahmins, Rajputs, Bania and so on. Lack of any caste census in 1931-2010 and as well as lack of caste information from SECC 2011 (Socio-Economic Caste Census) is a big hindrance to any study along caste lines. It will surprise many that there is no concrete population figure of Other Backward Castes (OBC) in India. It is very recently that national representative surveys have started gathering information on castes- which helps in estimating their respective population. I initiated the process of cleaning, sorting and most importantly categorizing the castes. This is a gigantic effort as there are thousands of castes in Indian society with almost no documentation from recent times. I provide basic demographic and socio-economic characteristics for few well known castes. Lack of any prior attempt is an issue to compare the estimated numbers. One robustness check is to compare the estimated population

¹Improving ease of doing business, simplifying the tax rules etc. were some other reasons for its abolition.

²In this paper, I use Forward Caste to denote the Hindu religion population who do not fall under Scheduled Caste (SC), Scheduled Tribe (ST), Other Backward Castes (OBC). Forward Caste is synonymous to 'General' class which is used often in India. The basic difference from the point of view of Indian government is that SC, ST and OBC are positively discriminated through reservation/quota in employment and educational institutions.

from two completely different surveys (conducted by different organizations). I intend to carry out this task, as it has huge potential to analyse the population at finer level. The big social groups - SC, ST, OBC, FC are quite heterogeneous, as detailed later, and hence a finer level analysis will help in policy recommendation on reservations (positive affirmative actions) in India.

This paper also contributes towards the literature dealing with the measure of inequality and polarization in societies by empirically utilizing the theoretical concepts of representational inequality, sequence inequality and group inequality developed in Jayadev and Reddy (2011) on Indian wealth data. Representational Inequality measures a degree of “segregation” of between different social groups along any attribute space.³ Sequence Inequality measures the degree of “clustering” based on hierarchy among social groups. Group Inequality measures the portion of between inequality in the total inequality in a society. I perform Representational Inequality and Sequential Inequality analysis and find the existence of both which has been termed as ordinal polarisation. The population share of all the perceived lower caste groups (SC, ST and OBC) in top decile of wealth and consumption is lower than their overall population share. The situation has deteriorated over the last 40 years. Further a high level of within inequality is found in terms of wealth and consumption which has brief mention in Borooah (2005) also.

Taking a detour from the inequality series estimation, but remaining within the realm of inequality, I estimate the assortative mating in Indian couples. Assortative mating (AM) refers to the willingness to marry/cohabit with a person of similar attributes. AM is usually measured along dimensions like education, employment, education etc. (Cancian and Reed, 1999; Eika, Mogstad, and Zafar, 2014; Frmeaux and Lefranc, 2017; Greenwood et al., 2014; Kremer, 1996; SCHWARTZ, 2010). AM is a demographic phenomenon which has been observed for a long time now. The recent surge in studying AM in economics mainly comes from its potential impact on economic inequality. The debate on whether AM impacts economic inequality is still unsettled. A non-negligible effect of 3%-9% on measured household earnings inequality is estimated in French working couples (Frmeaux and Lefranc, 2017). The effect is higher at 10%-20% on potential household earnings. On the other hand Olivo-Villabril (2017) do not find impact of AM alone in US data and it only acts as amplifier of the underlying inequality in wages across educational groups. Greenwood et al. (2014) similarly concluded that for AM to impact HH level of inequality, married woman must work. If this is so, it could lead to higher inequality in India in future. Currently the labor force participation of women in India is much lower than developed world. Only 20-25% of women have some paid employment in the couples’ dataset. On the other hand the level of estimated AM is very high and comparable to France and US. If the labor force participation increases in future, it will fuel the household level income inequality and thereby will have impact on intergenerational mobility (Schwartz, 2013).

Talking of marriage in India, one can’t escape caste. Caste is the first checkpoint in Indian arranged marriages.⁴ Education, employment, wealth, earnings etc take secondary position to caste. The estimated Inter-caste marriages (ICM) in India is at 5-6% level (Ray, Chaudhuri, and Sahai, 2017). I find the ICM very persistent in society with change of 1 pp in 30 years. I anticipate the level of ICM to increase in near future with increased enrolment of women in higher education and/or increased labor force participation due to the interplay of demographic peculiarities and differential trajectory of educational attainment curve among different castes and sexes.

3 Data

There is neither any all-India level administrative household wealth data nor a caste census data publicly available, which makes the task challenging. Fortunately there are many all-India level large sample surveys to estimate statistics on wealth. Also under demographic particulars in surveys, some information on caste is available. Certainly surveys are not the best in capturing the information perfectly, but it is the best what can be used as of now. In this chapter I will describe in detail the different datasets I have used. Some datasets (like IHDS, NSS Consumption) are quite well known and have been well exploited by economists. I will briefly go through such datasets. However, some datasets are less known, like NSS-wealth datasets, which I describe in detail.

Since the study covers a long time period, several rounds of same surveys are used. Since these surveys are supposed to capture the same information, so at different points, one would expect them to be comparable. Unfortunately, often there are changed definitions, which makes a direct comparison difficult. I have made an effort to make different rounds of surveys comparable, but it is not possible in some cases. For example, there could be different type of biases coming from conduction of surveys by different agencies. One prime example in Indian NSS surveys is that on several occasions half of the sample is surveyed by agencies affiliated to central government and half by state government agencies. Due to differences in availability of resources the efforts in surveying can be different leading to some biases. Due to lack of

³Wealth and Consumption in this paper.

⁴Arrange marriage means marriage fixed by parents. Parents’ role in marriage in India is still intact. Maximum marriages in India are decided by parents/family members with no to very limited role of grooms and brides.

time, I have not tried to assess such differences and assumed the bias due to such differences to be negligent. This is usually an accepted practice in literature. All other changes which I make are presented below in different survey blocs.

Since wealth surveys are prone to differential level of under-reporting (with high level of under-reporting coming from land and gold asset holders) and high concentration of wealth distributions at top end, I re-estimate the distribution using the Forbes data (Jayadev, Motiram, and Vakulabharanam, 2007).

3.1 NSS- All India Debt and Investment Survey (AIDIS)

The NSS AIDIS are decennial surveys for the years 1961, 1971, 1981, 1991, 2002 and 2012. It is the primary data source for generating the wealth inequality series. The survey began in 1951 when the RBI (Reserve Bank of India) started All-India Rural Credit Survey with the main objective of identifying the demand and supply of credit in rural India to formulate banking policies and schemes. Information on assets and incidence of debt on rural households were collected to assess the demand side of credit. The next round of survey called All-India Rural Debt and Investment Survey (AIRDIS) was conducted in 1961-1962. This round of survey, for the first time collected details on financial assets, but the survey was still restricted to rural areas. Nationalisation of Banks in 1969 gave a new shift to credit policy, focussing on the hitherto untouched people involved in entrepreneurship, retail trade, professional works etc. and this led to decision to extend the survey to urban areas in the third round of survey in 1971-72. This round also saw an organisational change which brought into place NSSO (National Sample Survey Organisation) to conduct the survey. Unfortunately due to some sampling issues the urban data was never published.

Methodology:

The methodology adopted for the collection of data in NSS AIDIS is different from the well known NSS consumption surveys. Each selected household is visited twice, once during the first half of the survey period and then during the second half of the survey period. The surveys are conducted during the calendar year for the reference period which matches with agricultural year of the country.⁵ The asset and liability on a certain fixed date (reference date) is ascertained. The reference dates are mid-point of the reference period. Depending on different rounds of surveys, the methodology for ascertaining the asset value changed slightly.

In 1961-62, the asset and liability of the household is derived on a fixed reference date (i.e. 31st Dec 1961 which is mid-point of the reference period July 1961-June 1962) by ascertaining the stock of assets/liabilities as on the date of survey and their transactions/flow during the period of date of reference and the date of survey. In 1971-72 the survey directly collected for the fixed reference date (30th June 1971 and 30th June 1972). In 1981-82, the information on the stock and flow of assets was similar to 1961-62, with changes in valuation.

The details of stratification in different rounds of surveys are provided in Appendix (Stratification in NSS AIDIS). The important thing to note is that before the 1991 round of survey, households were stratified only on the basis of land possession (primary asset of household wealth). Later research found that the liabilities/debts were not captured properly due to stratification based only on assets. The correction was made from 1991 onwards in the stratification method using land possession and indebtedness status of households in rural areas. For urban areas, monthly per capita expenditure and indebtedness status were used.

Assets:

Assets include all items owned by the household which had some money value. The survey has collected the information on inventory of assets and liabilities on a fixed reference date. The definition of assets has gone through changes in different years. Physical and Financial assets are the two broad categories. In 1961-62 total assets included the value of 8 different types of assets- i) Ownership rights in land ii) Special rights in land iii) Buildings iv) Livestock v) Implements, Machinery, Transport Equipments etc vi) Durable household assets (life more than a year and which can't be purchased at a nominal price) vii) Dues receivables on loans advanced in cash/kind. viii) Financial assets-Government securities, National Plan savings certificates, shares etc. Importantly no information on cash was collected. In 1981-82, Agricultural implements and non-farm machineries including transport equipments were collected separately. However, crops standing in the fields, cash in hand or stock of commodities were not considered as assets. In 1991-92, non-farm business equipments and all transport equipments were collected separately. Unlike previous years there was an effort to collect information on cash in hand of the household. Similar definition and categorization was used in 2002 survey. Bullions and Ornaments were also collected as part of assets. In 2012-13, survey excluded some of the assets which it had collected before. Household durables were excluded on the pretext of valuation concern. Bullions and ornaments were also not collected.

⁵For e.g 1961-62 survey collected information from Jan-Dec 1962 for the reference year July 1961-June 1962. Agricultural year in India is from July to June. Similarly for 1971-72 survey, the survey was conducted during Jan-Dec 1972 for the reference period July 1971-June 1972. And so on for later years of 1981, 1991, 2002 and 2012

Valuation of Assets:

Estimating the total wealth requires not only the number/amount of assets but also their valuation. In 1961-62, all the values of the physical assets were evaluated using the fixed average market value prevalent at the time of the first round of the survey. Dues receivables were evaluated using average wholesale prices. Shares were valued at their paid-up value and all other financial assets were evaluated at their face values. There was a slight change in 1981-82 in the valuation. The value of physical assets owned on the reference date was simply the value on the date of survey minus the value of transaction between the reference date and the date of survey. Essentially it means different prices for the same assets were used.

Due to the lack of book value for valuation of assets for household sector, the following procedure is followed:

- i) Value of physical asset acquired prior to the 30th June 1981 (1991, 2001, 2011) was evaluated in its existing condition at the current market price prevailing in the locality on the date of survey if the asset is owned on the date of survey or on the date of disposal if the asset is disposed of during the reference period in a manner other than sale.
- ii) In case the asset is sold/purchased during the reference period, sale/cost price is considered as value of the asset. If the asset is acquired by way of construction, the expenditure incurred on construction is taken as its value.
- iii) If the asset is acquired other than purchase, then the value of the asset in its existing condition as prevailing in the locality at the time of acquisition is noted.
- iv) If the asset is acquired during the reference period and is also disposed of during the said period, the disposal value is reported

Availability of the data: Micro-individual survey files are available for the last three rounds of survey- 1991-92, 2002-03 and 2012-13. For previous surveys, only source of information is annual reports. The 1981-82 and 1971-72 reports are available⁶. The 1961-62 report is not digitised yet and is available in hard copy format.⁷ The samples are fairly big in size and vary for different years. The assets are collected at household level in all the surveys. I decided to work with individual level wealth as it provides a better understanding of wealth distribution from the perspective of inequality study. For example, a statement like top 10% of the households own 50% of wealth is a weaker statement to make in studying inequality if those 10% happened to represent 50% of total population. Household level inequality hides the per-capita level information, which is better suited for inequality statistics. Since there is no standard method to divide the wealth within household, I equally split among adult household. This will hold true for most of the physical wealth (like land, building, transport etc) which is synonymous to public goods within household. This procedure allows global comparison of inequality statistics. Equal split within a household is a big assumption in Indian society- where women usually do not own wealth because of various reasons like customary transfer of wealth from father to son, biased gender inheritance laws⁸ and general gender discrimination in India, but usually have a say (even if unequal) in managing at the household level. The social norm of primogeniture also leads to unequal share of wealth - major share of inheritance going to the eldest son. The definition of adult population is chosen as (> 20 years) for all the analysis.

In all the analysis related to wealth using AIDIS datasets, adult individual is the basic unit of reference. One of the issue is that the tabulated data for surveys are provided at household level. For converting the wealth values at individual level, I use decadal growth rate of household adult size from next round surveys, for which micro-data is available. The decadal rate of change in HH adult size in 1991-2001 is assumed to be same in 1981-1991 by decile. I use a correction factor- to make the estimated data consistent at macro level population. Full detail is given in Appendix "Converting from Household Level to Individual Level".

Justification for using total wealth instead of net wealth Further the distribution is estimated with household per capita total wealth and not with household per capita net wealth. Net wealth is total household wealth minus total household debt. The debt level from AIDIS are suspected to be unreliable due to reasons of strong tendency of under-reporting of liability, issues related to sampling methodology and relative increase in state sample compared to centre sample Chavan (2008) and Narayanan (1988). Comparing the debts handed out by commercial banks, cooperatives and other lending agencies, Gothoskar (1988) found under-estimation of 40% in 1971-72 round and 50% in 1981-82 round. The other reason to use total wealth is that for pre-1991 surveys the information is only available in tabulated form which restricts estimation of distribution of net wealth as the classification is by total asset ownership and not by net wealth status (Subramanian and Jayaraj, 2008). Due to these issues, I work with household total assets instead of household net assets.

⁶<http://www.mospi.gov.in/download-reports>

⁷The library of College of Agriculture Banking, Pune, has the report and it is accessible on prior appointments.

⁸Before Hindu Succession Amendment Act 2005, women were deprived of inheriting land and other properties.

3.2 NSS- Consumption Surveys

I use the well known quinquennial NSS surveys of Consumption for years: 1983, 1987, 1993, 1999, 2004 2009 for which micro-files are available to produce consumption decile shares. For previous years I use the estimation from Datt (1997). I intend to increase the series by using the micro-data available for later rounds of consumption survey.⁹ . It is only after 1993 that NSS consumption data has information on OBC.

I winsorize the data at lower end of consumption at 1 percentile level. It can be treated as assigning some minimum consumption/transfers to the households who claim to have zero or minimal level of monthly expenditure. For different years the minimum value varies from 0.2-0.4 times the Poverty line.¹⁰

The changing methodologies are present in different NSS annual reports and they have been well scrutinized in literature, Deaton et al., 2001; Sundaram and Tendulkar, 2001. One of the major debated issues is based on the reference period for different consumption items. Till 50th Round (i.e. till 1993) of survey, NSS consumption survey were based on Uniform 30 days reference period. Since 30 days recall period for food items was criticised for being too long to recall, it was decided to change the recall period. In 1999, the survey followed two recall periods- Mixed Reference Period, where some items like clothing, educational expenses etc. were asked for 365 days reference period and food items with two recall periods (7 days and 30 days). This led to contamination of the reporting. There is certainly a threat in inequality statistics too. For 1999 year, I use 30 days recall period, ignoring the contamination argument. Tarozzi (2002) has shown that the effect of questionnaire change on the 30-days reports for food¹¹ is not large. Hence while making any comparative analysis across caste groups it is important that changing the recall period should not impact different castes differently.

3.3 IHDS- Indian Human Development Survey

IHDS - is a nationally representative panel survey for the years 2005 and 2011. It is a joint venture of University of Maryland and National Council of Applied Economic Research (NCAER). The survey collects information on income which is not captured in other household surveys in India. Further it provides information on Brahmin caste- a subgroup within Forward caste- which is unavailable in NSS surveys. This paper uses the cross-sectional aspect of the dataset¹² for two purposes.

First to produce general socio-economic characteristics of different castes in India and perform a comparative analysis using the latest round of survey- IHDS 2011-12.¹³ The survey sampled 42,152 households with 2,04,568 individuals interviewing the head of the household to get this information. I create caste group based on religion and caste codes based on the following combination- a) Dalits (SC), Adivasis (ST) and OBC's (Other Backward Castes) are coded as they are, regardless of their religion; b) Next Brahmins are coded as Brahmins. c) Next all Hindus¹⁴ who are not categorized above are coded as Forward Caste (FC); d) Next, all the Muslims, who are not yet coded are coded as Muslims. e) All the rest of the population is grouped as Others.

I would like to highlight that this classification is different from the one given in IHDS. I chose this categorization, since SC, ST and OBC have some provisions of positive discrimination¹⁵ from central and state governments which directly impacts the educational and income outcomes.

Second I use the same year survey to estimate the level of Inter-Caste Marriages (ICM) in India and the level of economic and education level assortative matching. For this purpose I follow a combination of religion and caste code provided by IHDS. The difference from the above categorization is mainly on account of classifying "Muslims OBC" into Muslims instead of in OBC. Since the focus here is on marriage, and there is strong adherence towards within religion marriage, I use this categorization.

ICM is identified based on the answer from eligible women respondents to the question- "Is your family the same

⁹Post 2009, there has been one more round of consumption expenditure survey in 2011-12, Social Consumptions in India 2014 and Domestic tourism expenditure for the year 2014-15. The last two surveys are not targeted to collect detailed consumption as they used to before, but the survey asks question of monthly consumption per capita.

¹⁰ For example, the minimum consumption level in survey in different years at all-India level stands at: 2009- Rs. 306.2, 2004- Rs 187.6; 1999- Rs 176; 1993-94- Rs 93.25; 1987- Rs 48.3; 1983- Rs 24.5

¹¹Food forms the important component of expenditure in lower deciles

¹²I intend to use the panel structure of the dataset in my future research on understanding the evolution of different economic outcomes for different castes

¹³I refer IHDS 2011 in the whole document. The survey was conducted between Nov 2011-Oct 2012

¹⁴There are 5 major religions in India-Hinduism, Islamism, Christianity, Buddhism, Sikhism

¹⁵usually termed as reservation in India- as 15%, 7.5% and 27% of total seats are reserved for SC, ST and OBC respectively in all the government educational institutions and government employment positions. SC, ST and OBC are sometimes also referred as reserved castes.

caste as your natal family?” A negative response is categorized as Inter-Caste Marriage (ICM). The survey interviewed 39,523 ever-married women aged between 15-49 regarding health, education and marriage status. One of the issues to study ICM using this survey is that separate castes of husband and wife is not provided. There could be underestimation of ICM from IHDS dataset as inter-caste marriages are usually not disclosed in the society. National Family and Health Survey (NFHS) questionnaire asks for separate castes from husband and wives, however income information is missing. Also the social stigma towards inter-caste marriage can bias the response in NFHS too. Since one of the important objective is to estimate income level assortative matching in the society, I decided to use IHDS.

Assortative Mating (AM) analysis is performed on 34,713 couples. The age of couples are restricted to 15-60 years. All the retired, disables, young persons who can't work are not included in the analysis. It is important to highlight the different classifications of employment and education used in the analysis as they are ordered in nature. The AM analysis uses ordered categories, which induces some element of subjectivity.

Ordered categories of Employment: Two variables are used for the categories of occupation. First is Primary Activity Status, which identifies an activity as primary for the member of household. Retired, Student, Unemployed, Too young/Unfit and Others are removed. From the rest, I create 5 groups namely- 1) Salaried/Professional- combining Organized Business, Salaried and Profession. 2) Small Business/Artisan 3) Cultivators- Cultivators and Allied Agriculture 4) Non Agri Wage Labour and 5) Agri Wage Labor.

The second variable is based on Occupation, which uses National Classification of Occupation 1968 (NCO). which was asked from wage earners. One of the basic objective here is also to compare the estimates of AM in India with the level observed in France. There is an attempt to draw parallel with the 6 Professions and socio-professional categories (PCS) categories of INSEE France. The table outlining the comparison is provided in Table 43

The 7 created groups in decreasing hierarchy is:

- 1) “Professional” - comprising scientists, engineers, teachers, jurists etc.
- 2) “Admin, Exec, Managers”- elected and legislative officials, executive and managerial workers etc.
- 3) “Clerical”- clerical and supervisors, stenographers, conductors and guards etc.
- 4) “Sales” - merchants and shopkeepers, manufacturers, sales workers etc.
- 5) “Service Providers”- hotel keepers, maids, cooks, waiters, etc.
- 6) “Farmers, Cultivators”- Agricultural labourers, cultivators, fishermen etc.
- 7) “Labourers (Non-Agri)”- miners, tailors, carpenters, plumbers, construction workers etc.

Next, since occupation was available only for wage earners, I use the variable of primary occupation status to categorize the rest of the population in the above formed 7 occupational categories. For example, if occupation code is missing and primary activity status is cultivator I categorize the individual under 6 “Farmers, Cultivators” category. Organised Business (from Primary Status activity) are categorised under “Profession” as they are related to well established business (> 10 employees). Salaried are classified under clerical.¹⁶ Artisans and Independent workers are kept under “Sales”. This helped in increasing the sample size to carry out the analysis. There is a scope of improvement with re-categorization of occupation based on changed reality/perception in the society towards occupation. NCO 1968 code is old, but still used in IHDS 2011 survey- probably to keep consistency of panel data, which I follow here.

Ordered categories of education: Similarly for estimating education level AM, the categories are defined based on the questions related to completed years of education, highest degree obtained and whether the individual has even attended school. There are two sets of categories for education. One with 8 categories, namely-

- 0) “No Education” - zero years of education
- 1) “Less than Primary”- < 5 years of education
- 2) “5th Pass” - $\geq 5yrs$ and < 8yrs
- 3) “8th Pass” - $\geq 8yrs$ and < 10yrs
- 4) “Secondary” - $\geq 10yrs$ and < 12yrs
- 5) “H. Secondary and Diploma (< 3yrs)”¹⁷ - $\geq 12yrs$ and < 15yrs
- 6) “Bachelors (BA, Bsc, Diploma 3+)”¹⁸ - $\geq 15yrs$
- 7) “BTech, MBBS, MD, CA, PhD”¹⁹ - $\geq 15yrs$

The last two groups differentiate between elite and common higher educational degrees in education. The last group is synonymous to *Grand Ecole* of France, but since college information is not present, it is not perfectly comparable. The classification based on highest degree is imperfect.²⁰

¹⁶It is unclear why the salaried primary activity status have missing occupation status.

¹⁷Higher Secondary

¹⁸BA- Bachelors in Arts, Bsc- Bachelors in Science

¹⁹BTech- Bachelors in Technology, MBBS- Bachelor of Medicine, MD-Medical Degree, CA- Chartered Accountant

²⁰For example, A BA degree from St. Stephen's college will have more weight than a B. Tech from not well-known engineering college,

3.4 NFHS- National Family and Health Survey

Survey years 1992, 1998 and 2005 and 2015

4 Demographic Profile

India is the second most populous country after China and 7th largest in terms of area. India is full of diversity. India has 29 states with 1.3 billion population. The largest state of India has more than three times the population of France. It is one of the oldest civilisations in the world, and is the birth place of four existing religions of the world- Hinduism, Buddhism, Jainism and Sikhism. It also has the third largest Muslim population in the world. There are 22 national languages and a caste system with ~ 4000 castes. Religion, race, caste and language not only determine personal identity but also play an important role in many economic decisions, both at household level and at national level. It became a secular, democratic country and adopted a socialistic pattern of production after becoming an independent nation in 1947 and is now transitioning towards a capitalist system. All these demographic, social, political and economic factors have some contribution towards the existing “Indian form” of inequality.

This chapter outlines the evolution of basic demographic profile of India in 1901-2011. The hundred years time period provides a long term evolution of population structure. Understanding of inequality in India will be incomplete without taking into account its societal structure. The demographic patterns across different religion and social groups are estimated using survey data and the Census. Census was started in India in a regular manner from 1881 onwards. One has to be cautious while looking at the data from the pre-independence era. The British territory has seen a gradual expansion in India and some of the differences in population is due to different areas mapped under Census. The first attempt of conducting the census was between 1840-65 and it covered only the three big cities of that time- Bombay, Madras and Calcutta. Later at the time of independence of India in 1947, India was partitioned and two nations were formed out of it. The partition was along religious lines and resulted in a change of the population mix in terms of religion.

The total population of India grew from 253.9 million in 1881 to 1210.7 million in 2011. As per Thomson-Notestein-Blacker classical demographic transition, India was in Pre-transitional phase of very high crude birth rate (CBR) and crude death rate (CDR) at > 40 per thousand till 1920, and the population was almost stable. The small increasing population of Census data is mainly on account of increasing covered area. India entered into second phase of the Transition in 1920. The CDR started falling after 1920 but the CBR remained high till 1960's. The gap between CBR and CDR resulted in population growth. The life expectancy also increased from less than 20 years in 1911-20 to over 40 years during 1951-60. The decline in CBR started from 1960's (start of Middle transitional phase) and it dropped to 22.1 in 2011. The total fertility rate (TFR) reached just above replacement level at 2.4 in 2011. The period during 1970-2011, saw a consistent decline in CBR, but the difference between CBR and CDR resulted in population explosion during this phase. Now India is considered to be in the Late Transitional phase, where CDR has stabilised and CBR/TFR is continuously declining. The population of the country has increased 3.35 times its population in 1951. The rate of growth was lowest in 1921 and accelerated after 1951 with decadal increase of 20% during 1961-2001. It is only in last census decade that the rate of growth has declined to 17.7%

4.1 Population Share- Rural Urban divide

The urbanisation level in India has grown from a level of 10.8% in 1901 to 31.2% in 2011. It is still a rural dominated country and the urbanization level is much lower than what is experienced in developed countries. France US, UK all have urbanization level as high as 80%. The urbanization level is very low even compared to developing nations like- China (55%) South Africa (65%), Russia (74%). It is even lower than the urbanisation level in partitioned countries- Pakistan (38%), Bangladesh (34.3%). India still is a country of villages where more than 65% of the population resides. Table 1 provides the evolution of urban population share in the country in long run.

Table 1: Rural Urban Population Share

Census	Year	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991	2001	2011
Population %	Rural	89.16	89.71	88.82	88.01	86.14	82.71	82.03	80.09	76.66	74.28	72.17	68.84
	Urban	10.84	10.29	11.18	11.99	13.86	17.29	17.97	19.91	23.34	25.71	27.81	31.16

Source: Census data.

The share of Rural and Urban population in NSS consumption survey in 1983-2009 is provided in Fig. 26. We can notice a slight difference from the census figures. Urban area is slightly over-represented in the data. The definition

if applying to some non-technical positions in labor market.

of Urban area is the same as used in Census.²¹

4.2 Population Share by Religion

India has residents from all the religions. Islam and Christianity are imported religions to India and Buddhism, Jainism and Sikhism have their origin in India. The above mentioned 6 religions, makes up for more than 99% of Indian population Hindus are in majority in the country. One important remark on aboriginals or people from Schedule Tribe (ST) populations in the country is that before the census, they were categorized separately from Hindus. Post independence they are clubbed under Hindus. The estimate of ST population is provided later. They are captured in the census under Hindu religion. The population proportion is almost constant except a decline in Hindus share and increase in Muslims share.

Table 2: Religion-wise Population Share

Religion	Census Years																
	1840-65	1860-69	1867-76	1881	1891	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991	2001	2011
Aboriginals	0	0	0.28	2.53	3.23	2.92	3.28	3.09	2.36	6.58							
Hindus	71.67	75.96	72.93	74.02	72.32	70.37	69.4	68.56	68.24	65.93	84.1	83.45	82.7	82.3	81.53	80.5	79.8
Muslims	19.08	16.31	21.39	19.74	19.96	21.22	21.26	21.74	22.16	23.81	9.8	10.69	11.2	11.8	12.61	13.43	14.23
Christians	5.36	0.51	0.47	0.73	0.8	0.99	1.24	1.5	1.8	1.63	2.3	2.44	2.6	2.44	2.32	2.34	2.3
Sikhs	0	1.1	0.61	0.73	0.66	0.75	0.96	1.02	1.24	1.47	1.79	1.79	1.9	1.92	1.94	1.87	1.72
Others	3.9	6.13	4.32	2.24	3.03	3.75	3.87	4.08	4.21	0.57	2.01	1.63	1.6	1.59	1.6	1.86	1.95
Total %	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Total Population (mill)	1.7	103.1	191.1	253.9	287.2	294.4	313.5	316.1	350.5	386.7	361.1	439.2	548.2	683.3	846.4	1028.7	1210.7

Source: <http://dsal.uchicago.edu/copyright.html>.

4.3 Population Share by Caste

Define caste: Caste is a system prevalent in India society which can be thought of as a practised version of varna system enumerated in religious scriptures. There is enormous literature on caste in sociology and anthropology (Kannabiran, 2009). I will describe certain essential components of caste system which still have relevance today. I define in detail some features of the caste system.

First- Caste system is not only a Hindu religion phenomena anymore though it indeed has its beginning in Hinduism/Brahmanism way of living. It has now entered into all the religions in India, in one or other form. It is believed that Buddhism and Jainism, the two ancient religions of India were born in around 5th century B.C against the backdrop of high levels of social inequality that generated from discriminatory norms of Brahmanism of that time. With time these religions themselves acquired the prejudiced traits of Brahmanism. The decline of Buddhism in its own birthplace is often associated to this reason.

Second- caste system is not varna system. Varna system has 4 (or 5) categories namely Brahman (priests, scholars and teachers), Kshatriya (rulers, warriors and administrators), Vaishya (agriculturists, merchants) and Shudra (workers / small farmers / service providers), and it has its origin in Dharamshastras²². The fifth category consisted of Untouchables. The first three categories were allowed to wear the sacred thread and study religious texts. Even though caste system is believed to have come from varna system, both of them are not identical. Today society associates more with caste (or jatis in Hindi), which are thousands in numbers and vary regionally. (Srinivas, 1955) says that these thousands of castes can be clustered around varnas. This makes comprehension easier.

Third- Caste system has broad acceptance in the society. It is not a menace in the opinions of majority of the population. It acts as an identity and beyond that it provides some of the benefits which the welfare state provides (Srinivas 2004). For all the practical purposes, there is very less competition inside this system to outrank others, because caste is determined by birth. The top-most rank is reserved forever for the perceived creators of the system - Brahmins.

Fourth- Caste system provides a form of social status to every caste. Indeed, as with any form of status, it comes at some cost. Some castes derive higher status tag at the expense to lower status tag to others. However, it is more generous in the sense that the higher status has been distributed to fairly large proportion of population. 5% Brahmins and around 5-7% Rajputs (from Kshatriya of varna), Bania (2%) are considered higher castes. It could be surmised that the most important and unique factor which provides sturdiness to the caste system is the fact that it assigns relative position to every caste in the society. Even though Rajputs come second in hierarchy, they are on top of 90% of the population. Similarly Banias derives utility by being higher in status to 88% of the population. And this way

²¹The definition of Urban area depends on 3 criteria namely- 1) Population threshold (≥ 5000) and density (400 sq.km) 2) Proportion of men in agriculture sector (75%) 3. Presence of Municipality/Municipal Corporation.

²² Manusmriti- the first one to be translated in English by Britishers in 1784 to rule Hindus in India. I wonder the impact of this colonial act in reinforcement of caste system in the form it is today.

the caste system has created a situation where all the castes (except the ones at bottom) derive some kind of superior social status compared to other castes.

Overall the debate on caste system has remained contentious. The most famous debate in recent history is the Gandhi-Ambedkar debate on caste. Gandhi (considered as father of the nation) was against the discriminatory nature of caste system. He coined the term Harijans (in Hindi it means-children of God) for untouchables and worked throughout his life for their upliftment. However, he was not against the caste system itself. According to him, division of labour is inevitable in society and there was no need to overthrow the caste system which was based on division of labour. On the other hand, B. R. Ambedkar (architect of Indian Constitution) was against the stratification of society based on caste. He made a distinction between division of labour and division of labourers, where the first one is choice-based and the later imposed by the caste system. He said that outcaste is a product of caste and unless caste system is uprooted from its base, the discrimination based on caste will remain. Hence he demanded complete overthrowing of the system. In principle, I would agree with the Ambedkar stance on the caste system, however one cannot ignore the reality of the society. The difficulty to overthrow the system is its huge approval in society. One of the peculiar aspects of the caste system, which has become the strongest pillar of its existence, is caste endogamy. Everyone is supposed to marry within one's own caste. This paper finds strong evidence of caste homogamy in the society, where 94% of total marriages are within-caste marriages.

Caste became part of census in full fledged manner in 1901 census. The historical census data at caste level is available online at Digital South Asia Library (University of Chicago)²³. Table 3 provides top 20 castes which existed in 1901 and 1911.²⁴The last census which gathered caste data and is available in public was in 1931. Indian government wanted to stop caste-based census due to the fear of re-enforcing the caste identity. However provision of affirmative actions were made for Scheduled castes (Dalits) and Scheduled Tribes (ST) since 1951. Scheduled caste (SC) was not a single caste entity but a list of castes. The people belonging to castes in SC were considered untouchables during that time. They formed the lowest rung in the caste hierarchy. For the implementation of the clauses of positive affirmation enshrined in the constitution, census gathered information on SC and ST. Later in 1990, India extended the caste-based positive affirmative actions (also referred as reservation policies) for Other Socio-Economic Backward Class of the society. These are the population often coming from Other Backward Castes (OBC) castes. Again OBC is not a single caste but a list of thousands of castes. When the reservation policies was promulgated for OBC's, there was a major problem in estimating the population of OBC's, because there was no census information. OBC population was estimated from 1931 census using the population growth. Facing the problem of lack of information on castes, Indian government conducted the Socio-Economic Caste Census in 2011 to update the knowledge on different castes. However, unfortunately the data has not been made public.

²³<http://dsal.uchicago.edu/copyright.html>

²⁴The data of caste for 1921 and 1931 census is much elaborate and is not in csv format. I am working to handle it

Table 3: Population Share - Census

1901			1911		
Caste	Population Size	Population %age	Caste	Population Size	Population %age
Shekh	28,708,706	9.16	Sheikh	32,131,342	10.25
Brahman	14,893,258	4.75	Brahman	14,598,708	4.66
Chamar	11,137,362	3.55	Chamar	11,493,733	3.67
Ahir	9,806,475	3.13	Ahir	9,508,486	3.03
Rajput	9,712,156	3.1	Rajput	9,430,095	3.01
Jat	7,086,098	2.26	Burmese	7,644,310	2.44
Burmese	6,511,703	2.08	Jat	6,964,286	2.22
Maratha	5,009,024	1.6	Maratha	5,087,436	1.62
Teli and Tili	4,025,660	1.28	Kunbi	4,512,737	1.44
Kurmi	3,873,560	1.24	Teli and Tili	4,233,250	1.35
Kunbi	3,704,576	1.18	Pathan	3,796,816	1.21
Pathan	3,404,701	1.09	Kurmi	3,735,651	1.19
Kumhar	3,376,318	1.08	Kumhar	3,424,815	1.09
Kapu	3,070,206	0.98	Kapu	3,361,621	1.07
Napit	2,958,722	0.94	Mahar	3,342,680	1.07
Mahar	2,928,666	0.93	Koli	3,171,978	1.01
Jolaha	2,907,687	0.93	Hajjam	3,013,399	0.96
Bania	2,898,126	0.92	Lingayat	2,976,293	0.95
Kaibartha	2,694,329	0.86	Gond	2,917,950	0.93
Lingait	2,612,346	0.83	Jolaha	2,858,399	0.91
Koli	2,574,213	0.82	Palli	2,828,792	0.9

Source: <http://dsal.uchicago.edu/copyright.html>.

As per Census data²⁵ the proportion of SC in population has increased from 14.67% in 1961 to 16.6% in 2011. During the same time period the proportion of ST has increased from 6.23% to 8.6%. Apart from natural growth rate of SC and ST, classification of more castes/tribes into SC/ST²⁶ has contributed in the increase of their population share. Apart from the Census, surveys too provide valuable information on the estimates of population. The Census is conducted once in 10 years and the NSS consumption surveys are conducted almost every 5 years. Hence NSS surveys in a way complement the Census and can be utilised to study mid-year variations in the population profile. Post 1999 NSS surveys allow the breakdown of the population into OBC also. I categorize the population based on a combination of religion and caste which has been mentioned in a section of the IHDS dataset. The same definition has been used in NSS also.²⁷ Fig. 1a depicts the evolution of population share of SC, ST, OBC and FC since 1983. We can see that SC and ST are slightly over-represented in the survey (SC- 20% and ST -10%). We also observe an increasing share of OBC population share and corresponding decreasing share in population of FC. Non-Hindu population share declined in 1999, because of the creation of the OBC category. OBC Muslims are categorized under OBC.

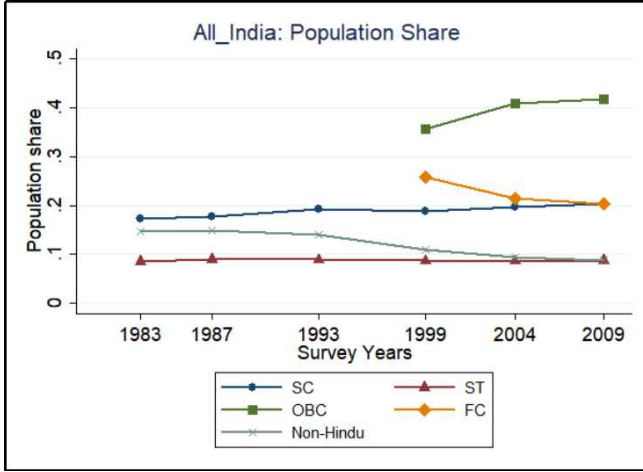
²⁵<http://censusindia.gov.in/>

²⁶The recent most addition is in 2015. <http://pib.nic.in/newsite/PrintRelease.aspx?relid=115783>

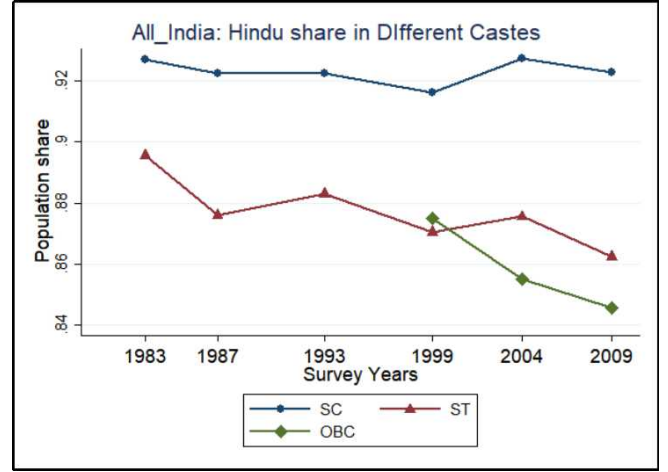
²⁷1. SC(Scheduled Caste) 2. ST(Scheduled Tribe) 3. OBC(Other Backward Castes) 4. FC(Forward Castes-Hindus who are not SC/ST/OBC) 5. Non-Hindus

Figure 1: Population Share from NSS

(a) Proportion of SC/ST/OBC/FC



(b) Share of Hindu in different castes



The share of different caste groups in Rural, Urban sector separately is provided in appendix Fig. 27. We observe that FC community is disproportionately concentrated in Urban areas and STs are mainly confined to Rural areas.

The second subgraph 1b shows the proportion of Hindus in SC, ST and OBC. It is important to re-iterate here that SC, ST, OBC which are reserved communities in India (with positive affirmation) are not restricted to just the Hindu population. Caste has intruded into other religions also. We see an increase of other-religions in ST and OBC. In 1983, Hindus comprised almost 90% of ST which dropped to 86% in 2009, whereas for the OBC the share of Hindu population dropped from 88% to 84%.

IHDS datasets allows to split FC category into Brahmins and rest of the FC, which help in estimating the population share of Brahmins ($\sim 4.86\%$) and some of their general demographic and socio-economic characteristics. Table 19 provides the percentage of population in different caste groups. The information on Brahmin community which is considered as the highest caste in the caste hierarchy makes this table interesting. As there is no census on castes, surveys are the only source to gauge their population share or other characteristics. Another important observation is that child and young population among SC, ST and OBC is higher than their respective population, whereas adult and old population of FC is higher. This hints towards decreasing fertility level of FC.

Another nationally representative dataset namely, NFHS(National and Family Health Survey) allows to further split the FC into different castes.²⁸ Table 20 provides the population shares. I split FC into Brahmins (4.7%), Rajputs (proxy for Kshatriya- including Marathas 4.9%), Bania (2.1%), Kayasth (.6%) and Others (9.3%). The population share of Brahmins here is close to what we get from IHDS dataset- which is a good sign for the categorization. The observation on fertility holds for other FC groups too. They have higher proportion of adults and are older compared to their overall population share.²⁹

Comparing with 1901,1911 Census we see almost same proportion of Brahmins after almost 100 years! Rajput+Maratha population proportion is also very close to what it was a century ago. Bania population is higher in NFHS, but it could simply be due to categorization of more castes into this category. Business as a profession has been adopted by other castes and those castes have now been associated with Bania community.

The next few figures in the Appendix (28 and 29 show the proportion of SC/ST/OBC in different religions in surveys. Pre-1993 surveys have no categorisation such as the OBC. In Hindus the share of SC and ST is almost stable at around 9.5% and 19-21% respectively and they are slightly over-represented in the survey compared to Census estimation. For later rounds we see in 1999, OBC's formed 38%, 30%, 20%,13.6% in Hindu, Muslims, Christians and Sikhs respectively which increased to 43%,43%,25%,21% in 2009 in the same order. The increase in OBC population is in all religion which is chiefly is due to reclassification of more castes into OBC. The potential demographic reasons like - growth rate, migration etc. usually would take longer time to translate into such big changes.

²⁸Currently the categorization is based on google searches, crowdsourcing etc. It is prone to biases and hence I am working on basing these categorization on some historical authentic sources.

²⁹All the castes categorized into Brahmins, Rajputs, Bania and Kayasths are categorized is given in appendix.

4.4 Economic Characteristics of different Castes

IHDS and NFHS datasets are utilized to compare the socio-economic characteristics of different castes. IHDS datasets has the additional benefit of being a panel dataset to understand changing conditions over 20 years. Also it is the only dataset which captures income. I intend to perform a panel data analysis in my future research. However, for the purpose of this paper, I restrict my analysis to a cross-sectional one.

4.4.1 Economic Aspects

Table 4 shows the mean values of basic economic and educational characteristics of different caste groups. The values are given at both household and per-capita level. It provides a clear picture of stratification in the society. The second last column “Others” contains the rest of the population which belongs to the Non-Hindu, Non-Muslim religion group and does not fall under SC/ST/OBC. This is a rich minority cluster (with only $\sim 1.5\%$ population) in terms of all the economic and educational parameters. The last column shows the all-India statistics.

Income: Average annual income of an household in India is Rs.113,222(Rs. $\sim 9,435$ per month). The annual income of ST and SC group stands at 0.7 times and 0.8 times lower than the all-India average income. OBC and Muslims both have around 0.9 times household income of the overall average income. Forward castes (FC), have average household income at 1.4 times the all-India income (with a slight difference between Brahmin and Non-Brahmin). There is sequential inequality (SI) based on average income with ranking $ST < SC < Muslim < OBC < \text{OVERALL} < FC(Non - Brahmin) < FC(Brahmin) < Others$ trend. We see a similar trend at per-capita level of annual income. The standard deviation which is not present in this table is highest for Others followed by FC(Non-Brahmin), FC(Brahmin), SC, Muslim, OBC and ST.

Wealth: Wealth/Assets in IHDS is simply the average of indicator (=1) of the presence of for 33 different household durable goods like TV, Air Conditioning, 4-wheeler etc. It has a range of 0-33 and is defined at HH Level. The SI based on assets across different caste groups is visible with ranking $ST < SC < Muslim < \text{OVERALL} < OBC < FC(Non - Brahmin) < FC(Brahmin) < Others$, which is similar to income trends.

Consumption: The consumption data shows similar pattern with OBC and Muslim groups near the all-India average, SC and ST below average and both Brahmins and non-Brahmins performing better.

Table 4: Wealth,Income and Consumption Mean Values-IHDS 2011

	SC	ST	OBC	FC(Brahmin)	FC(Non-Brahmin)	Muslim	Others	OVERALL
Annual Income of HH (in Rs)	89,356	75,216	104,099	167,013	164,633	105,538	242,708	113,222
Per Capita Annual Income (in Rs)	19,032	16,401	21,546	35,303	36,060	20,046	56,048	23,798
Annual Consumption of HH (in Rs)	87,985	72,732	108,722	146,037	143,497	102,797	181,546	109,216
Per Capita Annual Consumption (in Rs)	18,740	15,860	22,503	30,869	31,430	19,525	41,924	22,956
ASSETS	12.7	10.2	14.7	18.2	17.9	13.3	22.2	14.6
ASSETS2005	12.2	9.9	14.2	17.5	17.1	12.9	21.2	14.1
Highest Adult Education	6.7	5.9	7.8	11.5	10.3	6.6	11.6	8.0
Highest Male Education	6.3	5.6	7.5	11.3	9.9	6.0	10.5	7.6
Highest Female Education	3.9	3.3	5.0	8.6	7.8	4.6	10.0	5.3

Source: Author's calculation using NFHS IHDS 2011 datasets. Design weights are used to estimate these values.

Using NFHS 2005, we can compare between more caste groups in FC. Table 22 in Appendix presents the wealth index provided in the dataset, categorizing the population in 5 brackets from Poorest to Richest. We see that 50% of the Brahmin, 31% of Rajputs, 44% of Bania and 57% of Kayasth fall in richest category. For other caste groups only 5% ST, 10% SC, 16% OBC, 17% Muslims fall in richest category.

I find a high level of sequential inequality in society, which implies clustering of social groups. FC is clustered more towards higher income/consumption/wealth values. OBC and Muslims are in the middle and SC/ST cluster towards the lower end. This clustering of different groups is parallel to the caste hierarchy with FC the highest caste and SC/ST the lowest in hierarchy. Economic ranking follows caste hierarchy, making caste a valid stratification in the society. It is important to keep in mind that the standard deviation is also very high in FC groups, i.e. not all are well off in that group. The clustering of social groups is not perfect.

4.4.2 Education

The last three rows of Table 4 depict the highest level of education for adults, males and females among different caste groups.³⁰ Brahmins have highest adult education followed by Others and FC (Non-Brahmins). The comparison for adult education across different community gives- $ST < Muslim < SC < \text{OVERALL} < OBC < FC(Non - Brahmin) < FC(Brahmin) < Others$ The difference between Brahmin and ST is 5.6 years, Brahmin and Muslims is 4.9 years, Brahmin and SC is 4.8 years, Brahmin and OBC is 3.7 years and Brahmin and Other FC is 1.2

³⁰0-16 with 0 denoting no education, 12- Higher Secondary 15- Bachelors, 16 above Bachelors.

years. Muslims who are almost near the overall average in economic parameters fall behind the SC group in terms of education.

Taking note of the highest male and highest female education pattern separately, we observe same trend among caste groups as we saw for the highest level of adult education. This shows the different trajectory of education curves among different caste groups. Another important point to note is that among all the social groups, women have lower mean education than men. There is a **secular gender discrimination** in the society. This shows a different trajectory of education curves between men and women in every social group. The interaction of different trajectories of education curves among social groups and among sexes are potentially a crucial determinant in shaping up inter-caste and inter-religion marriages. I explore this in detail in chapter on Assortative matching. Comparing among females, FC (Brahmins and non-Brahmin) communities have the highest average years of education, implying these caste groups are sending more of their girls towards schools and colleges.

Using NFHS 2005 I compare the education level of adult in other Forward castes [Appendix Table 23]. Kayasth with 12.3 years of education is highest, followed by Brahmins (11.9years), Bania (10.3 years), Rest of FC(9.16yrs) and Rajput (9.05 years). The qualitative comparison is same as found in IHDS datasets, though there is slight change in average years of education.

Within SC/ST/OBC Analysis: I look within the SC/ST/OBC category to see if there are differences across different religion groups. The table 5 presents the relative performance of the groups with respect to all-India average values. We can see that within ST, Christians have 1.6 times income and assets than all-India average and their educational level is better than many other groups. Muslim ST's economic parameters are closer to all-India average but education wise they are behind. Hindus and Other/No religion ST's which forms 78% and 12% of all ST's are the worst performing groups.

In the SC group, Hindu (93%) and Others (6%) are two major groups and we see that SC's from other religions have outcomes at par with the all-India level. Hindu SC's have worse outcomes and their mean assets have declined from 2005, which is a worrying issue.

In OBC, the two big groups are Hindu (82%) and Muslim (16%). Hindu OBC's have better educational averages than Muslim OBC's. On the economic scale they look almost the same. The small group of other religion in OBC's are in no way backward.

Table 5: Within SC/ST/OBC Relative to all-India: IHDS 2011

	ST				SC			OBC		
	Hindu_ST	Muslim_ST	Christian_ST	Others_ST	Hindu_SC	Muslim_SC	Others_SC	Hindu_OBC	Muslim_OBC	Others_OBC
Annual Mean Income of HH (in Rs)	0.58	0.97	1.6	0.54	0.78	0.81	0.9	0.91	0.88	1.38
Per Capita Mean Annual Income (in Rs)	0.58	0.86	1.69	0.67	0.78	0.75	0.97	0.92	0.74	1.51
Annual Mean Consumption of HH (in Rs)	0.63	1.27	1.29	0.44	0.8	0.99	0.86	0.97	1.09	1.26
Per Capita Mean Annual Consumption (in Rs)	0.64	1.12	1.37	0.55	0.8	0.92	0.93	0.98	0.92	1.39
ASSETS	-4.9	-1.21	1.71	-5.82	-2.2	-1.08	0.9	-0.08	0.43	4.9
ASSETS2005	-4.66	-1.07	1.39	-5.46	-2.1	-0.88	0.85	-0.09	0.46	4.48
Highest Adult Education	-2.36	-1.41	1.8	-2.99	-1.41	-2.13	0.25	-0.08	-1.13	2.05
Highest Male Education	-2.19	-0.48	1.68	-2.98	-1.35	-2.26	-0.22	0.07	-1.2	1.48
Highest Female Education	-2.3	-1.24	2.4	-2.86	-1.51	-1.18	0.73	-0.28	-1	3.07

5 Wealth Inequality Series

This chapter illustrates the picture of household wealth in India. The core of the chapter deals with estimating the distribution of wealth in population. Instead of using gini coefficient as a single measure of inequality, the emphasis is on estimating different shares of wealth in top 10%, middle 40% and bottom 50% of the population (Piketty, 2014). This distribution of wealth is also examined among different caste groups. Previous chapter has already established the stratification in society along castes from simple economic parameters. In this chapter, the focus is entirely on wealth, which is associated with welfare, a tool for consumption smoothing, a source of power and contribution in total income through asset income (rent). Micro-files allow for more detailed analysis, like age and education profile of different strata of the society.

I have made an effort to look into the issue of wealth inequality from social point of view. Distribution of land is analysed in detail. Land happens to be the most important asset of Indian households in monetary value. India has tried in the past to achieve an equitable distribution of land with various land reform measures.³¹ Today the slogan of "land to the tiller" has died down in India partly due to decline in the number of big land sizes (population explosion has a major role in division of land plots, long drawn court cases by landowners to stall the land reforms etc.) and

³¹Zamindari Abolition Act 1954, Land Ceiling Act, Tenancy reform act etc.

partly because of arrival of new opportunities in manufacturing and service sectors. Unless India develops different forms of assets, the issue of unequal distribution of land can't be neglected. The distribution of wealth in different social/caste groups has been studied with an attempt to characterize the evolution of the pattern. Representational Inequality-i.e. lower presence of a certain caste group than its proportionate share in the population has been exploited in Section 5.2, apart from the top share statistics. Both across-castes and within-caste inequality have been examined in detail. The dichotomy between rural and urban sectors has been highlighted across all the outcomes.

The methodology in estimating the top share is still new and evolving. The limitation of surveys in capturing top wealth holders is dealt in two ways- 1) By using Forbes millionaires data and making an assumption of Pareto distribution at top percentiles. (Blanchet, 2016; Garbinti, Goupille-Lebret, and Piketty, 2017) 2) A novel attempt to re-estimate the inequality in a selected urban area³²

The results are accompanied by discussions which is an attempt to look for the reasons in past and current government's economic and social policies. The trends and evolution has been analyzed in detail. The discussions part in every section mentions potential future research.

The chapter begins with providing a description of the contribution of different assets in the total household wealth at all-India, Rural and Urban sector separately 6.

Table 6: Contribution of different assets in total Wealth

Contribution of Different Assets in Wealth (excluding Durable HH's)														
Type of Assets/Year	Rural						Urban				All-India			
	1961	1971	1981	1991	2002	2012	1981	1991	2002	2012	1981	1991	2002	2012
Land	66.49	69.92	66.92	68.12	65.00	69.84	38.21	40.04	41.62	45.26	60.06	59.33	56.59	56.36
Building	22.52	18.76	22.31	22.69	24.28	20.33	41.98	44.34	40.92	43.24	27.00	29.47	30.26	32.89
Livestock	8.00	6.81	5.39	3.58	4.35	1.55	0.94	0.48	0.47	0.10	4.33	2.61	2.95	0.75
Agricultural Machinery, Equipments and Transportation	2.99	2.83	3.99	3.99	3.99	2.70	5.90	5.36	6.76	3.17	4.44	4.42	4.99	2.96
Financial Assets	-	1.15	1.29	1.56	2.28	5.46	12.50	9.30	9.94	7.96	3.97	3.98	5.03	6.83
Loan Receivables		0.52	0.11	0.06	0.10	0.12	0.47	0.49	0.29	0.28	0.19	0.19	0.17	0.21

Looking at all-India verticals, we see that *Land* value share is the highest but is on a gradual declining trend. The share has declined from 60% in 1981 to 56% in 2012. Rural sector expectedly has land as the most valued asset and it has almost remained at around $67\% \pm 3\%$ since 1961. This hints towards continued existence of lower material wealth in rural sector. In Urban sector, land share is lower than all-India level but it shows an increasing trend from 38% in 1981 to 45% in 2012.

Building is the second most important asset in the wealth. It is surprising to note that Land and Building together form 90% of the total household wealth. The decline in land share is accompanied by increase in the Building value share. In Rural area, the share of building has hovered around $20\% \pm 2\%$ except in 2002 when it is at 24%. The rural village morphology allots only a small share of land area towards living and common space³³ Most of the land is dedicated to cultivation. In Urban area, building share is almost double than that of rural area at around $42\% \pm 2\%$. The stagnancy in the building share might look surprising as in many cities the advancement of apartment culture is accompanied by high rise in building prices. One possible explanation is almost equivalent increase in land prices which has not let the building share in total wealth rise.

Financial Assets has the third largest share and it has increased to almost 7% level in 2012 from 4% in 1981. This is in contrast to high level of financial assets in developed countries. For example in France and US the share of financial assets stood at 30.9% and 48.3% in 1979.³⁴ There is a declining trend in the share of financial assets in Urban areas which is surprising. To some extent the decline can be explained by the change of definition of financial assets (explained in data section) in different rounds of survey. Non-reporting/non-inclusion of rich households could be another potential reason.

The supremacy of physical assets over financial assets is undoubted. Financial assets is a very small small part of total household wealth.

Agriculture Machinery, Non-Farm Machinery and Transport Equipements comes next in the descending list of contribution to total wealth. It's share has declined from 4.4% in 1981 to 2.96% at country level. In Rural area this component remained stable at almost 4% for 3 decades before dropping to 2.7% in last survey. Urban area, has seen a decline from 5.36% to 3.17%.

Livestock used to be around 3-4% of total wealth till 2002 before dropping to 0.75% in 2012. This decline is solely attributed to Rural areas where we see a decline of 3 percentage points. Interestingly in 1961, the share of livestock was more than the share of financial assets in 2012. The Livestock survey of NSS in 2012, found a decline of 26 millions (around 10% of 2002-03 level) bovines in 2012-13. This could be the reason for low share of livestock wealth.

³²I am thankful for this idea to Professor Abhijit Banerjee.

³³The rural roads inside villages in Northern India are kept so thin that sometimes passing two cycles from opposite directions at one time is not possible.

³⁴Kessler and Wolff, 1991.

However, livestock from only those households which had some operational holdings of land³⁵ were surveyed which could potentially driving this decline.

The evolution of average per capita wealth (APCA) is given in Table 7.³⁶ We can see APCA is increasing since 1961 and the increase is faster in recent years. In Rural areas, till 2002 there is only a modest increase from Rs.50.8k in 1961 to around Rs.180k in 2002. This figure jumped to Rs.390k in 2012 which is $\sim 117\%$ growth in a decade or $\sim 7.5\%$ annual growth rate from 2002-12. Similarly in Urban area, we see a steep increase after 2002. The APCA in Urban area has increased from 272k in 2002 to 904k in 2012, implying an increase of 232%, or $\sim 12.7\%$ annual growth rate. The Urban-Rural gap in APCA is consistently increasing since 1981. The ratio of Urban APCA to Rural APCA has increased to 2.32 in 2012 from 1.25 in 1981.

Table 7: Per-Capita Average Wealth

Average Wealth - AIDIS					
Year	Price Level 2012		Nominal Level		Ratio=Ur/Rur
	Rural	Urban	Rural	Urban	
1961	50,836		1,369		
1971	56,118		2,748	-	
1981	86,388	108,785	10,874	13,693	1.26
1991	146,575	195,096	36,860	49,062	1.33
2002	180,655	272,443	96,011	144,794	1.51
2012	390,223	903,895	390,223	903,895	2.32

Source: Author's calculation using NSS AIDIS datasets. Design weights are used to estimate these values. Wealth Price Index is used to bring the prices from nominal to 2012 price level.

Overview of wealth along Caste in AIDIS

Surveys have usually slightly different composition than the population from Census population. Table 24 shows that ST formed 8-9% of total population and SC formed 18%-19% of total population. The information on OBC is present only for 2002 and 2012 survey. The proportion of OBC increased from 40.28% in 2002 to 43.57% in 2012 i.e. almost an increase of 3 percentage points. Correspondingly we observe a 2% decline in FC share and 1% decline in Muslim. I check the share of total wealth and total land value share and found that SC/ST/OBC have lesser share than their population share. Table 26 presents the representational inequality for different castes.

At all-India level: SC suffers the worst in total wealth share as it owns only around 7-8% of total wealth which is almost 11 percentage points (pp) less than their population share. ST owns 5% – 7% of total wealth which is around 1-2 pp less than their population share. OBC group owns $\sim 32\%$ of total wealth in 2002 which increased only marginally in 2012 resulting in overall worsening of the gap relative to population share (-7.8% to -10.2%), due to considerable increase in their population share. On the other hand FC group share has shown an increase from 39% to 41% in their share in total wealth. Relative to their population share this group improved the gap from +14% to +18%.

Rural-Urban: The overall picture of relative gap between total wealth share and population share for different castes is similar in both Rural and Urban area as found at all-India level except few differences. First rural area seems more favourable for OBC group and less favourable to FC. In Urban area the sign of gap changes for ST group, i.e. they own more wealth than their population share. In 1991, 2002 and 2012, the relative gap is +1.12 pp, +2.24 pp and +1.72 pp respectively presenting urban area to be favourable to ST. I discuss representational inequality by decile in detail later.

Land has an important role in shaping up the overall wealth inequality in the country. The motivation to deal land inequality separately lies in the fact that half of the population still depends on land to get their primary income. The movement of people from agriculture towards manufacturing and service sectors is slower than the movement of GDP of India from agriculture to service sector. The contribution of agriculture and allied services declined from 51.45% in 1951-52 to 14.1% in 2011-12, whereas people dependent on agriculture was $> 50\%$ of the total population in 2012.

There is less representative inequality in *land area distribution* than what we saw above in total wealth. Table 27 provides the representational inequality at overall level among different caste groups. The total land area with ST which was 3.6 pp more than their population share in 1991 increased to 10.81% in 2012 at all-India level. There is no surprise that ST own more land as majority of them still closely live with nature in forests with large land area. The big increase in 2012 survey is surprising. Forest Rights Act (2006) was implemented from December 2007, with the provision to grant legal recognition to the rights of traditional forest dwelling communities- almost all of whom will

³⁵Operational holding is different from Owned land. Land which is utilised for agricultural or other purposes by way of lease, tenancy etc is called operational holding.

³⁶I use WPI index to bring the prices at 2012 level for comparison across years.

be ST³⁷. Land title rights was one of the rights. As per (Aggarwal, 2012), 1.1169 millions claim covering 3% of the forest area was recognized till 30th Apr 2011. OBC group has around 38% land area. SC is the worst in land area ownership. Second if we look at the *land value distribution* we go back to the same level of inequality as we observed in total wealth. The total land value share is 2 pp less for ST and Muslim, 11 pp less for SC, 7.9 pp less for OBC however 14.7 pp more for FC. This suggests the poor land valuation of the land owned by ST and OBC. The poor land valuation could be due to poor land fertility in rural areas or less developed land area in urban areas. After highlighting the basic Representational Inequality (RI) at caste group level, I move to present the concentration of total wealth in the population- at all-India level, rural and urban area separately.

5.1 Concentration of Total Wealth

According to New World Wealth Report, in India, the cumulated wealth of all High Net Worth Individuals (HNWI) increased from \$ 310 billion to \$588 billion and their numbers increased from 84k in 2008 to 153.4k in 2012.³⁸ HNWI's are individuals owning net assets of more than \$1million (=Rs 60,000,00) value. Correspondingly, in the same time period, as per Reserve Bank of India report, the decrease in the population of BPL (Below Poverty Line; Monthly consumption below Rs.1000) was from 407k to 269k. The rate of increase in HNWI's was 82% compared to reduction rate of BPL population by 24%.

This section deals with estimating the top share of wealth historically using surveys and Forbes millionaires list.³⁹First I use micro-data from the NSS AIDIS surveys (for the years 1991, 2002 and 2012) and tabulated data from NSS AIDIS annual reports (for the years 1961,1971 and 1981) to estimate the top share. The tabulated data provides certain brackets of wealth with total population and total wealth in that bracket. I use non-parametric Generalized Pareto Interpolation method to generate a continuous distribution of wealth from the limited information in the wealth intervals (containing average and number of individuals falling into that bracket). Micro-data from the surveys in itself provide the full distribution. However, I test Generalized Pareto Interpolation programme by estimating the full distribution from it and comparing with the survey distribution for 1991-2012 surveys. The difference difference in decile level shares from generalized percentiles method and survey shares is positive in lower deciles and negative in top decile.⁴⁰ This provides a robustness check that the estimates of the top decile share are conservative and it is not overestimated. At the end the results presented here keep the results from generalized percentile method to compare across years, with no induced bias due to different methods.

5.1.1 Two Contrasting Ends

In this section comparison is drawn between the top and the bottom percentile population to show the contrasting nature of wealth inequality. At all-India level, top 10% of population had 45% of total wealth in 1981 which increased to ~ 58%, an increase of 15 percentage points (pp) in 30 years. There is a major jump in 2012 from 2002 almost 10pp. On the other hand looking at bottom 10% Fig. 2b we see, total wealth share is less than .6 % for all the years and for both sectors. Wealth inequality is lower in rural areas than in urban areas and we see an improving trend (i.e. increase in wealth share of Bottom 10%) in both rural and urban areas.

The existence of micro-data from NSS-AIDIS survey from 1991 onwards help in robustness check with different combinations like gross individual wealth equally split in all the household members, net individual wealth equally split in adult members of households and net individual wealth equally split among all the household members. The variations are not very far from the numbers presented in the main text (which is gross individual wealth split at individual level). In the discussion section later, I explain the differences from different variations.

Rural: Looking into rural sector top decile share see a slight decrease/stagnation in the period of 1961-1981 at ~ 43%, which gets withered away in later decades when the top decile share jumped to 51.2%. In 60 years, the change in top decile share is of +8 pp.

Urban: The top decile share in total wealth stood at 55% in 1991 which first declined to 52.5% in 2002 and then gained 7pp to reach to 60% level. In 1991-2002 the top 10% wealth share decreases only in Urban areas while it remains constant in Rural areas. Nevertheless after 2002 the 10% wealth share increases faster in urban areas than in rural areas.

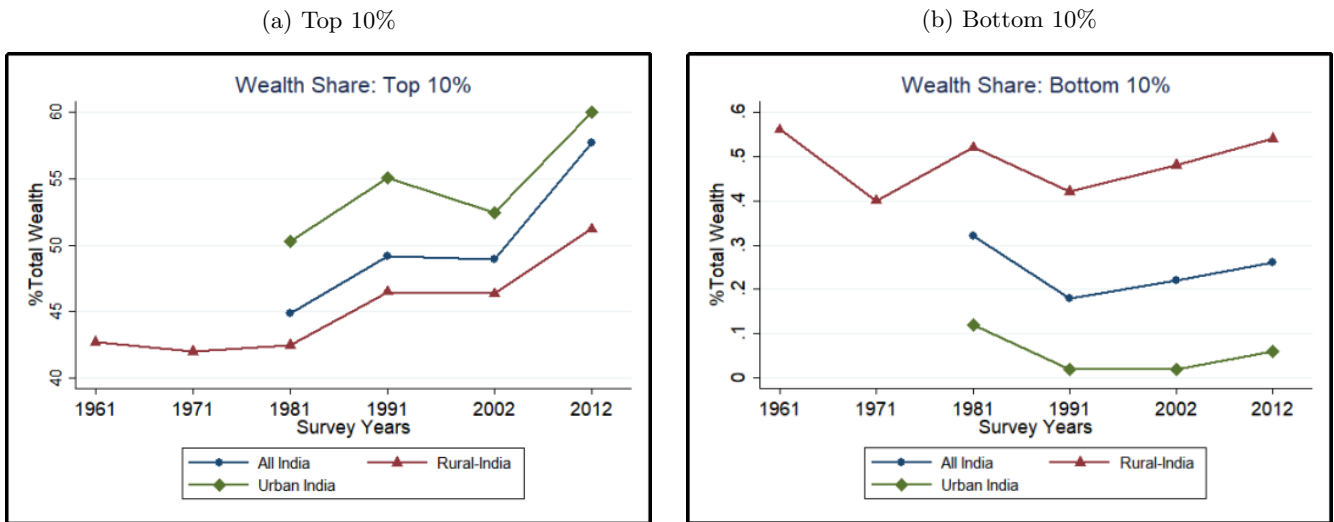
³⁷Though the Act is not only restricted to ST, the condition that last 4 generation should have operated the forest land will automatically remove all other caste groups.

³⁸In 2018 the numbers of HNWI's stands at 3,30,400

³⁹I am pursuing to understand the methodology of HNWI's list of Capgemini and if possible using their dataset to overcome the limitation of few individuals in Forbes dataset.

⁴⁰The reduction in top decile share is 1.76 pp ,1.7 pp and 2.6 pp in 1991, 2002 and 2012 respectively using generalized percentile programme.

Figure 2: Wealth Share:Top 10% and Bottom 10%

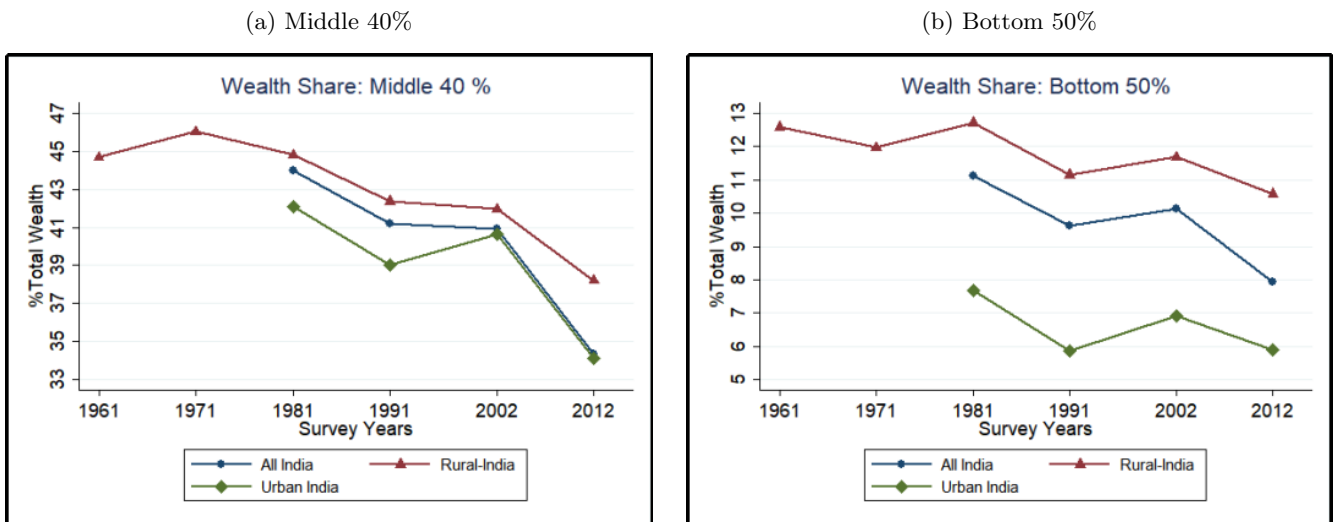


Note: Author's calculation using NSS-AIDIS datasets and reports. Rural India series is from 1961. Urban India series information is available only from 1981.

5.1.2 Weak Base

The share of Middle 40% (which is defined as population in 6th-9th deciles) and Bottom 50%(1st-5th deciles) of the population, can be seen in Fig. 3a and Fig 3b respectively. We see that middle wealth population in Rural India used to own 45% total value of rural wealth in 1961 which has decreased to ~ 39% in 2012. The wealth share almost equals the population share of the middle 40%. In Urban sector, Middle 40% share has declined to below 35% in 2012 from ~ 42% in 1981. We see a jump in 2002 in Middle wealth population which is in contrast to the decline of Top 10% share we saw above. At all-India level, the share of middle wealth population is now closer to urban sector level at ~ 35%. The share of bottom 50% (Fig. 3b) in Rural India has decreased from 12.6% in 1961 to 10.5% in 2012 which implies a drop of 15.9%. Inequality in Urban regions is more extreme. In those regions the bottom 50% owns only 5.9% of total wealth. At all-India level the share of this section stands at 8%. The condition of half of the population of country is dismal in the share of total wealth.

Figure 3: Wealth Share:Middle 40% and Bottom 50%



5.1.3 Rich Topping

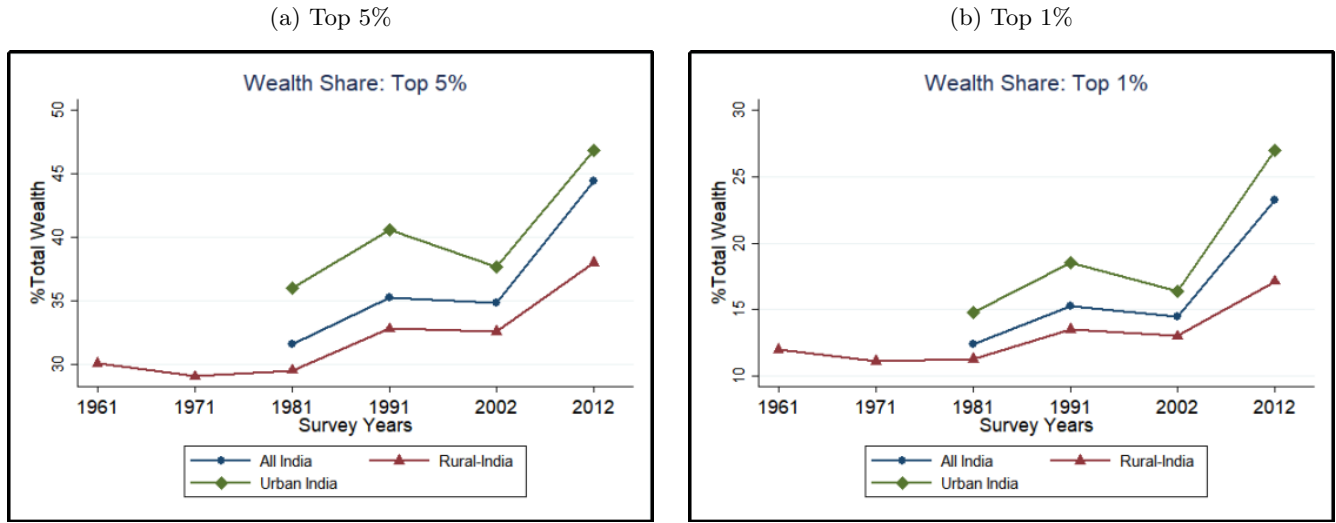
The inequality within top decile is visible from the graphs Fig. 4a and 4b. Wealth is heavily concentrated at the top. *Rural:* Out of the total asset share with top decile in Rural area, top 5% population owned 70.4% in 1961 and 74.3% in 2012. Looking at the evolution of portion of top 1% population in total wealth share of top 10% population, we see first a decrease from 1961 to 1981 and then an increase post 1981 to reach at 33.5% ($=17.14 \times 100 / 51.2$). The period of 1961-71 and 1971-81 saw a small decline which happen to be the years when land reforms were implemented in

India. Land Ceiling Act implementation in 1960's might have some role to play. Under Land Ceiling act, different ceilings on the area of land ownership was specified in different states and land above the ceiling was confiscated by the government to distribute among poor farmers. The small decline gets nullified by a jump in 1991, followed by a decade of stagnation and in the last decade increasing to its highest level.

Urban: The concentration at top is higher in Urban area than in rural area. The portion of wealth share of Top 5% population in the wealth share of top decile, increased from 71% in 1981 to 78% in 2012. Similarly Top 1% population captured 29.4% in 1981 to 44.5% in 2012 of the total wealth share of top decile. The year of 2002 is an aberration, when wealth share of top 5% and top 1% (and also of top 10% share) saw a major decline in urban area. This is the time period just after the implementation of liberalization in India. The decrease in the wealth share at Top 10% is distributed across all the lower deciles. All other 9 deciles saw an increase in their wealth share. Rural area seems to be un-affected in this period in terms of wealth distribution.

All-India: Since the proportion of rural wealth in total wealth saw a decline from 68.5% in 1981 to 45.1% in 2012, we see the top share curve is now more closer to Urban area. The concentration at Top 5% and Top 1% shows a similar trend as in Urban area.

Figure 4: Wealth Share:Top 5% and Top 1%



It has been well documented by Subramanian and Jayraj (2006) that top wealth is not captured properly in the dataset. One can easily see this from the maximum wealth captured in the dataset. The ratio of maximum wealth in the Forbes list to the maximum wealth in survey is 3250, 16507 and 3279 in 1991, 2002 and 2012 respectively.⁴¹ We see that the ratio is very high for the year 2002, compared to 1991 and 2012. The total wealth from the Forbes list comes from only 1, 5 and 46 individuals in 1991, 2002 and 2012 respectively owning 0.17%, 1.72% and 2.65% of the total survey wealth in the same order. There is indeed the issue of under-representation in the survey. Hence to overcome this issue, I use Forbes millionaires list to correct the top share for 2012 and 2002.⁴²

Top shares post correction I: There are several variants to correct the top shares using some external source such as Forbes list. One simple method is adding these few individuals in the dataset. This will certainly change the distribution to some extent, but the biggest demerit is that this method ignores all individuals between the poorest individual in the Forbes list and the richest individual in the survey. Put differently, there is no way to know the corresponding weights on the rich individuals. Hence I refrain from using this technique. The second method, which I use here, is based on the assumption that wealth distribution follows Pareto distribution at top end of the distribution (Blanchet 2016). This approach doesn't tell exactly the threshold above which the distribution will be a good approximation. Hence I take multiple thresholds ($p=95$ means I replace top 5% survey distribution with Pareto generated distribution) as a robustness check. Forbes millionaires list provides a list of very few individuals (2002-5; 2012-46) but can be used to generate the distribution utilizing the convenient property of Pareto distribution that the curve of log (wealth) and log (rank) follows straight line. I simulate the wealth using millionaire's list till the threshold (p) and replace the top wealth distribution from the survey with this synthetic Pareto distribution. The top share post this correction is present in the Table 8.⁴³

⁴¹The historical exchange rate is taken from Foreign Exchange Dealers' Association of India. Exchange rate conversion from \$ to Rs: 1991- 24.5, 2002- 48.4 and 2012- 54.4. The Forbes wealth is assumed to be at individual level, however some of the names in the Forbes list is mentioned as Birla family. Unfortunately the household size is not present in Forbes data. The potential concern is that if the wealth is truly at household level, the discrepancy between Forbes and survey will be inflated.

⁴²For the year 1991, there is only 1 individual present in the Forbes list. For prior years, I intend to use the data of top 0.1% wealth holder from estate tax returns, which is still under preparation (Kumar, 2016)

⁴³I also intend to check the top share with the Net wealth instead of Gross wealth and using Household size. The variations will be within 3-5%.

We can see the top decile share increased by 16pp to reach 64.9% for 2002 and 11.1 pp for to reach 68.8% for 2012 for $p=95$. The increase is higher in 2002 is not surprising as we saw above that the ratio of maximum wealth in the Forbes list to the maximum wealth in survey was higher in 2002 than in 2012, suggesting that 2002 did worse in capturing the wealthy individuals. We see from the Table 8 that as the threshold (p) value (i.e. the p value where we replace survey distribution with Forbes based Pareto distribution increases, increases the estimated top share decreases. For $p=99.9$, the increased share for top decile is 6 pp and 4.4 pp for 2002 and 2012 respectively.

Table 8: Top Wealth Share

	Top Wealth Share 2002						2012					
	10%	5%	1%	0.10%	0.01%	0.001%	10%	5%	1%	0.10%	0.01%	0.001%
Pre-Correction Shares	48.94	34.86	14.46	3.72	0.86		57.7	44.4	23.2	10.06	5.24	
Post Correction Shares												
Threshold p for Pareto												
95	64.93	55.08	38.17	22.44	13.01	7.35	68.81	58.68	40.7	23.95	13.91	7.88
96	64.29	54.26	37.72	22.3	13	7.38	68.74	58.56	40.62	23.92	13.9	7.89
97	63.53	53.29	37.08	22.09	12.97	7.43	68.15	57.8	40.15	23.83	13.95	7.97
97.5	63.02	52.63	36.57	21.91	12.93	7.44	67.8	57.33	39.79	23.73	13.95	8.01
98	62.32	51.74	35.79	21.61	12.86	7.45	67.25	56.6	39.14	23.53	13.94	8.05
98.5	61.5	50.69	34.71	21.18	12.73	7.44	66.64	55.8	38.33	23.26	13.91	8.11
99	60.2	49.03	32.71	20.33	12.44	7.39	66.03	54.99	37.29	22.88	13.83	8.15
99.5	58.13	46.38	29.21	18.5	11.71	7.19	64.53	53	34.54	21.62	13.47	8.16
99.9	54.96	42.31	23.85	13.94	9.6	6.39	62.12	49.81	30.07	17.81	12.02	7.86

It is important here to emphasize that there are some limitations in the methodologies, however comparison with other literature show that the results are quite close. There are some data limitations while using surveys or even Forbes- like under-reporting, non-covering etc. Lack of wealth price index is a big issue in comparing the evolution of per-capita wealth in the country.

Top shares post correction II: Experiment: I present this correction method which is an alternative way and do not require assumption of Pareto distribution. The basic idea is to correct the land value in the survey dataset using the market land price from the website portals which facilitates seller to advertise his/her plot of land/apartments and prospective buyers can search for these posted advertisements.⁴⁴ I use one of the websites to create a dataset of cities and associated land prices (per area)⁴⁵. I then use housing price index from National Housing Bank to convert the prices from website to 2012 level. I use this deflated land rates to correct the land values in NSS AIDIS survey dataset for the year 2012, using the land owned area. Rest all the asset values are kept untouched. The share of Top 10% share increases to 80% which is not very far from wealth share of top decile (73%) in USA in 2014. My calculations are upward biased due to several reasons- first I correct only for 30 cities out of 42 cities available in the NSS AIDIS datasets (for which housing price index was available). These big 30 cities covers all the metropolitan cities like Delhi, Mumbai, Kolkata etc. which have probably the most expensive land rates in India. Second, the land rates gathered from the websites is not representative for the whole city due to two reasons. I use only one website where only some of the sellers post their advertisement. Second I only take five land rates per cities and took average to get the land rate of the whole city. A proper analysis will require using multiple websites and more samples of land rates which I intend to do in future research.

5.2 Representation of Different Castes by Decile

After establishing general findings on wealth inequality, this subsection explores if there is any particular social group which is benefited more by this inequality. If the economic inequality is independent of social inequality one will expect the probability of falling into different wealth deciles/percentiles to be same for any social group. Any digression from this hints towards interaction of social and economic inequality. I use the term representational inequality (RI) from Jayadev and Reddy (2011) symbolizing segregation among social groups in a given attribute space (which here is wealth). The basic idea is to compare the population share of a given caste group in a given wealth decile with its share in overall population. So for example if OBC population is 45% of the overall population, a perfectly representational equality will imply representation of 45% OBC population in all the deciles. Any deviation from this will lead to Representational Inequality. The statistics I estimate to compare can be given by:

$$RI_j^{YCD} = \frac{Popshare_j^{YCD} - Popshare_j^{YC}}{Popshare_j^{YC}} \quad (1)$$

⁴⁴In recent years many of such websites have sprung in India like 99acres.com, magicbricks.com etc.

⁴⁵I randomly pick 5 land rates from the website per city and took mean of it. This sample is nowhere representative. This is just an experiment to present the method

where $Popshare_j^{YCD}$ denote the population share of caste C in decile D in year Y for sector $j \in (Rural, Urban, India)$. A higher representation (than the population share) for a caste group in top decile, will increase representational inequality. That caste group is the beneficiary of the inequality. On the other hand higher representation (than the population share) of a caste group in the bottom deciles also increase RI, but that caste group is in a way victim of the inequality.

The RI statistics are present in the Table 9 for three categories- Top 10%, Middle 40% and Bottom 50%. A positive value imply more than the proportionate population share. Looking at the top decile in both Rural and Urban sector, we see that only FC⁴⁶ have positive RI or in a sense beneficiary of the inequality. All other social groups SC/ST/OBC and Muslims have negative RI, which means they are under-represented in top decile. Looking at Middle 40%, we can see that ST, SC and Muslims have negative RI in both 2002 and 2012. Again only FC has positive RI. In Bottom 50%, only FC has considerable negative value of RI, OBC is close to zero. SC and ST have quite high value. The RI pattern is similar in Rural and Urban area.

Table 9: Representational Inequality

		2002			2012		
	caste_group/deciles	Bottom 50%	Middle 40%	Top 10%	Bottom 50%	Middle 40%	Top 10%
INDIA	ST	0.36	-0.32	-0.72	0.4	-0.36	-0.75
	SC	0.37	-0.31	-0.81	0.32	-0.26	-0.79
	OBC	-0.03	0.06	-0.21	-0.05	0.08	-0.17
	FC	-0.34	0.27	0.91	-0.3	0.21	0.98
	Muslim	0.19	-0.14	-0.51	0.12	-0.07	-0.46
	Others	-0.63	0.05	4.58	-0.58	0.01	4.4
RURAL	ST	0.34	-0.33	-0.71	0.34	-0.32	-0.76
	SC	0.35	-0.34	-0.84	0.28	-0.26	-0.75
	OBC	-0.06	0.08	-0.09	-0.1	0.12	-0.04
	FC	-0.38	0.37	0.89	-0.27	0.22	0.94
	Muslim	0.18	-0.17	-0.45	0.11	-0.06	-0.57
	Others	-0.71	0.05	6.88	-0.6	-0.11	7.44
URBAN	ST	0.34	-0.28	-0.45	0.34	-0.29	-0.45
	SC	0.37	-0.25	-0.75	0.36	-0.23	-0.81
	OBC	0.07	-0.01	-0.32	0.06	-0.01	-0.27
	FC	-0.24	0.15	0.57	-0.28	0.17	0.67
	Muslim	0.22	-0.12	-0.61	0.24	-0.16	-0.49
	Others	-0.5	0.11	2.11	-0.46	0.07	2.17

Source: Author's calculation. The numbers are comparable across caste and years as they are standardised. A positive value implies more population share in that class group than their population share in overall population.

In summary, we observe that that FC is relatively more present in top 10% and Middle 40% of population where almost 90% of the wealth is concentrated. SC, ST and Muslims have relatively more presence in the Bottom 50%. OBC population is evenly distributed across wealth deciles. The representational inequality is silent on the relative wealth share. It only captures the representation of caste groups in a pre-decided bins, which here is chosen as deciles. Next, I analyse the wealth share under Rural and Urban separately.

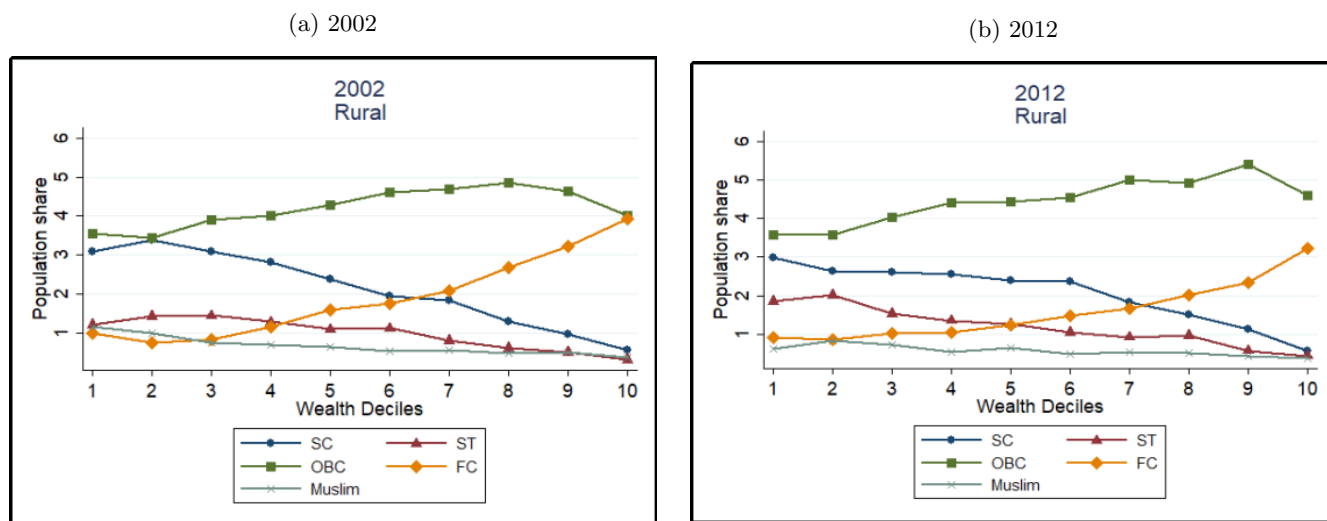
5.2.1 Rural

Fig 5a and 5b show the percentage share of different castes in different deciles in Rural areas. For the 2002-03 survey, in the top decile 40% belong to FC and 40% belong to OBC. The representation of OBC in top decile is slightly less than than their population share in the survey (which is 41.97%), however FC is almost double represented in top decile (21.27% population share). While in 2012, we see a slight drop in the representation of FC(to 32%) and a corresponding increase in OBC(45.9%) in the top decile. Not to miss the fact that their population share of OBC and FC changed to 44.2% and 17.6%. Hence relative to their total population share, the share of FC represented in the top decile remains almost constant while OBC share improved slightly. The combined share of FC and OBC in the top decile has almost remained same in a decade (79.43% in 2002-03 and 78.14% in 2012-13). The representation of SC, ST and Muslim is lower in the top 10% (in fact in all top 5 deciles, their representation is lower) than their population share. The situation is the worst for SC. Their representation in Top 50% is almost 41 percentage points (pp) lower than their population share in 2002.⁴⁷ The representation of ST and Muslim in Top 50% is lower by 15.5 pp and 9 pp respectively than their population share in 2002. in 2012 SC are under-represented by 20.4 pp, ST by 29 pp and Muslims by 5pp in the top 50%. This implies a deterioration of the relative wealth of ST.

⁴⁶ Apart from Other which is a small rich minority group

⁴⁷ So correspondingly they are over-represented in Bottom 50% by 41 pp.

Figure 5: Proportion of different castes-Rural

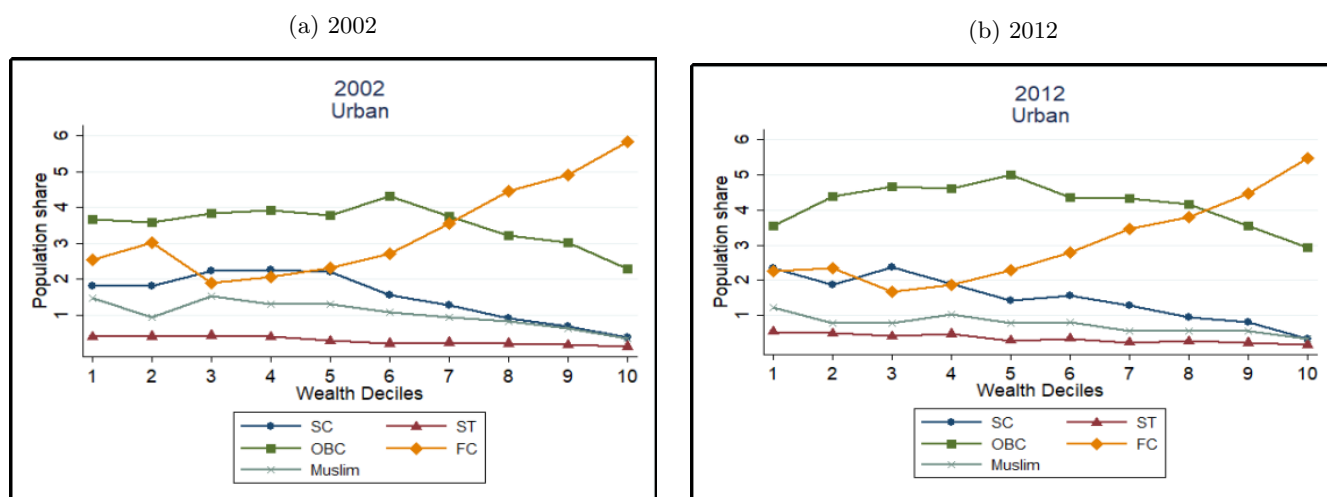


Note: This graph shows the proportion of population in different gross wealth deciles using NSS AIDIS datasets in Rural area.x axis is wealth deciles, yaxis is population share. “Other” population is not presented here as their population share is very small.

5.2.2 Urban

The results in urban areas are very different from the results on rural areas, just discussed. Fig 6a and 6b show the percentage share of different castes in different deciles in Urban areas. For 2002-2003 survey, in the Top 10% population, 58% are from FC and 23% are from OBC. The representation of OBC in top decile is 12.4 pp less and for FC is 25 pp more than their respective population share in the survey. In 2012-13 the increase in OBC representation to 29% and decrease in FC representation to 54.7% is in line with their population change. The combined share of FC and OBC increased from 81.3% in 2002-03 to 84% in 2012-13, while their combined population saw an increase from 68.6% o 71.9%. SC, ST and Muslim share in top decile are not only below their total population share in top decile but also lower in combined five top deciles. Similar to rural areas SC is skewed in the lower tail of the wealth distribution. Its representation in Top 50% is 27.7 percentage points (pp) lower than its population share in 2002. The under-representation of ST and Muslims in the top 50% is 5pp and 13pp respectively in 2002. In 2012-13 the representation for SC, ST and Muslim in 2012 lies below their population share by 24.2 pp, 4.8 pp and 9 pp respectively i.e. an improvement in SC and Muslim.

Figure 6: Proportion of different castes-Urban



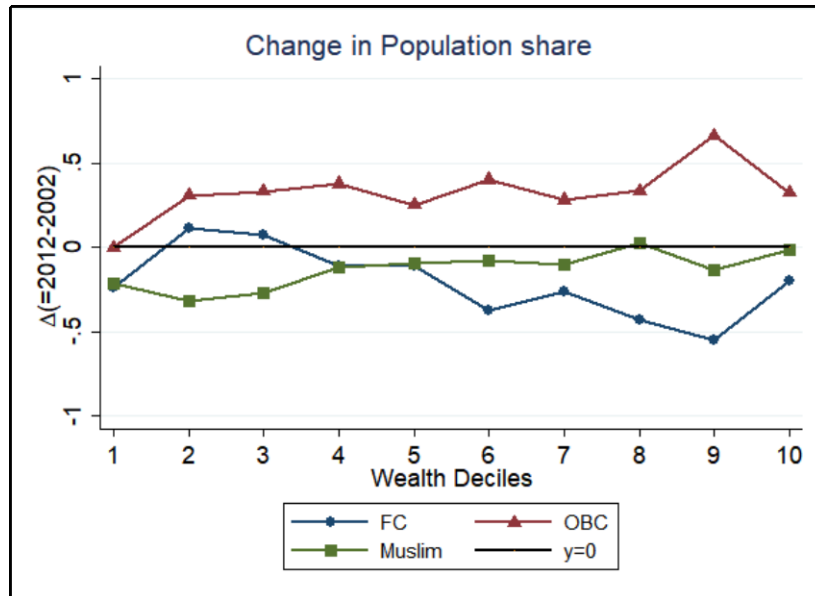
Note: This graph shows the proportion of population in different gross wealth deciles in Urban area.x axis is wealth deciles, y axis is population share. “Other” population is not presented here as the population share is very small.

5.2.3 Discussion

Some important takeaways from the above observations are:

First SC has the lowest wealth ownership in urban and rural areas. There has been some improvement over a decade but the condition is far from satisfactory. Second, FC is still overly represented in Top 10% or Top 50% of the population. Third I want to highlight the issue of **reclassification of castes**. We observed an increase in the relative representation (RI- changed from -0.21 to -0.17) of OBC population in Top 10% from 2002 to 2012 time period. If the population share of OBC in overall population had remained same in the two time periods and there was no scope of inclusion or exclusion in OBC, then this result would have suggested an improvement in OBC caste group. However, OBC group of caste is a floating list of castes. Mandal Commission in 1980 charted out the criteria based on which OBC status is conferred to a caste for receiving the reservation benefits.⁴⁸ National Council of Backward Commission (NCBC)- a statutory body, hears the petition from different castes for the request of inclusion into the OBC list. Time and again different castes have been included in the list while there has been no exclusion of the castes from the list in last 17 years (Annual Reports of NCBC). This is probably one of the main reasons behind the increase in the population share of OBC. Fig. 7 depicts the change in the population share of OBC in different deciles. We see a positive change in the population share for OBC across all the deciles. Correspondingly there is decline in population share of FC, with major declines towards higher deciles. There is also a decline in Muslim population which is concentrated more in lower deciles than in higher deciles. This further raises another concern whether relatively rich FC castes are being included into OBC category.

Figure 7: Change in population share between 2002 and 2012



Note: Author's calculation from NSS AIDIS surveys. It represents that OBC population has seen a positive change in population vis-a-vis drop in population share in FC and Muslim groups. y axis denotes the change in population share in time period 2002-2012.

There could be other possibilities of increase in OBC population like -increase in population share is due to sampling issues, higher growth rate of OBC, higher Inter caste marriage into OBC, classification of other communities (like SC/ST) into OBC.

5.3 Total Wealth Inequality within Caste

Till now, I have treated the broad social groups/castes SC, ST, OBC, FC and Muslims as one homogeneous groups- because government affirmative actions does so. However, this is far from social reality. Each one of them is composed of hundreds of different narrower version of castes (jatis) in hierarchical fashion. This multifarious layering could create heterogeneity on economic aspects. Several other reasons like- accessibility to education, different level of awareness and ability - can create heterogeneity. In this section, I try to analyse the distribution of wealth within every caste. In a big country like India, even if groups are small, this wouldn't be a reason not to study them. Further looking at inequality only across castes gives an impression of economic uniformity within the castes, which will be put into question in this section.

⁴⁸A score is calculated based on social, educational and economic criteria. Castes which score above a fixed point gets OBC status. Higher weight is assigned to social criteria. This ensure socially backward castes have higher chance of getting the status.

5.3.1 Top Deciles

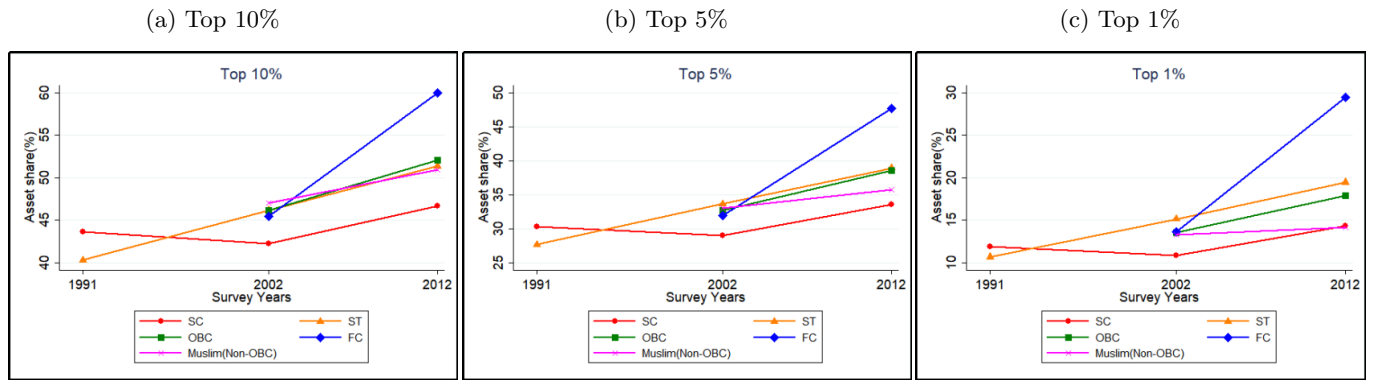
Fig. 8a, 8b and 8c show the share of wealth with Top 10%, Top 5% and Top 1% respectively. We can see an increasing trend of within caste inequality in all the castes. The divide between FC and rest of the caste is more prominent in 2012 than before. Interestingly within ST the allocation of wealth to Top 1%, Top 5% and Top 10% population increased than that of in SC in 2012 which was opposite three decades earlier.

Within *Forward Caste*, Top 1% within FC owned almost 13.6% of the total FC wealth in 2002 which increased to 29.4% in 2012. Top 5% owned 32% in 2002 which increased to 47.6% in 2012. And Top 10% now owns 60% of the total FC wealth. This is a drastic change in ten years which needs more enquiry. The inequality within FC group is the highest.

Within *ST*, Top 1% owned 10.7% of total ST wealth in 1991, which increased to 15.1% in 2002 and further to above 19.5% in 2012. Top 5% share of total ST wealth increased from 27.7% in 1991 to 38.93% in 2012. Top 10% of share increased from 40.3% in 1991 to 51.4% in 2012. This caste group has seen a consistent rise in within caste inequality. Within *SC*, Top 1% owned 11.9% of total SC wealth in 1991, which decreased to 10.83% in 2002 and then increased to 14.4% in 2012. Top 5% share of total SC wealth increased from 30.3% in 1991 to 33.5% in 2012. Top 10% of SC share increased from 43.6% in 1991 to 46.7% in 2012. SC has the lowest level of within caste inequality in terms of top decile share.

Within *OBC*, Top 10% share increased from 46.2% in 2002 to 52% in 2012.

Figure 8: Share of Top Deciles Within Caste



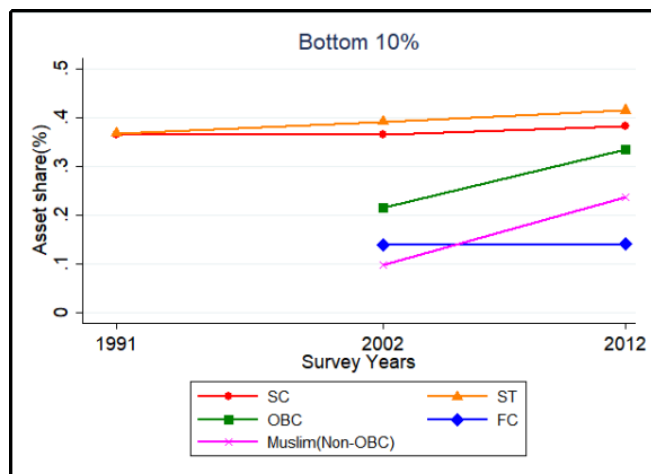
Source: Author's calculation using NSS AIDIS dataset. This graph shows the wealth share in Top 10%, 5% and 1% within different caste groups.

5.3.2 Lower Deciles

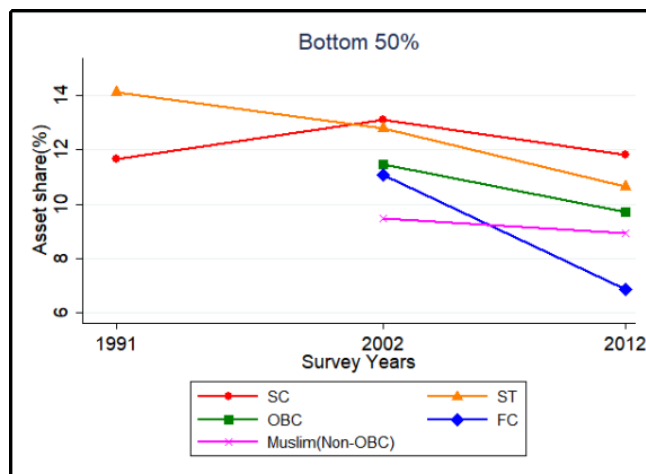
Fig 9b shows that Bottom 50% of the population has lost 2-4 pp within all the caste categories. The major decline of their share in FC followed by ST, OBC and SC and Muslims. Fig 9a shows that there is slight improvement of wealth share in Bottom 10% population within every caste. The gain is highest in Muslims (.14 pp) followed by OBC(.12 pp).

Figure 9: Share of Lower Deciles Within Caste

(a) Bottom 10%



(b) Bottom 50%



Source: Author's calculation using NSS AIDIS dataset. This graph shows the wealth share in lower deciles within different caste groups.

Inequality is quite high in every caste group. The threshold wealth by decile and caste is provided in the Table 28. One way to intuitively visualise these numbers is to look at how many times of wealth an individual needs to enter into 10th decile than to enter into 2nd decile in given wealth distribution within every caste group (i.e. ratio of threshold of 10th decile and threshold of 2nd decile). In 2002, a person from ST group needed 23(=118,060/5625) times more wealth to reach to top decile from the threshold of second decile. Similarly for SC, OBC, FC and Muslim one required 22,32,45 and 57 times respectively to enter into 10th decile.

5.3.3 Discussion

We see a very high within caste inequality which has increased from 2002-2012. It hints that treating these big administrative castes as homogeneous groups is far from reality. Looking at the entire distribution, all the population from 2nd to 9th have lost their share within castes. The major gain is in the top decile and a minor gain in the bottom decile within every caste. The inequality within FC has increased and it is potentially one of the reasons behind uneasiness among certain Forward caste groups in country and their demand for OBC status- which helps them to gain access into reservation benefits (positive affirmation). “Jat” an agrarian community from North India, are demanding the OBC status in Central Government list of OBC^{49, 50}. Patidar (people with well known surnames Patel) group in 2015, started agitation for similar demand which became the central issue in 2017 state's election.⁵¹. These caste groups are not considered socially backward and highly likely to have better economic outcomes. The surprising demand for OBC from these communities can be rationalized through the close threshold ownership and a sense of competition for scarce resources. From Table 28 we see close threshold values of FC 8th and 9th decile and OBC 10th decile. Lets call these two groups A and B. The population share group A and group B is almost same ($A=2*1*20\%$; $B=.1*40$). There are two differences between A and B. Group A is considered upper caste in the society and hence may have some advantage in terms of better connections to get into jobs. Group B has reservation benefits in government jobs. If the utility from being in upper caste reduces due to any reason than and the perceived utility from being in OBC group increases group A will demand for OBC status. Since in reality there are many castes in group A with different hierarchy, if not whole group A, some castes from A will sure demand for OBC status. The perceived benefits come from expectation of lesser competition in OBC. The demand for OBC status can lead to reclassification which is probably one of the reason for the increase in OBC population. The unmet demand might die down or might resurge again with change in the government. This is what I call Forward caste backlash can be seen in different forms in the society today.

There is a large literature finding that people evaluate their life relative to individuals that are similar to them. For example, Luttmer (2004) finds that subjective well being reports depends on whether a respondent earns well relative to her neighbors. He uses US data. Furthermore, using data from former soviet countries,(Senik, 2009) finds that colleagues and former schoolmates are very important for relative income comparison. Lets assume that the same

⁴⁹There are two different lists- Central level OBC list, which makes a caste eligible for reservation benefits in central universities and central government jobs and a corresponding state level OBC list.

⁵⁰https://en.wikipedia.org/wiki/Jat_reservation_agitation

⁵¹https://en.wikipedia.org/wiki/Patidar_reservation_agitation

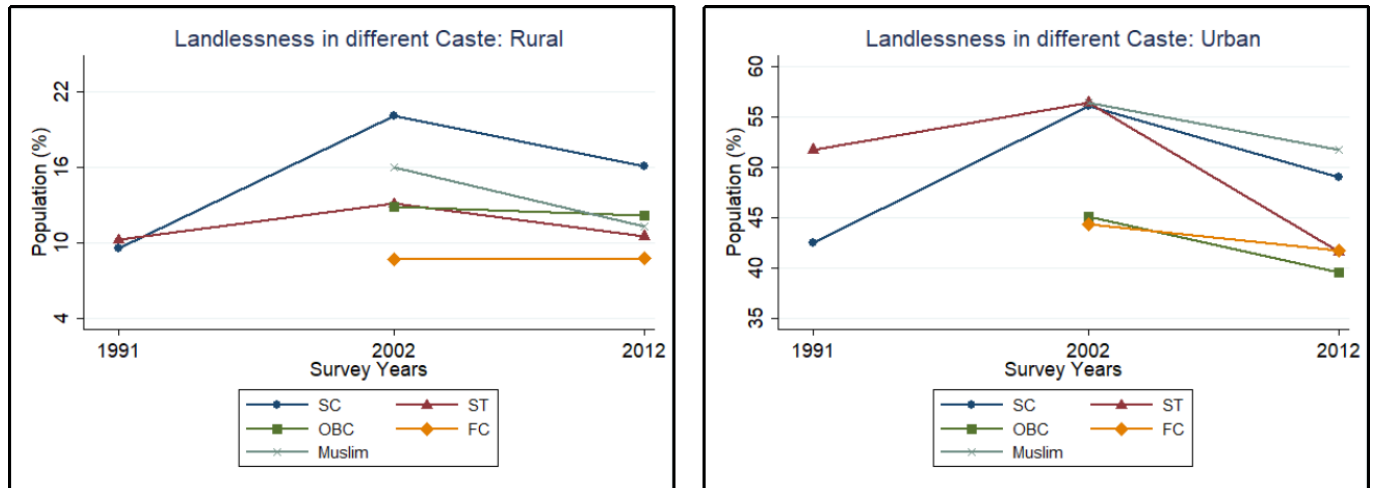
mechanism holds in India. Then, if income and wealth are correlated, people with similar wealth might work together or go to the same schools. Thus FC individuals from the 8th decile of the FC wealth distribution might work together with OBC individuals from their respective 10th decile. In this case FC individuals might feel deprived, when their OBC colleagues with similar wealth level get a prestigious position in administration due to reservations. Maybe they just don't consider members of the FC tenth decile as a relevant reference group.

5.4 Total Land Inequality

Land forms the major part of asset value and contribution of land in total wealth has almost remained at same level in last 60 years. In this section I analyse inequality coming from unequal distribution of land ownership. NSS AIDIS provides information on land area owned and its valuation⁵². We first look at the landless population in the data. Landlessness in India has increased from 15% to 24% (Fig 31). The landlessness in rural area (< 14%) is much lower than landlessness in urban area (45%). In Figure 10 I plot the percentage of population with no land by caste groups in rural and urban sector separately. The landlessness among SC increased by 10 pp in 1991-2002. Similarly ST saw a rise in landlessness in this decade by 3 pp. In Urban areas the pattern is similar. I discuss the big increase in landlessness among SC and ST in the period from 1991 to 2002 in detail below. One of the main tools used in liberalization process was de-licensing. Before 1990, India was almost a closed economy. Setting up an industry required license from the government and the allocation of land to the industries was decided while giving licenses. Post liberalization companies are free to choose their location and will obviously try to target the land which can be easily acquired at cheaper prices (Sardana, 2012). Since ST and SC are vulnerable to shocks and relatively more poor their bargaining position might be bad. This can explain selling of land at low prices. A second reason for the post 1991 increase in landlessness is industrialisation. Economic progress of certainly will involve diversion of agricultural or relatively unproductive/unused land towards construction of roads, rails, industries etc. This will in general increase landlessness. But we see higher rate of landlessness in SC community than in any other communities. The reason might be that the land was acquired without proper compensation. ST community is more attached to forest areas where acquiring land is more tedious due to complex regulations related to environment and restrictions on buying land managed under Scheduled Tribe area by outsiders. This might be the reasons for reduced rate of landlessness among ST.

Among other communities, FC has the lowest level of landlessness in both Rural and Urban area. We see a drop in landlessness across all castes in 2012 from 2002. The drop can be due to the fact that land area was captured till 3 decimal points in 2012 but only till 2 decimal points in 2002.⁵³

Figure 10: Share of Population - Decilewise



5.4.1 Concentration of Land Within Land-Ownning Population

I restrict the analysis of concentration of land area and land value among the population with non-zero land area. I perform deciles level analysis, with both land area based deciles and land value based deciles. Both have their own advantages. Land value based deciles are more intuitive as individuals with more land value will be categorized into higher deciles. However, the reasons for the change in land value share from one year to another can be due to multiple reasons like- appreciation/depreciation of the existing land (without any change in land area), buying/selling of land

⁵²One should keep in mind that land valuation was different in 2002 and 2012

⁵³Though for the analysis to be comparable I have rounded of the land area to two digits, it will not correct the behavioural differences in answering the question by respondents. More people will report non-zero land area, thereby reducing the landlessness.

or due to the change in the method of collecting information. It is difficult to narrow down to one reason or interaction of all the factors. I intend to explore more in my future research. Land area based deciles, on the other hand will classify large land owners in terms of area into higher deciles. A forest dweller with large land area will be categorized into top decile.

Deciles based on Land Area: Table 10 shows the top 10% of population based on Land Area in Rural area occupied 48.4% land area in 1991 which increased slightly to 49.4% in 2012. The corresponding land value of this share of land was 33.9% in 1991 and increased to 38.2% in 2012, i.e. an increase of 4.7 pp. The change in land valuation could be due to either increase in the valuation of the existing land or purchase of the more valued land. Unfortunately I can't elaborate more on this further because of lack of necessary information in the survey datasets. In Urban area, the top 10% of the population occupy 83.1% of the total land area in 1991 which decreased to 77% in 2012. However, the land value share (i.e. for the land owned by top 10% population) increased from 27.4% to 34.3%.

The entire distribution of land area in Rural sector seems quite stable. The bottom 50% population share of land area has remained almost constant at 6.8% in Rural area. The value increased from 15% in 1991 to 17.9% in 2012. For Middle 40% (i.e. 6th-9th decile population) the share remained constant at 44% with land value decreasing from 51% in 1991 to 44% in 2012. Urban area is more dynamic. Bottom 50% doubled their land share from .9% to 1.8% in 2002 but saw a decline in next decade to reach 1.4% land of total land area. The land value increased too from 27.5% in 1991 to 30.8% in 2002, however declined to 23% in 2012. Middle 40% occupied 16% of land in 1991 which decreased to 11% in 2002, then increased to 21% in 2012. The land value follows similar trend which first decreased from 45.1% in 1991 to 39% in 2002 and then increased to 42.7% in 2012.

Table 10: Land Area and its Value distribution

Percentage of Land Area Ownership									
	Bottom 50 %			Middle 40%			Top 10%		
Year	1991	2002	2012	1991	2002	2012	1991	2002	2012
India	4.10	4.01	3.84	45.83	41.33	35.88	52.18	53.01	54.48
Rural	6.68	6.76	6.79	44.98	44.47	43.78	48.34	48.76	49.43
Urban	0.86	1.77	1.36	16.05	11.27	20.96	83.09	86.96	77.68
Percentage of Land Value Ownership									
	Bottom 50%			Middle 40%			Top 10%		
Year	1991	2002	2012	1991	2002	2012	1991	2002	2012
India	21.67	24.80	28.83	43.72	42.99	41.67	32.50	33.88	35.30
Rural	15.01	15.92	17.85	51.12	48.47	44.00	33.87	35.61	38.14
Urban	27.53	30.80	23.02	45.13	39.39	42.66	27.35	29.81	34.32

Source: Author's calculation using NSS AIDIS datasets. Design weights are used to estimate these values. The deciles are based on land area.

Decile based on Land Value: The top 10% of the population based on land value, owned 53% in 1991 which increased to 61.5% in 2012. The figures are quite close to the concentration of total wealth share we saw before. The corresponding share of land area at all-India level has decreased from 31.6% in 1991 to 29.1% in 2012. For the middle 40% classes the land value varies more than the land area. The land value has decreased $\sim 5 - 6$ pp in both Rural and Urban area. Moreover there is a decline of ~ 2 pp in the Bottom 50% of the population. Even though this decline is smaller than the decline in Middle 40%, the base is also very small, hence the percentage change is higher.

Table 11: Land Value and its Area distribution

Percentage of Land Area Ownership									
	Bottom 50 %			Middle 40%			Top 10%		
Year	1991	2002	2012	1991	2002	2012	1991	2002	2012
India	16.04	16.40	21.53	52.40	50.30	49.34	31.55	33.30	29.13
Rural	15.70	16.11	19.22	53.13	50.39	49.97	31.16	33.50	30.81
Urban	14.60	11.51	16.94	44.28	41.08	42.55	41.13	47.41	40.50
Percentage of Land Value Ownership									
	Bottom 50%			Middle 40%			Top 10%		
Year	1991	2002	2012	1991	2002	2012	1991	2002	2012
India	8.11	8.54	6.21	38.84	38.68	32.28	53.05	52.78	61.50
Rural	8.00	8.81	6.81	39.61	39.09	33.67	52.39	52.10	59.52
Urban	8.90	9.12	6.84	36.09	37.20	32.85	55.01	53.68	60.32

Source: Author's calculation using NSS AIDIS datasets. Design weights are used to estimate these values. The deciles are based on land value.

5.4.2 Discussion

The above analysis is only among the population with strictly positive land owned. Within those population, we saw that land area is highly concentrated in Urban areas to the top 10% population. However, big land doesn't mean valuable land too. In urban areas, usually the land in city center/City Business District (CBD) is very costly and in the fringes of city land is cheaper. That's why we see that 77% land owned by top decile by area (Table 10) in urban areas is only 34% of the total land value and a meagre 1.36% land area with Bottom 50% has 23% total land value in 2012. In rural area the gap is not so high as most of the land has been already been brought into agriculture. The thrust of agricultural revolution since 1960's may have to certain extent levelled out the variability of utility of the land in rural area.

Cross Ownership: 2012: The survey data in 2012 allows to do a cross ownership analysis, ie. how many Rural area residents own land in Urban area and vice-a-versa. Table 29 provides the cross ownership in Land Area and Land Value terms. We see that only 0.81% of total HH's from Rural areas own some land in Urban areas, where we can see that 15% of total HH's from Urban areas own some land in Rural Areas. In terms of Land Value, 14.75% of rural land is owned by residents in urban area while only 1.2% of the Urban Land is owned by the rural residents. Family migrating to urban areas, usually retain their lands in rural areas, either with some family members taking care of the land or leasing the land to agricultural tenants. In terms of land area, 4.3% of the urban Land is owned by the rural Residents, whereas 12.3% of the rural land is owned by urban residents. The discrepancy in land value and land area hints that rural residents have bought cheaper land in urban areas. Such lands could be far from the city center, or less developed within the city.

Employment Type and Caste Overlap: In this section I check the employment type of different caste groups for 2002-03 and 2012-13 survey.⁵⁴ The employment type is based on the major source of income during 365 days preceding the survey of the household head. In Rural area, the 5 classifications in 2002 include Self-employed in agriculture and non-agriculture; Labor (agriculture and non-agriculture) ; Others. In 2012 in Rural areas we have an extra classification of Regular/Wage salary. In Urban areas, there are 4 employment types, namely- Self employed, Regular Wage/Salaried, Casual Labor and Other.

Rural Landscape: A look into the Table 30 and 31 show the percentage of employment type in different caste group i.e. where different castes are employed. We observe an increase in agricultural self-employment in 2012 from 2002 across all castes (except Muslim) in Rural area. In 2012, 51% of the ST and FC population are engaged in Self-Employed in Agriculture category, followed by 45%, 32%, 29% of OBC, Muslim and SC population. The second biggest employment type in Agricultural Labor which saw almost 10 pp decline and again the decline is present in all the castes. Unfortunately regular wage salary was not captured for rural area in 2002 to make a comparison in time. Highest proportion of FC are employed in salaried job, followed by SC.

Table 32 and 33 show the distribution of caste group representation in different employment types. There is not much change in a decade. Self-employment in agriculture has higher (than their respective population) proportion of FC, OBC and ST. Also a higher proportion of OBC and Muslims are employed in non-agricultural self-employment. Agricultural labor is engaging much higher proportion of SC's than their proportionate population (11 pp difference).

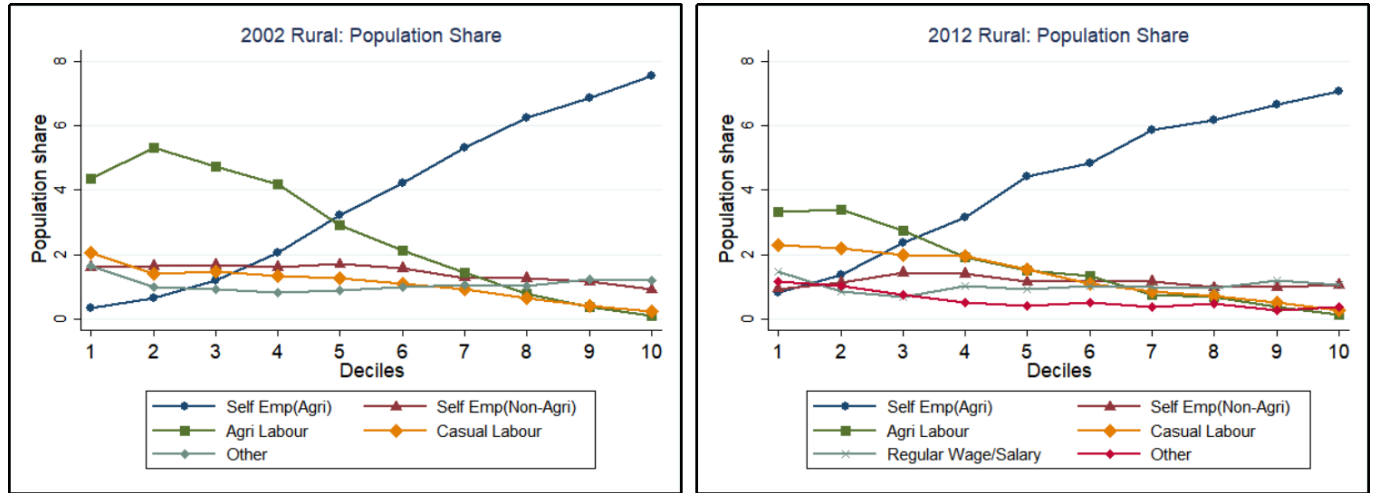
Urban Landscape: In Urban area (from Table 30 and 31) maximum people are employed in regular wage type (40%). 45% of SC, ST and 50% of FC are employed in regular wage sector. On the other hand only 36% of OBC and 32% of Muslims are in wage employment. The next higher employment type is self-employment at 37% in 2002 which saw a decline to 32% in 2012, and the decline is across all the caste groups. Higher proportion of SC(25%) and ST(23%)

⁵⁴Tabulated data before 1981 doesn't allow the analysis at caste-employment type analysis. 1991 census has no OBC information. 2002-2012 decade is important to analyse as Indian economy saw a sustained very high economic growth rate.

population are engaged in casual labor, and the proportion has increased from 2002 to 2012. From table 32 and 33, we see higher representation of FC in regular wage employment in both years. There is an observed drop in the share of FC, and a corresponding increase in OBC share, which again hints towards re-classification of FC castes with higher share in regular wage employment.

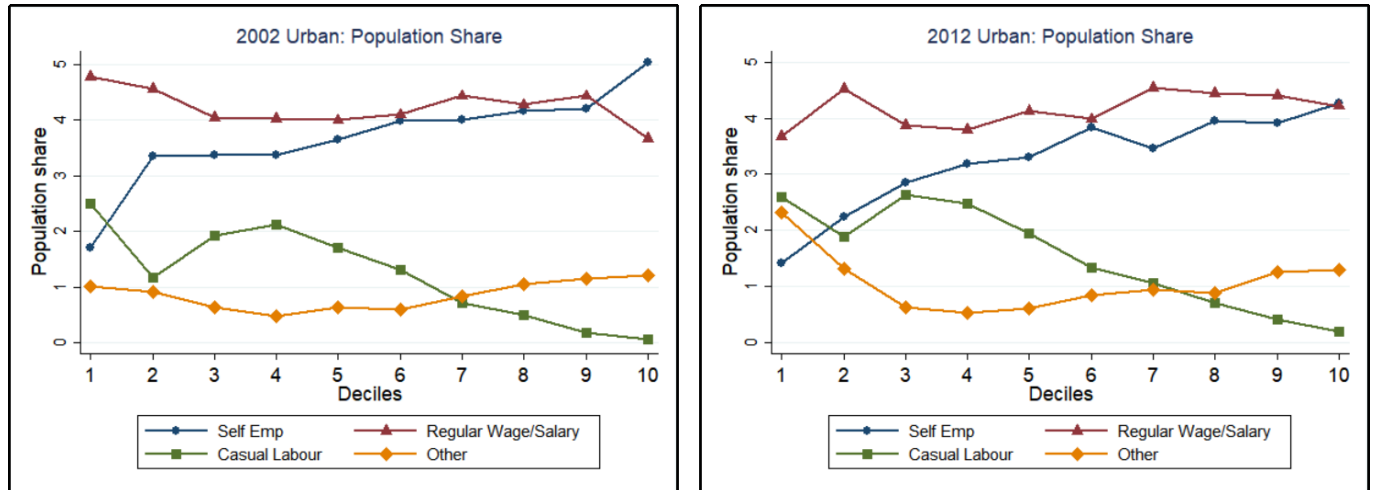
Next, I check the profile of employment type in different deciles. In Fig. 12, we observe that self-employed households in agriculture are concentrated in higher deciles. On the other hand, almost all the agricultural labourers households are confined to lower deciles. Combining this with the above findings that almost 50% of the SC population is employed in agricultural self-employment is surprising. The simple reason is a clear divide in the type of caste present in self-employment. The rich section of self-employment comes from FC and poor section of self-employment is from SC. It is clear from the top wealth decile caste composition in Table 34. In Rural area within top decile, (which has 60% wealth share) we find only 5% SC compared to 34% FC with employment type Self-Employed in Agriculture in 2012.

Figure 11: Share of Population - Decilewise



In Urban area, self-employed and regular wage earners are predominant in top deciles as we can see below.

Figure 12: Share of Population - Decilewise



Age-Deciles Analysis: Micro data in 1991-92, 2002-03 and 2012-13 allows to perform an age analysis in different deciles. Table 12 provides for mean and median age in different wealth deciles. We see increasing mean age with deciles in all the years.

Table 12: Age Profile in Wealth Deciles: India

Age Profile in Wealth Deciles (India)												
	Deciles/Year	Life Expectancy	1	2	3	4	5	6	7	8	9	10
Mean Age (yrs)	1991	58.4	23.84	24.12	24.01	24.34	24.6	24.64	24.69	24.72	25.05	25.97
	2002	63.4	24.73	24.95	25.25	25.12	25.36	25.67	25.95	26.38	27.22	28.45
	2012	67.3	25.85	26.69	27.7	27.14	27.5	28.11	28.69	29.21	29.9	32.75
Median Age(yrs)	1991	58.4	22	22	21	21	21	20	20	20	20	20
	2002	63.4	22	22	22	22	22	22	22	22	23	24
	2012	67.3	23	24	25	24	24	25	25	26	27	32

Age-decile table for Rural and Urban area separately is present in Appendix (Table 36 and 37).

Education-Deciles Analysis: I perform education analysis⁵⁵ in different wealth deciles. Table 13 provides the results for all-India. Since there is an overall increase in the education level in the country, we see an overall drop in population with zero education in all the deciles during 1991-2012. 1991-2002 decade saw an increase (12 pp) of 5-12 yrs of educated population in top decile. In 2002-12 decade we see a decline (-6.2 pp) in population share of “Below Secondary” education in top decile and increase of “Above Higher Secondary” educated (i.e. > 12 years education) by 8.8 pp increase. This is in line with the increase in share of salaried class share in top decile.

Table 13: Education Profile in Wealth Deciles (India)

Educational Profile- Decile-wise Percentage (India)											
year	Deciles	1	2	3	4	5	6	7	8	9	10
1991	Illiterate	52.29	66.28	64.48	60.47	56.46	54.67	51.64	47.41	42.15	34.84
	Below Primary	16.79	15.37	15.74	16.51	17.81	17.51	17.71	18.3	18.91	18.42
	Below Secondary	21.27	13.29	14.12	16.63	19.14	20.1	21.08	22.87	24.14	25.08
	Secondary/Higher Secondary	7.38	3.52	4.15	4.84	4.99	6	7.14	8.14	10.25	13.85
	Above Higher Secondary	2.27	1.53	1.41	1.53	1.6	1.72	2.41	3.26	4.53	7.79
2002	Illiterate	43.98	56.15	53.19	49.87	47.23	43.36	40.76	37.09	30.61	23.7
	Below Primary	15.78	16.1	16.52	17	16.71	16.94	16.79	16.2	16.05	13.87
	Below Secondary	26.91	20.93	23.18	24.36	26.38	27.76	28.91	29.14	29.68	29.48
	Secondary/Higher Secondary	10.79	5.6	5.72	7.31	8.01	9.83	10.81	13.47	17.42	21.33
	Above Higher Secondary	2.55	1.22	1.39	1.46	1.66	2.12	2.74	4.1	6.23	11.62
2012	Illiterate	39.71	44.36	41.08	39.82	37.77	34.12	30.68	27.27	23.69	16.13
	Below Primary	17.2	19.35	18.37	18.55	18.16	16.84	17.53	16.8	15.23	12.67
	Below Secondary	25.49	23.47	26.18	25.53	25.75	27.72	26.5	26.8	25.51	23.26
	Secondary/Higher Secondary	14.71	10.25	11.58	13.18	14.85	16.64	19.05	22.02	24.91	27.51
	Above Higher Secondary	2.89	2.57	2.79	2.94	3.47	4.67	6.23	7.11	10.66	20.44

Tables 38 and 39 in appendix provides the similar information in Rural and Urban area resepctively.

6 Consumption Inequality Series

Consumption can be more closely related to welfare of a person in a big developing country like India where large population do not own any wealth and have no income. 270 million⁵⁶ of individuals were living on under less than a dollar a month in India in 2010-11. The following analysis is based on NSS-Consumption datasets. Consumption inequality statistics, similar to wealth inequality statistics is used to analyse the distribution of consumption in the population. Consumption Inequality Series is produced for analysing the evolution historically. The flow of the chapter is similar to previous chapter for easy readability.

Zero/Low consumption values in the datasets is winsorized at one-percentile level of consumption, to provide minimum consumption level.⁵⁷ In all the analysis consumption is per-capita consumption which is total consumption divided by total population in the household.⁵⁸

The evolution of mean per capita expenditure (MPCE) level at Rural and Urban level is provided in Table 14.

⁵⁵There are 5 educational categories based on 1991-92 categorization in NSS AIDIS. “Illiterate” is the section with 0 years of education, “Below Primary” has upto 5 (< 5) years of formal education. “Below secondary” category have people with 5-8 years of formal education. “Secondary or Higher Secondary” implies 8-12 years of education and “Above Higher Secondary” implies > 12 yrs of education.

⁵⁶Planning Commission Press Note: July 2013

⁵⁷1 percentile level in the dataset is lower than the poverty line for every years.

⁵⁸In previous chapter on wealth inequality, per-adult wealth was used.

The nominal Rural MPCE in 1983 was Rs. 111 which has increased consistently in nominal terms to reach Rs 982 in 2009 (8.2% per annum in long run⁵⁹). At fixed price, the growth rate is 1.5% per annum. *Urban* MPCE is also consistently increasing in both nominal and real terms. It stands at Rs.1786 in 2009, which is 1.92 times rural MPCE. The long run growth rate of MPCE in nominal terms is 9.24% per annum and at fixed price level it is 1.82% per annum. The difference in the growth rate between urban and rural area explains the increasing gap between urban and rural level of consumption.

Table 14: Evolution of MPCE in Rural and Urban area

MPCE (in Rs)						
Year	Price Level 2010			Nominal Level		
	Rural	Urban	Ratio=Ur/Rur	Rural	Urban	Ratio=Ur/Rur
1983	809	1,186		111	163	
1987	776	1,209	1.56	158	246	1.56
1993	825	1,342	1.63	281	458	1.63
1999	897	1,578	1.76	486	855	1.76
2004	849	1,599	1.88	559	1,052	1.88
2009	928	1,786	1.92	928	1,786	1.92

Source: Author's calculation using NSS Consumption datasets. Design weights are used to estimate these values. For Rural area, Agricultural Labour Consumer Price Index (CPI), and for urban area, Industrial Worker CPI is used which is published by Reserve Bank of India annually.

MPCE across Caste Groups: Table 15 provides the MPCE for different caste groups. For every round of survey, the ranking of caste groups based on MPCE follows the order of *ST < SC < Muslim < OBC < FC* for years after 1999. For previous years, OBC information is not available. The order for previous years is *ST < SC < Muslim < Hindu*. In 2009, the MPCE of FC (Rs. 1684) was 1.45 times the MPCE of overall population. In the same year, the MPCE of OBC (Rs 1065) was at 0.92 times, Muslim (Rs. 1032) at 0.89 times, SC (Rs. 887) at 0.87 times and ST (Rs. 854) at 0.74 times of the MPCE of overall population. Interestingly, the ratio of the consumption of SC and ST with respect to the MPCE of overall population has remained stable in last 26 years. This raises a serious question if the Indian government policies targeting the SC and ST are enough. This finding is important for the debates over ongoing reservation issues in India.

There is no convergence across caste-groups in MPCE. Tracing the evolution of percentage difference (PD) i.e. $((MPCE_c - MPCE_{overall}) * 100 / MPCE_{overall})$ between MPCE of caste groups with the Overall Population over the years helps in understanding the convergence. It is simply a measure of how far the mean consumption of a social group is from the mean consumption of overall population. Table 15 provides the values of PD under the heading “% change with Overall MPCE” for different castes. Firstly the gap is positive only for FC and negative for all other caste groups. The gap is increasing years over years. We see that ST group always had 25-31% less MPCE than overall population. For SC group the difference has increased from -20.7% in 1983 to -23.5% in 2009. For OBC group, the difference is consistent at around -8% in all the years. Lastly for Muslims we see a decline of 3.5 pp from 2004-2009. This shows that the inter-caste-group disparities in MPCE level is actually diverging. The findings are also supported in NSS Reports (Report on MPCE across Social groups, Report No-544)

Table 15: Evolution of MPCE across Caste Groups: India

year/caste_group	MPCE						% change with Overall MPCE						
	ST	SC	OBC	FC	Hindu	Muslim	Overall	ST	SC	OBC	FC	Hindu	Muslim
1983	520	584			796	683	736	-29.35	-20.65			8.15	-7.2
1987	622	670			925	790	850	-26.82	-21.18			8.82	-7.06
1993	661	695			977	812	890	-25.73	-21.91			9.78	-8.76
1999	705	772	901	1324		899	988	-28.64	-21.86	-8.81	34.01		-9.01
2004	694	794	933	1446		938	1016	-31.69	-21.85	-8.17	42.32		-7.68
2009	854	887	1065	1684		1032	1160	-26.38	-23.53	-8.19	45.17		-11.03

Source: Author's calculation using NSS Consumption datasets. Design weights are used to estimate these values. The values are at price level 2010. For Rural area, Agricultural Labour Consumer Price Index (CPI), and for urban area, Industrial Worker CPI is used which is published by Reserve Bank of India annually.

The evolution of MPCE in different caste groups in rural and urban areas separately is provided in Table 40. The gap is negative for SC, ST and OBC. In long run, PD has remained almost stable for both ST and SC in rural area. In urban area, the PD has decreased for SC by 7 pp from 1983 to 2009 and decreased for ST. The improvement in MPCE in urban area for ST is impressive. The PD of Muslims in urban area is more than double than in rural area and in

⁵⁹26 years during 1983-2009

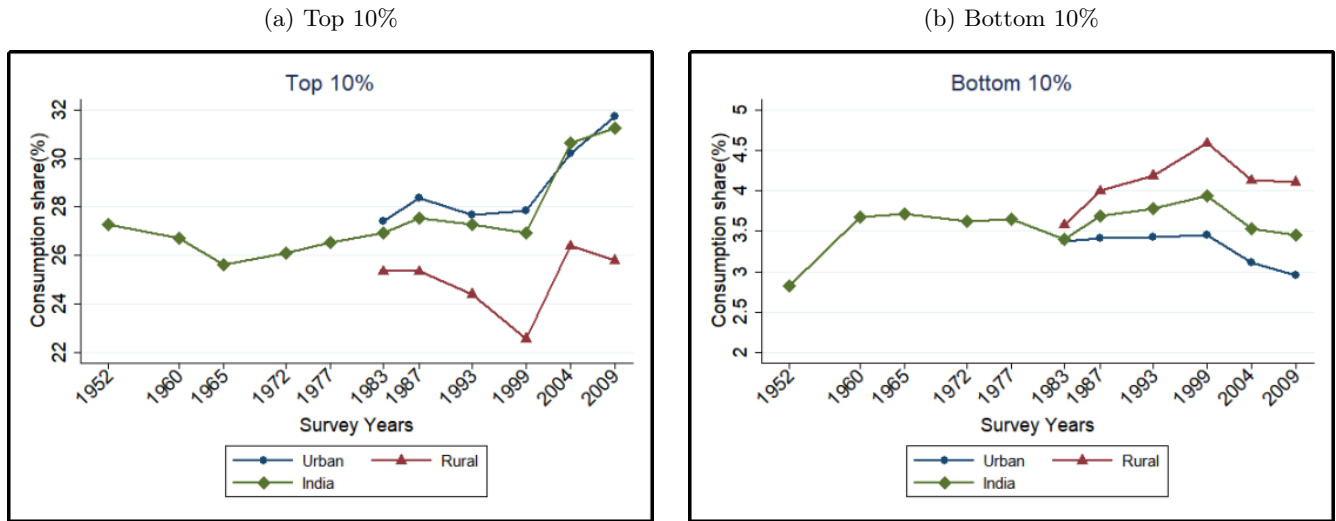
2009 there is a decrease in PD for Muslims. The decrease is mainly coming from rural area. The positive PD for FC has increased from 34% in 2000 to 45.2% in 2009.

6.1 Concentration of Consumption

6.1.1 Two Contrasting Ends

Fig 13a provides the consumption share of top decile population in India. In long run, the share of Top 10% population has increased from $\sim 28\%$ in 1951 to $\sim 32\%$ in 2009. The increase might be due to change in the methodology to capture the data. Similarly the bottom decile share of consumption in Fig 13b is also stable at 3.5-4% consumption share since 1951. In urban area, the share of top 10% is very close to all-India level. Rural area has relatively less concentrated consumption in top decile. There is a gap of 6 pp between rural and urban area in the share of Top 10%. Comparing with wealth share of Top 10% from previous chapter we see that consumption distribution more equitable. Top decile share in wealth is almost double that of consumption. However, the information on wealth and consumption is coming from different samples, so the two sets of population in top deciles are not same. Based on NSS AIDIS datasets which asks for the consumption too, Zacharias and Vakulabharanam (2011) finds overlap of only 50% in top quintile. In other quintiles, the overlap is less than one-third. There is no reason for the two population to be exactly same. The preference to save and the marginal propensity to consume play a role in the consumption behaviour of a person.

Figure 13: Consumption Share:Top 10% and Bottom 10%



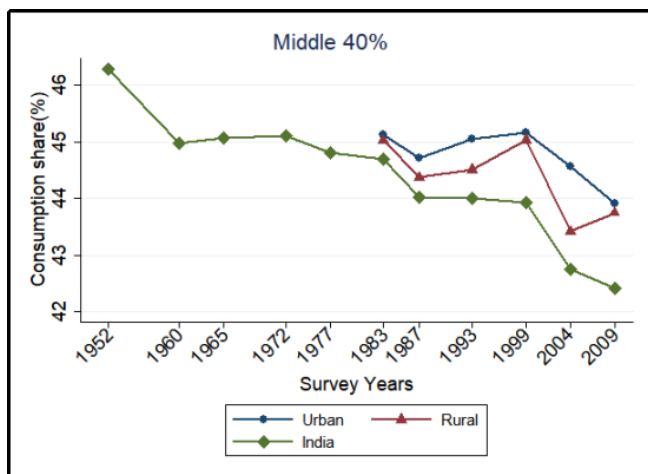
Note: 1951-1983 data has is used from Datt (1997). Post 1983, Author's calculation using NSS-Consumption datasets

6.1.2 Weak Base

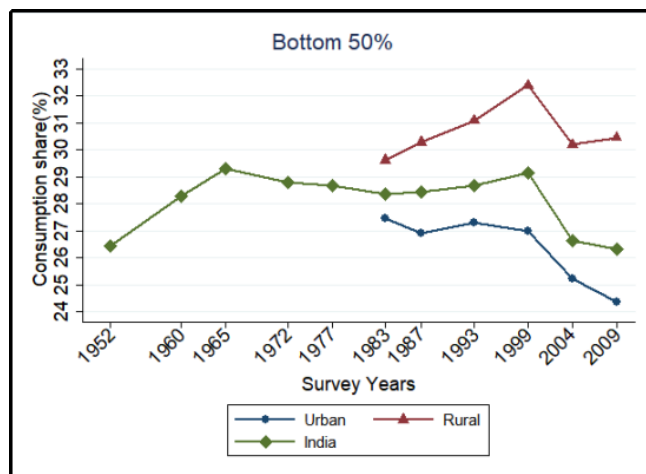
In long run, the share of Middle 40% section of the society has decreased from 46% to 42% at all-India level (Fig 14a) The share of Bottom 50% of the population was stable at $\sim 28\%$ till 1999. There is a dip next year and the share dropped to 26%. From the Fig. 14b we can see that steeper decline in rural areas than compared to in urban area. The consumption share of Bottom 50% is close to the consumption share of Top 10%. The stability observed in consumption distribution profile in long run is notable and is different from the trends observed in income or wealth distribution.

Figure 14: Consumption Share: Middle 40% and Bottom 50%

(a) Middle 40%



(b) Bottom 50%



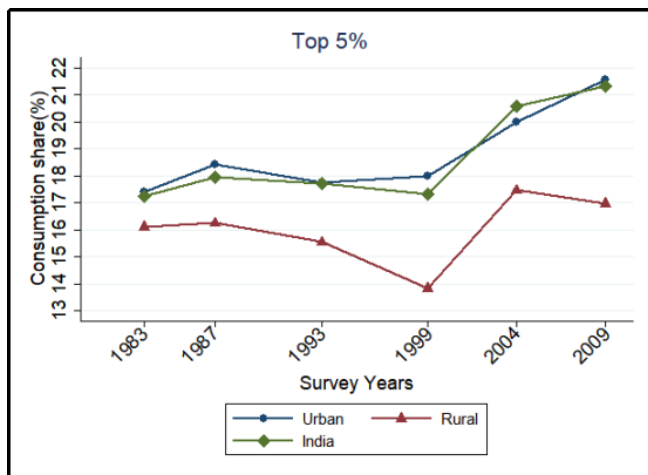
Note: 1951-1983 data has is used from Datt (ibid.). Post 1983, Author's calculation using NSS-Consumption datasets

6.1.3 Rich Topping

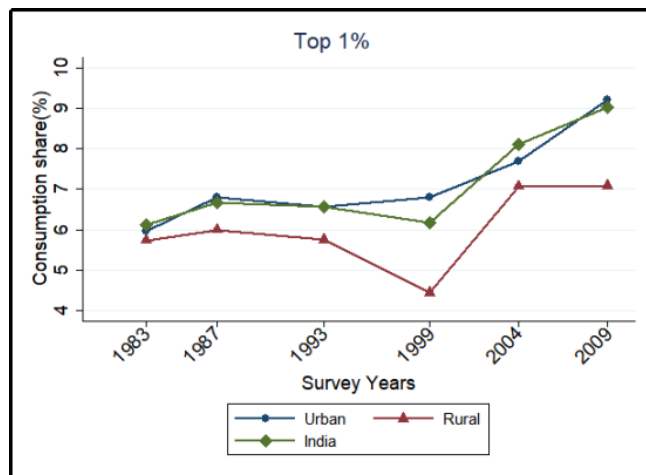
Top 5% consumption share has increased from 17% in 1983 to 21% in 2009. The increase at all India level is mainly due to Urban areas as we see in rural areas, the share of top 5% is almost constant. In Top 1% the share has increased from 6% to 9%.⁶⁰

Figure 15: Consumption Share: Top 5% and Top 1%

(a) Top 5%



(b) Top 1%

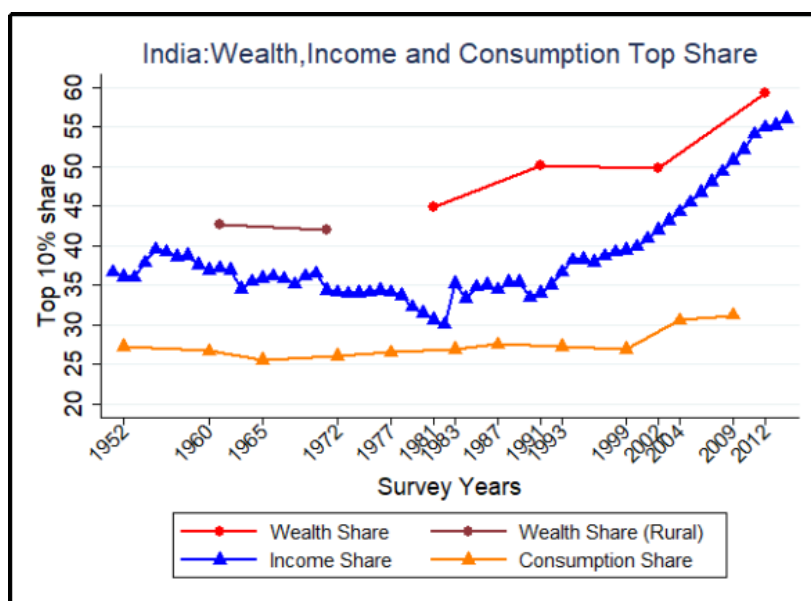


Note: Author's calculation using NSS-Consumption datasets

Discussion: I compare the combined wealth, income and consumption inequality series in India. Fig 16 presents the wealth, income and consumption inequality series since 1951. We can see from the graph that income curve falls between the wealth and consumption share. In this figure, wealth share is pre-corrected wealth share i.e. simply from the NSS AIDIS survey, hence it is a gross under-estimation. I intend to explore in detail the combined evolution of these three distributions.

⁶⁰The sample size for 1% in the dataset for all-India is 2.4k in 1983, 3.97k in 1987, 4k in 1993, 3.2k in 1999, 2.1k in 2004 and 1.6k in 2009.

Figure 16: Consumption Share:Top 5% and Top 1%



Note: Income inequality series is from Chancel and Piketty (2017).

6.2 Representation of Different Castes-Decile wise

This section deals with Representational Inequality defined in Jayadev and Reddy (2011) in Consumption, first at all-India level and then at Rural and Urban level separately. As in wealth inequality section, I look at the representation of different castes in different deciles and compare it with their population proportion.

6.2.1 All India

At all-India level, the proportion of population of ST, SC, OBC and FC in NSS consumption datasets are provided in Table 41. We see that ST population share is almost constant at around 8.7%, SC population share increased from 17.2% in 1983 to 20.3% in 2009. OBC is captured after 1999 survey and it increased from 35.7% to 41.8% in a decade. In this decade FC share decreased from 25.8% to 20.4%. Muslim share is also constant till 1993, and the steep decline in 1999 is because of categorization of OBC Muslims into OBC category. From 1999 to 2009, we observe a drop of 1.5 pp of Muslim category. The trend of the change in the population share of OBC, FC and Muslim is same as we saw in NSS AIDIS datasets. It further supports the evidence of categorization of castes from Muslims and FC to OBC.

Fig 17 shows the representation of different caste groups in top and bottom consumption deciles. The population share of ST in top decile has remained at constant at 2-3% during 1983-2009. It is 5-6 pp lower than their population share. ST representation in bottom decile has remained constant around 17% with an temporary increase in 2004 to 21.8%. The population share of SC in Top decile has remained constant at 6-8% which is 13-14 pp lower than their overall population share. In bottom consumption decile, SC population share (26-30%) is higher than its population share (20.3%).

The OBC population share in top decile is at 23%, 29% and 30.5% in 1999, 2004 and 2009 respectively which is 12.3 pp, 11.9 pp and 11.3 pp lower than their population share. Since the population share is changing in OBC due to categorisation of more castes into OBC category, it is not clear if the increase in top decile share is due to improvement in condition of pre-existing OBC's or inclusion of relatively well-off castes into OBC. Since the NSS consumption datasets are different cross sections i is not possible to do a panel analysis to capture the effect. In bottom decile the population share of OBC is almost equal to its population share.

The FC population share in top decile is at 53.5%, 47.1% and 45.8% in 1999, 2004 and 2009 respectively which is 27.7 pp, 21.3 pp and 25.4 pp higher than their proportionate share. The decreasing population share in top decile needs more scrutiny as it could simply be the counterpart of the re-classification into OBC as discussed above. In bottom decile the population share of FC is 18 pp, 15.3 pp, 14.4 pp lower than its overall population share in the years 1999, 2004 and 2009 respectively.

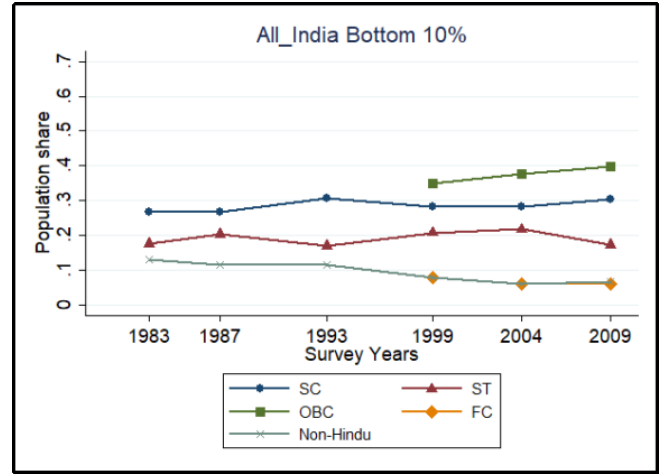
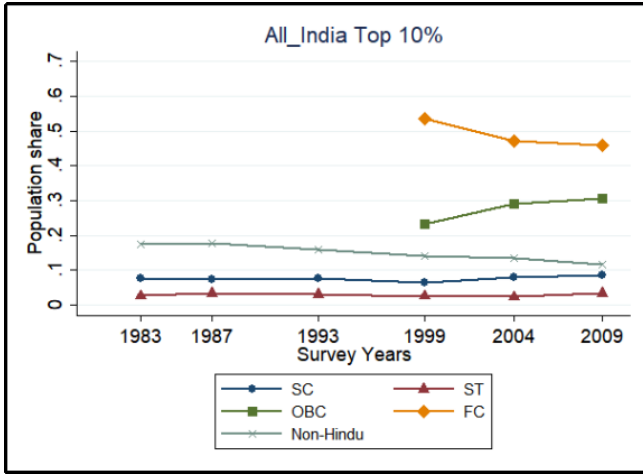
Muslim representation has to be seen in two phases. Before 1999 they are under-represented in top decile by 3.5 pp and after re categorization of OBC Muslims into OBC, the under-representation drops to 3.36 pp, 2 pp and 2.64 pp

in 1999, 2004 and 2009 respectively. In bottom decile, Muslims are consistently under-represented implying bulk of the population is present in Middle 40%. This also shows the poor economic condition of Muslims in general because even after taking out the so-called Other Backward Castes, the representation is poor in top decile.

Figure 17: Population Share of different castes in top and bottom deciles

(a) Top 10%

(b) Bottom 10%



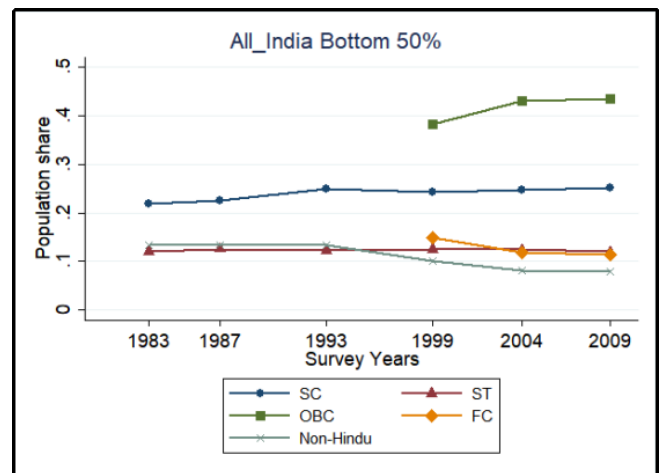
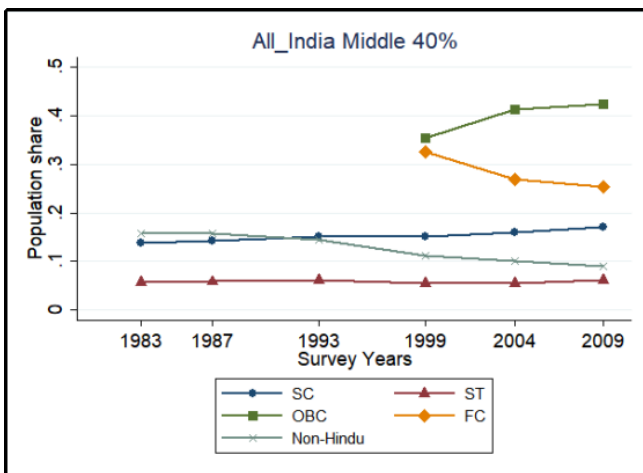
Note: Author's graph using NSS Consumption datasets. A higher population share in top decile than the overall population share for a caste makes it beneficiary of the inequality. Similarly if the presence of a caste is higher in bottom decile, it makes it victim of the inequality.

After analysing the Top and bottom decile, I look at the middle 40%(i.e. 7th – 9th deciles) and Bottom 50%(i.e. 1st – 5th deciles) to get an understanding of representation inequality in entire distribution of consumption. Fig 19 provides the evolution of the representation of different caste groups in Middle 40% and Bottom 50%. In Middle 40% ST representation has remained around 5-6% which is less than 2-3 pp than their population share. In Bottom 50% the representation has remained stable at around 12% which is 3pp higher than their overall population share. The representation of SC in Middle 40% is stable at around 3 pp lower and in Bottom 50% it has remained in 4.5-5 pp higher than their overall population share. The representation of OBC in Middle 40% is almost in line with their population share, where there is over-representation of 2 pp in Bottom 50%. The FC is over-represented in Middle 40% is higher by 5-7 pp with a decreasing trend.

Figure 18: Population Share of different castes in Middle 40% and Bottom 50%

(a) Middle 40%

(b) Bottom 50%



Note: Author's graph using NSS Consumption datasets.

6.2.2 Rural Urban Comparison

In Rural area, ST/SC/OBC population share is higher and in Urban area FC/Muslim population share is higher than in the population share at all-India level. FC population share is double in Urban area (32%) than in Rural area

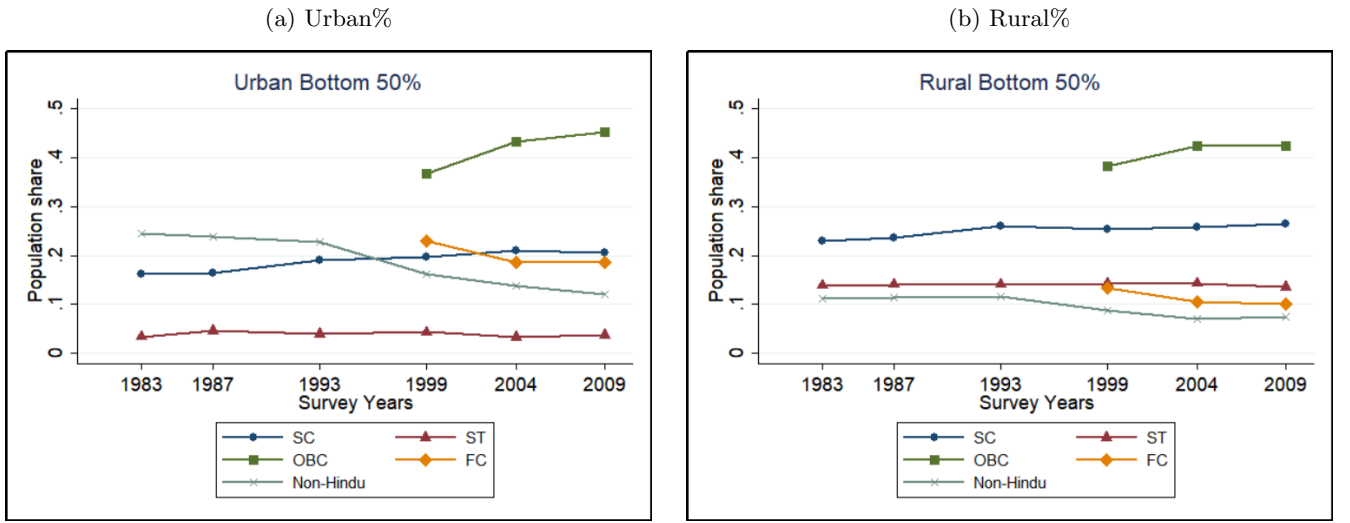
(16-17%). To compare in Rural and Urban area, I create a statistic of normalized Representational Inequality which is defined as below:

$$RI^{ycd} = RI_{Rural}^{ycd} - RI_{Urban}^{ycd} = \frac{Popshare_{Rural}^{ycd} - Popshare_{Rural}^{yc}}{Popshare_{Rural}^{yc}} - \frac{Popshare_{Urban}^{ycd} - Popshare_{Urban}^{yc}}{Popshare_{Urban}^{yc}} \quad (2)$$

for the year y and the caste c in decile d. $Popshare_j^{ycd}$ represents the population share of caste c in decile d in year y. $Popshare_j^{yc}$ is the population share of caste c in year y in j where $j \in (Rural, Urban)$. A positive RI^{ycd} will imply a higher representation in decile d in Rural area compared to Urban area in year y for caste c. The numbers are present in Table 42. I compare the rural urban analysis using this table and the figure which I will mention below.

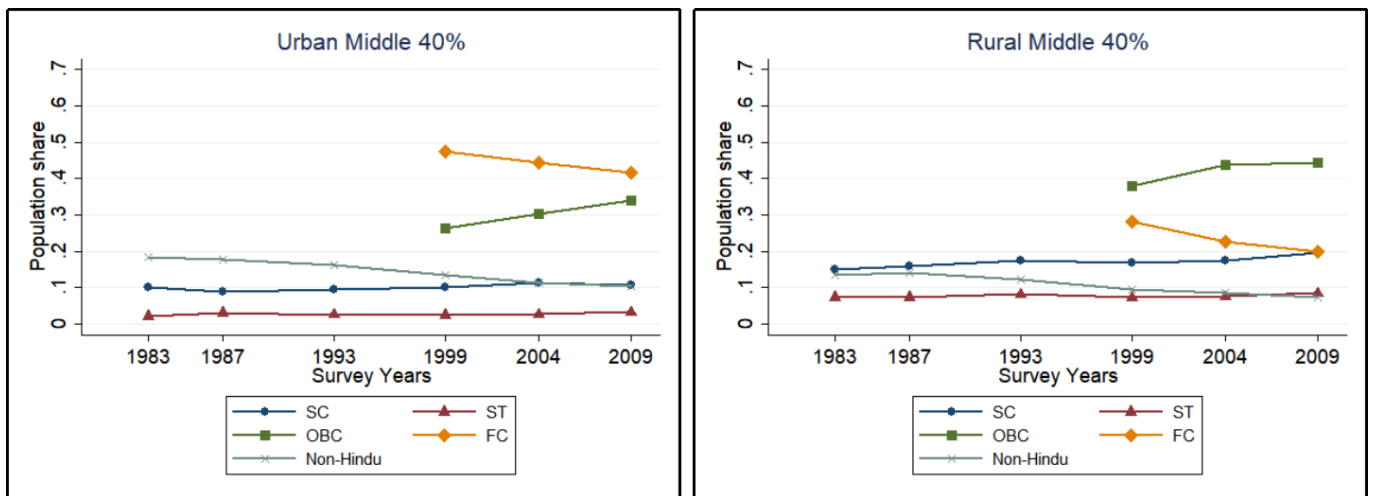
Bottom 50%: (Figure 19) shows the population share of different castes in Bottom 50% in urban and rural area. We see that rural area has higher representation (than Urban area) of ST in all rounds of the survey. SC has less representation in Rural area in all rounds of the survey. OBC has less representation in Bottom 50% in Rural area than in Urban area in 1999, 2004 and 2009. FC had lesser representation in Rural area compared to Urban area in 1999 but increased in 2004 and 2009 surveys.

Figure 19: Urban Rural:Bottom 50%



Middle 40%: Figure 20 shows the presence of different castes in Middle 40% urban and rural area. We see that Rural area has lower representation (than Urban area) of ST in all rounds of the survey. SC caste had same relative representation in 1983 in Rural and Urban area, but the representation declined in Urban area in later rounds of survey. In 2009 the difference is maximum. OBC has higher representation in Rural area than in Urban area in 1999, 2004 and 2009. FC has lesser representation in Rural area for all the rounds.

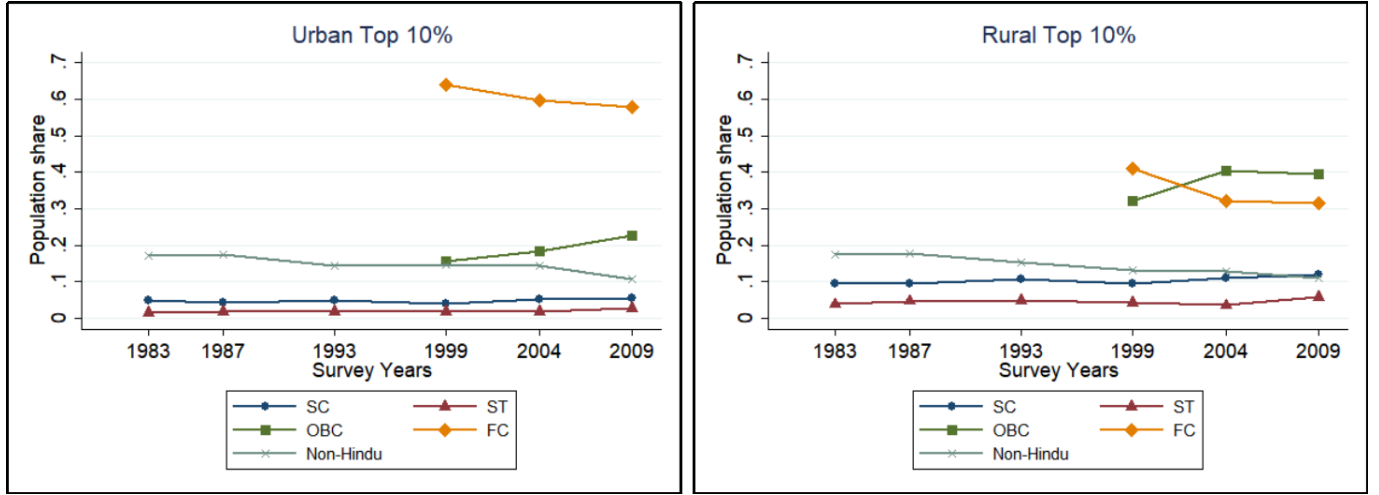
Figure 20: Urban Rural: Middle 40%



Top 10%: Figure 21 shows the presence of different castes in Top 10% urban and rural area. We see that Rural area

has lower representation (than Urban area) for ST in all the rounds of the survey. SC has higher representation in Rural area than in Urban area. OBC has higher representation in Rural area than in Urban area. FC has higher representation in Urban area for all the rounds.

Figure 21: Urban Rural: Top 10%



We can conclude that Urban area has remained more favourable to FC and ST than rural area whereas opposite is true for SC and OBC. We can also say that urban area attracts richer FC and ST, whereas rich community from SC and OBC remain in rural area.

Further in a given year, one can analyse the strength of representation among different castes in a given decile from the term $\frac{Popshare_{Rural}^{ycd} - Popshare_{Rural}^{yc}}{Popshare_{Rural}^{yc}}$. Positive value will denote more representation (than overall population share) and the magnitude will signify the strength. For example, in 2009, ST(-0.46) and SC(-0.46) are equally under-represented in Rural areas in top decile followed by Muslims and OBC. Only FC⁶¹ has positive sign in top decile. Similarly in Bottom 50%, ST is over-represented than SC followed by SC in Rural area. In Middle 40% Rural area, the strength of over-representation is higher for FC (0.24) than OBC (.04). On the other hand, SC(-0.22) is less under-represented than ST(-0.12) in Rural area in Middle 40% category.

6.3 Within Caste Inequality

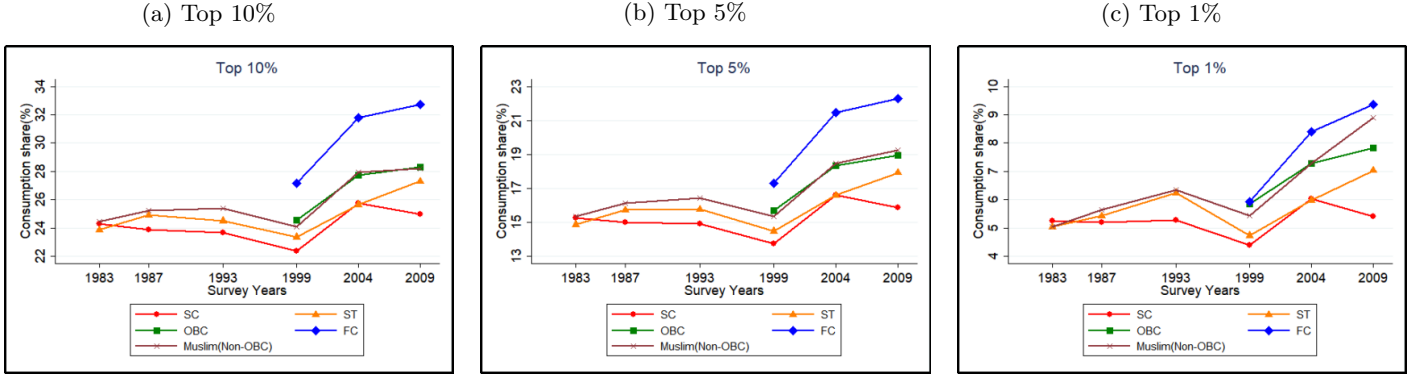
This section deals with analysing the inequality within every caste group. As in other parts of the sections in this chapter, we observe stability in within caste inequality too over the years in every caste. Generally there is slight increase in the share of top decile population in every caste group and corresponding decline in consumption share of Middle 40% and Bottom 50%.

6.3.1 Top Deciles

Figure 22 shows the share of Top 10%, Top 5% and Top 1%. FC has highest within caste inequality as we had observed in the case of wealth also. SC has the least within caste inequality. Every caste group saw a decline in the top shares in 1999 and then increase in later years. From 2004-2009, only within SC group there is a decrease in consumption share of top decile population.

⁶¹Others also has positive value- those who are not categorized above. Since the population share is low, this community is not compared.

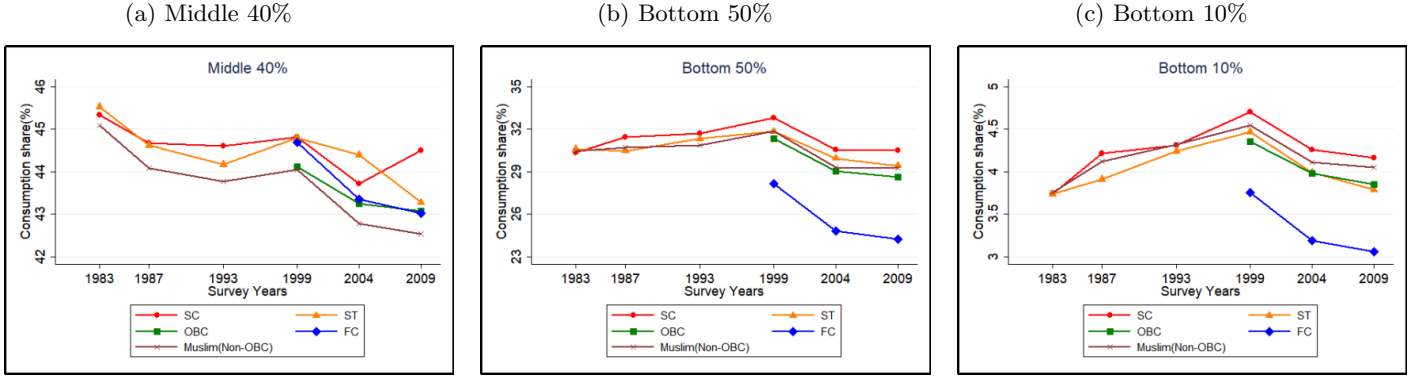
Figure 22: Share of Top Deciles Within Caste



6.3.2 Lower Deciles

Figure 23 show the consumption share of Middle 40%, Bottom 50% and Bottom 10% within different castes. FC is again an outlier with minimum share with its Bottom 50% and Bottom 10% population. For all other castes we observe an inverted U shape with a peak in 1999 when the consumption share of Bottom decile was the highest in entire period of 1983-2009. We observe a rise in the share of Middle 40% in SC, going from 2004-2009, which is the counterpart of the decline we observed above in its Top 10% population.

Figure 23: Share of Lower Deciles Within Caste



Discussion: NSS Consumption datasets provide a very stable consumption patterns both in terms of distribution in different deciles and distribution across different caste groups. There is unequal distribution but that inequality is quite stable. We do not find any kind of convergence in the relative consumption share of different castes. SC and ST group should have improved over time due to presence of positive discrimination. Lack of relative improvement show that India needs to do more to close the gap.

7 Inter Caste Marriage

Marriages in India are very different from France or other developed world countries. Hence I first outline some prominent characteristics of marriages in India. The generalized version which is presented here is not valid for all the communities in all the parts of India. Some of the marriage rituals are specific to religions and hence they vary in different religions. Many aboriginals or Scheduled Tribes in India have their own cultural practices of marriage. Marriage as a cultural practice is also different in different parts of the region. Inter-cultural transmissions among different societies, religions, linguistic groups etc. are constantly changing the way marriages are perceived. Modernization (or Westernization) has its own impact through the changing role of man and women as a couple. IT revolution is making its impact through spread of matrimonial web sites where marriages are fixed online! All this diversity and dynamism in the society makes any kind of generalization very challenging. My attempt here is to convey a broad idea of the marriage practices to the reader.

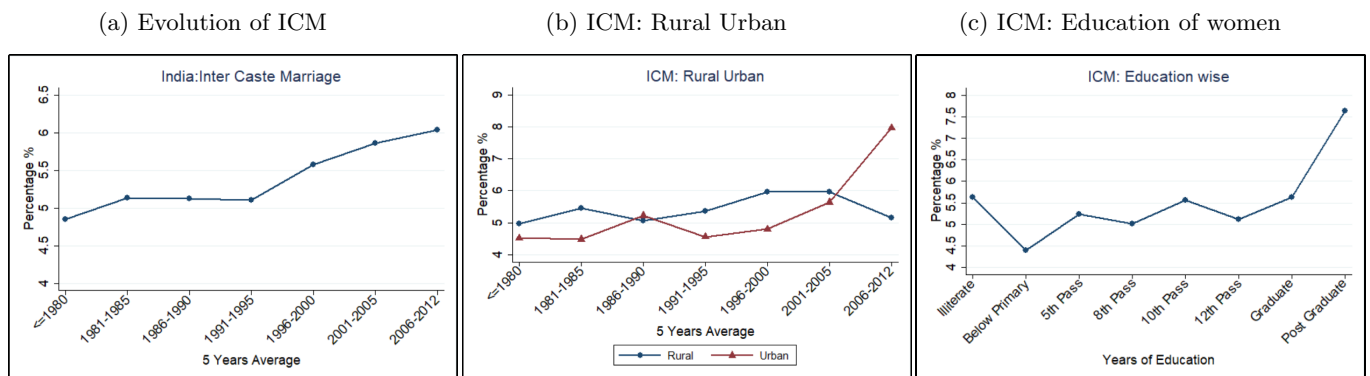
The role of parents in Indian marriages is very high. It would not be exaggeration to assume that all the marriages in India are fixed by parents. The decision making in choosing partners by oneself is minimal. There is no discrimination between female and male. Surveys like IHDS of NFHS captures this peculiarity. To illustrate, in IHDS 2011 survey, one of the questions asked to eligible married women is- "How long knew husband before marriage". 82%

of the women response was less than a month. On the question, “Who chose your husband?”, 95% responded with Parents or other relatives. This might sound surprising to many but fixing of marriages by parents is almost prevalent in all the religions, castes and regions. Such marriages are now being called Arranged marriages in colloquial terms. The other type of marriage is termed Love Marriage where a person chooses partner oneself. The role of youth to decide on their marriage is increasing but within the purview of arranged marriages (Banerjee et al., 2013).

Dowry is widely prevalent in entire country, even though it is illegal⁶². Dowry prospered in India from the arrival of modernisation -increasing average wealth and wealth dispersion- in the presence of caste system (Anderson, 2003). It was not the case a few decades ago. Dowry was limited to a few upper castes society in North India till 1980s. Due to economic, social and demographic transitions, it has spread into communities of South India and tribal populations of Central Indian states (Rao, 2006). Dowry is so rampant that it is considered as one of the main economic reasons for son preference in India (Basu, 1999). Patriarchy and patrilocal system ensure that the bride goes to the grooms house after marriage. Marriage is a widely accepted social norm. Not marrying is rare. The sense of continuing the family name from one generation to other is strong. Divorce is still a social stigma in the society and divorce rate is very low 1 per 100 marriages (compared to 55 for France, 46 in US)⁶³. Inter-religion marriages are rare at 2.1% (Das et. al 2011) After religion the next important identity for marriage is caste. The origin of caste endogamy probably is inspired from Hindu scriptural texts. Not all Inter-caste marriages (ICM) are prohibited in religious texts. “Anuloma” marriage (or Hypergamy)-wherein a man of higher caste marry woman from lower castes is acceptable. “Protoloma” marriage (or Hypogamy)- in which men of lower caste marry women of higher caste is not permitted (Anderson, 2003). However, in practice ICM is rare. I estimate the level of ICM’s using IHDS 2011 datasets.

The level of inter-caste marriage in India is at the level of 5-6%. Fig. 24 provides the level of ICM in different marriage cohorts. We see that almost similar level of percentage of ICM’s in different cohorts. This presents a very high level of assortative matching (AM) along caste, at a level of 0.95. It is 58% higher than the education level AM in France. The level of inter-caste marriages based on five-years moving average by marriage year has seen an increase of only 1 pp from 1980’s to 2006-12 which shows persistence. Further the change is observed after 1990. This depicts presence of a strong caste homogamy which hints towards evidence of strong own-caste preferences. Banerjee et al. (2013) has found horizontal preference, i.e. preferences for marrying within caste in rich educated middle aged cohort.⁶⁴ Looking at the ICM in rural urban regions separately, ICM’s are higher in rural areas except in recent years, where ICM’s in urban areas has jumped to 8%. Looking at patterns of ICM with different education level we see an increasing trend with education level. Post Graduates have 7.5% ICM’s but they form only 1.7% of the total population. However, trends do not reveal the causality. A detailed econometrics analysis using the same dataset found no statistical significant impact of education, urbanization and caste group on ICM rejecting modernisation theory (Ray, Chaudhuri, and Sahai, 2017). This is in contrast to findings from US where studies have shown increase in inter-faith marriages among educated cohorts (Qian, 1997).

Figure 24: Inter-Caste Marriages: IHDS II



Note: Author’s calculation using IHDS 2011 datasets. ICM is defined on the survey question asked to eligible ever married women- “Is your family the same caste as your natal family?”

NFHS datasets, another nationally representative survey in its 2005-06 (III survey) sampled 99k ever-married women aged 15-49. The survey has collected separate narrow level of castes (jatis) for husband and wife. However, unless the caste categorization is perfect, using the name of caste to designate inter-caste marriage can lead to over estimation due to many castes with same hierarchy but different names. However, one can look at administrative level castes-

⁶²Dowry Prohibition Act, 1961

⁶³<https://sites.google.com/site/indiandivorcelaws/divorce-rate-in-india>

⁶⁴Precisely speaking there is difference in preferences and actually observed marriages due to many exogenous factors in society

namely SC, ST, OBC and Others (FC). Das et. al (2011), using NFHS- III (2005-06) data has analysed the level of ICM at above defined caste groups. Their estimated level of ICM is at 10% level in India with varying levels in different states (e.g 22% in Punjab, 21% in Kerala, 2.5% in Rajasthan and Tamil Nadu). The higher level of ICM from NFHS datasets could be due to two reasons. Firstly the difference can arise from the different definitions of inter-caste marriage. IHDS survey asks a direct question to eligible women about inter-caste marriage, which has the potential to be interpreted as narrower version of caste (jati). I check the level of inter-caste marriages among the broader caste groups using IHDS dataset and for no caste group the ICM rate is above 10. Secondly difference is that NFHS-based study is restricted only to Hindu religion couples. I intend to extend the analysis of ICM using the narrow definition of castes from NFHS in my future research.

Discussion: i) The level of ICM in India is very low and the rate of its increase is also very slow, but societal changes are usually a slow process. The rate indeed is context and society dependent. I draw a parallel with Inter-ethnic and inter-racial marriages in USA- which has seen an increase from 3% in 1967 to 17% in 2015 among newly wed couples⁶⁵, i.e. 14 pp change in 50 years. Compared to this ICM rate of increase in India is much lower at 1% in 30 year. There are obviously many differences⁶⁶ and hence no conclusion is sought after from this comparison. Several other cultural changes are taking place in the realm of marriages within Indian, like spread and inflation of dowry in different communities, change in kinship structure in South India (fewer marriages within same family). They all probably have faster rate than the rate of ICM.⁶⁷ Interestingly there are central (Inter-Caste Marriage Scheme 2018) and state government (for example Anjugam Ammiar Ninaivu Intercaste Marriage Assistance Scheme 1967 of Tamil Nadu) schemes with monetary benefits to promote inter-caste marriages, but they are hardly publicised. The reason may lie in that a large part of the society do not see caste and its associated rules as regressive Srinivas (1955). The issue of ICM is not even covered in Indian school textbooks, even though negative aspects of caste-based discrimination are covered.

ii) It would be interesting to study patterns of ICM from gender perspective in different castes. Women historically have been discriminated against in investment towards their education, in all the castes groups. All the survey show less educational attainment among women than men in every caste group. However, women's participation in education in country is improving with higher enrolments. As per Government Report 2008, Gender Parity Index⁶⁸ in primary schools have increased from 0.22 in 1950 to 0.88 in 2005 (Nisha Nair, 2010). In coming years the gap is expected to reduce in higher education too. This has the potential to bring attitudinal changes in women (and overall society) in demanding more autonomy in choosing the partners of their own marriage. This has the potential to increase ICM in the society. The studies on US marriage market has found an increasing level of inter-racial/ethnic marriages with education (Qian 1997). Ray, Chaudhuri, and Sahai (2017) found a positive and significant impact of husband's mother's education on ICM. The causal impact of own education on ICM is still not established in India. Lack of good data on inter-caste marriage could be the reason behind this.

I want to highlight another channel from increasing of women education in India, which can impact rate of ICM. This is through the interaction mechanism between men and women (simply the process of meeting someone) in places like higher educational institutions (or in labor market-where selection is educational/merit based). Assume a differential trajectory of educational attainment curve for women from different castes. First FC group women will see an increase in higher education enrolment. Given 20% FC and say 60% other backward castes (SC, ST, OBC including only Hindu religion) the probability of interaction of FC woman to a non-FC man will be ~ 3 times higher than⁶⁹ interaction with a man from FC. If we analyse at granular level of caste system (jati), then the probability of meeting a person from same caste is much lower- for example a Brahmin girl meeting a Brahmin boy vs Brahmin girl meeting non-Brahmin boy will be lower than 3. If more interaction is a way to fall in love and eventually marry, then one can expect a higher proportion of FC women marrying outside their caste. This will happen first in FC and then eventually percolate to other castes. There is another force which can counter the above described process. It comes from the FC men, who will now encounter more FC women in educational institutions. This may lead to an increased effort from them for interaction. External factors like individual preferences (especially women), rigidity of parents etc. will ultimately decide the level of inter-caste marriages. But if the role of women in choosing their partners increases with time as observed in Banerjee et al. (2013), it will lead to increase in the number of ICM in FC women in first phase.

iii) We can apply the similar logic at religion level instead of caste, i.e. educationally backward Muslims (in ranking of

⁶⁵<http://www.pewsocialtrends.org/2017/05/18/intermarriage-in-the-u-s-50-years-after-loving-v-virginia/>

⁶⁶To cite one- there were miscegenation laws- forbidding people of different races from marrying were abolished in 1967. It is clearly not a case in India, as India had never such laws. Within caste-marriage is enforced by societal pressure.

⁶⁷Unfortunately, the data to measure dowry or other cultural practices is very rare which has been highlighted by many of the works cited above.

⁶⁸ratio of Girls Gross Enrolment Ratio to Bous Gross Enrolment Ratio. Gross Enrolment ratio is the total proportion of children enrolled in schools in a given age-cohort

⁶⁹The probability will be lower than 3 as there also exists a different trajectory of educational attainment among men from different castes. This can be seen by the order of education outcomes $ST < Muslim < SC < OBC < \text{OVERALL} < FC(Non - Brahmin) < FC(Brahmin)$

educational outcomes they have second worst outcomes) in near future will catch up with overall population, via an increase in enrolment of both men and women in higher education. Since there is differential trajectory of educational attainment curve for men and women (Table 46), we can expect first the increased enrolment in men. This increase in Muslim men's enrolment in higher education will lead to their increased interaction with non-Muslim women. Since Hindus are 78% of the population the probability of interaction with a Hindu women is 4-5 times more than interaction with Muslim woman. Also at later time period more enrolment of Muslim women will follow the similar argument. The claim of love-jihad (a term describing alleged campaigns by Muslim men to target women belonging to non-Muslims communities for conversion to Islam by feigning love)- started in 2009 which has become a topic often covered in news media, could simply be an inter-play of demographics and differential educational trajectory in different communities and in different sexes.

7.1 Assortative Matching: IHDS 2011 dataset

I create dataset of married couples using IHDS 2011 dataset by restricting the age of couples between 15-60 years. I remove the couples where even one of the member in couple is retired, students or non-workers. Though I do not remove housewives from the dataset as more than 70% women have identified themselves as housewives. The final dataset is left with 34,713 couples representing 204 million couples at all-India level using the survey design weights. For all the analysis, weights are used. Couples are married couples.

7.1.1 Educational Characteristics of Partners

Table 16 provides for the educational profile for married couples. Rural couples form 68.6% of total couples. The average age of men is 40 years and women 35 years. There is clearly age hypergamy (husbands are older than wife) in the society with an age difference of 5 yrs. Age hypergamy is almost at similar level in both Rural and Urban areas. There is a big chunk of population with no education in the dataset. 22% of men and 39.6% of women have zero years of education. Further 9% of men and 7.2% of women have less than five years of education. Rural area have higher illiteracy. For simplicity, I will use the term "Low Level Education" for this two categories, "Medium Level Education" for the next two categories (5-8 years of education) and "High Level Education" (more than 8 years of schooling). 36% men and 32% women have Medium level education. 32% men and 22% women have High level of education.

Table 16: Partners: Educational Profile

EDUCATIONAL CHARACTERISTICS						
	TOT POP		RURAL		URBAN	
	Men	Women	Men	Women	Men	Women
N (Total Samples)	34713	34713	65.58%	65.58%	32.86%	32.86%
Total Couples	203,933,402	203,933,402	68.64%	68.64%	31.36%	31.36%
Age	40.1	35.1	39.57	34.71	41.25	35.96
Education (%of Population)						
No Education 0	21.97	39.61	27.06	48.28	10.84	20.65
Less than Primary	9.06	7.21	10.22	7.86	6.52	5.78
5th Pass	9.28	8.97	10.27	9.33	7.12	8.17
8th Pass	26.69	22.68	26.66	21.12	26.77	26.09
Secondary	14.33	9.9	12.6	7	18.09	16.22
H.Secondary and Diploma (<3yrs)	9.11	6.36	7.39	4.26	12.89	10.96
Bachelors(BA,Bsc,Diploma 3+)	6.88	3.88	4.1	1.6	12.97	8.87
BTech,MBBS,MD,CA,PhD	2.67	1.4	1.7	0.55	4.79	3.26
Total (%)	100%	100%	100%	100%	100%	100%

Source: Author's calculation using IHDS 2011 datasets. Design weights are used to estimate these values. Educational categories are defined in Data section.

7.1.2 Employment Characteristics of Partners

Table 17 provides the detail on employment characteristics for both men and women. Occupation types and Primary Activity status are present for more number of men than for women. Under *Primary Activity Status* we see 44% of men and 15.3% of women in Urban area have salaried or professional status. In rural area only 11% of men and 8.5% of women have Salaried status. These percentages are out of those couples for which information is available. Indeed very few women (22.7%) and almost 93% men have this information present.

Occupation Type: Almost all the men have some occupation type compared to only 32% women having some form of occupation. 40% of the men are non-agricultural labourers and 30% have occupation related to agricultural activity.

In comparison 31% of women (for whom occupation type is provided) are non-agricultural labourers and 51.9% of women are involved in agriculture related occupation. The representation in Professional class is almost same at 5% for both men and women. In rural area as expected higher percentage of men and women are involved in agriculture and we observe a difference of 20 pp with more percentage of women working as farmers/cultivators. This shows high contribution of women in agriculture sector in India. In urban area, higher share of both men and women are involved in Professional (16% women and 8% men), Sales (13% women and 27.8% men) and Clerical occupations.

Wage Earnings for only 71.3% men and 26.9% women is present in the data, with average annual wage income of men at Rs 66k and women at Rs. 22k. The annual mean wages for urban men is 2.63 times the annual mean wage of rural men. Similarly urban women earn 3.7 times more than rural women. Based on imputed wages, the average annual wage goes down as I assign wages to people who are not in wage employment. Average wage of men at all-India level drops to Rs 55k for men and 19.7k for women which is a drop of 16.6% for men and 11.2% for women.

Table 17: Partners: Employment Profile

EMPLOYMENT CHARACTERISTICS OF PARTNERS						
	TOT POP		RURAL		URBAN	
	Men	Women	Men	Women	Men	Women
Primary Activity Status						
%age of Total Couples Covered	92.69	22.71	90.89	26.11	96.63	15.29
Salaried/Professional	21.86	15.27	11.07	8.53	44.06	40.45
Small Business/Artisan	15.73	7.05	10.83	4.5	25.81	16.58
Cultivators	24.23	20.07	34.67	24.99	2.74	1.7
Non Agri Wage Labour	26.49	23.36	27.13	21.19	25.17	31.49
Agri Wage Labor	11.69	34.24	16.3	40.79	2.22	9.78
Total (%)	100%	100%	100%	100%	100%	100%
Occupation Types						
% of Total CouplesCovered	99.93	31.8	99.94	37.97	99.92	18.28
Professional	5.08	5.13	3.61	2.76	8.28	15.9
Admin,Exec,Managers	1.26	0.13	0.44	0.05	3.06	0.52
Clerical	6	2.54	3.13	1.59	12.28	6.84
Sales	14.88	4.37	8.97	2.47	27.81	13.01
Service Providers	3.96	4.85	2.97	2.2	6.11	16.86
Farmers,Cultivators	29.36	51.89	40.45	60.21	5.07	14.08
Labourers (Non-Agri)	39.46	31.09	40.43	30.72	37.39	32.8
Total (%)	100%	100%	100%	100%	100%	100%
Earnings						
% with positive income	71.3	26.88	71.86	32.23	70.08	15.18
Average Annual Income	66179	22182	44061	15000	115810	55547
Imputed Earnings						
% with positive income	99.09	99.86	99	99.89	99.29	99.78
Average Annual Income	55165	19702	36748	11735	95352	37156

Source: Author's calculation using IHDS 2011 datasets. Design weights are used to estimate these values. Employment categories are defined in Data section. Percentages are calculated out the total population for which data is present. So 21.86% of men have salaried/professional status out of 92.69% men with primary activity status provided in the data.

Educational, Employment and wage characteristics for couples in different religion and caste is provided in Appendix (Table 45 & 46).

Comparison between ICM and non-ICM couples: We see from the table 45 , 5.3% of the couples are ICM couples. ICM couples are slightly younger. They have similar level of education and also have similar occupation profiles. Average annual wage based on given earnings is 22.7% higher for ICM women and 3.4% lower for ICM men than non-ICM members. Based on imputed earnings the gap between ICM and non-ICM women earnings drop to 7.5%.

Comparison between couples from different caste groups⁷⁰ Table 46 provides the education, employment and earnings information along different caste groups. Age hypergamy is almost same at ~ 5 yrs in all castes, even if Dalits (SC) and Adivasis (ST) couples are relatively younger in age than Brahmins/FC. As we go from Brahmins to ST group (left to right in the table) we see a decreasing percentage of men and women with high level of education and an increasing percentage of men and women with low level of education. Muslims men and women have second worse educational outcomes below SC but above ST.⁷¹

⁷⁰A slight change in definition of caste groups. Here I give priority to religion over caste group in OBC. It means Muslims who are OBC are placed into Muslims and not OBC. This categorization is logical while studying marriage.

⁷¹The proportion of Muslim in SC, ST is very low as we see from Figure 28.

Similarly the proportion of men and women in salaried class decreases from Brahmin to ST caste group. The trend reverses in terms of reporting positive income from Brahmins till ST. 46% of ST women have reported some wage compared to only 10% in Brahmins. More women from ST and SC are employed in some form of occupation than women in FC. The hope of finding something different from caste hierarchical order, shatters if we look at average annual earnings. There is a strict decreasing trend in average annual wages from Brahmin (Men- Rs. 127k, Women-Rs. 66k) to ST (Men- Rs. 42.7k, Women- Rs. 16.4k) for both men and women. The resultant low wages are simply a manifestation of poor educational outcomes and more presence in low paying jobs. 63.5% of women in ST are involved in agriculture compared to only 21% from FC. In both men and women, the proportion involved in agriculture increases from Brahmins to ST. The findings are consistent with findings from NSS consumption datasets found in (Deshpande and Ramachandran, 2014). They too find FC on top, OBC in middle and SC/ST at bottom in terms of education, health and consumption.

The position of Muslims come closer to SC in terms of age. In terms of education outcomes, Muslim men and women fall between SC and ST. In terms of annual earning wages, Muslim men earning falls between OBC and SC and Muslim women falls between FC(Non-Brahmin) and OBC. Ranking of Muslims based on imputed earnings is similar to what we found in education.⁷²

7.1.3 Correlation Coefficients

In this subsection, I provide the results for estimated correlation coefficients and comment on the level of AM in Indian married couples. We saw in previous sections that in Indian marriages, ascribed status like religion and caste are very important. Here I check the role of achieved status in the marriages. In real life scenario, once a person attains the age of marriage from parents' and society standard, parents and other elders start searching his/her partner in their community. What are the characteristics a parent look for while searching the partner for his/her son/daughter? This subsection will try to find.

I estimate Spearman, Polychoric and Pearson correlation coefficients which are described in detail in Appendix 9.3.5. The estimated coefficients are provided with 95% confidence interval. Higher values of correlation coefficients will imply more AM along that attribute.⁷³

Education:

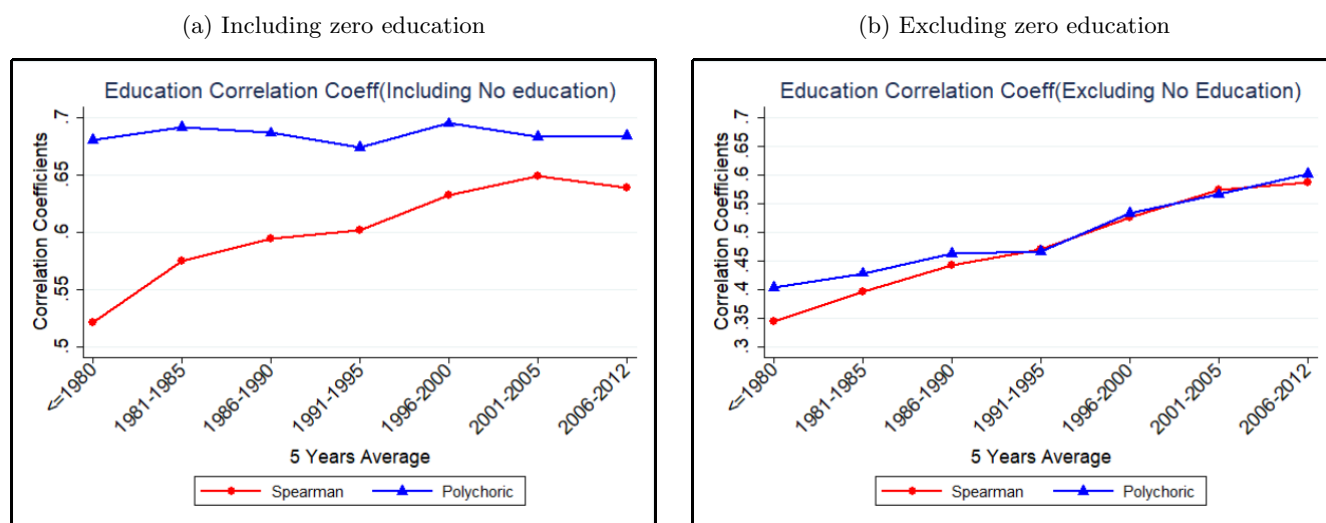
Own Education- Table 47 presents the estimates of AM for own education- for overall population, split into Rural and Urban areas, ICM and non-ICM, 4 largest Religions and the caste groups. The correlation coefficient for overall population is .63 for Spearman and .64 for Polychoric. The correlation is quite high and is around the level what one finds in developed country like France. The level of coefficients is ~ 0.6 in all caste groups or religion. Rural area has 11.9% lower coefficient than urban area. There is no difference in ICM and non-ICM couples. One of the possible reasons for this high level correlation is matching of couples having zero education. I re-estimate the coefficients in the second block (of the same table) by removing the cases when both members of couple have zero education years. As expected, the coefficient for overall population decrease by 20.3% to 0.5. The decrease is higher in rural area at 31.3% than in urban area at 9%. Muslims population correlation coefficient decreases the highest among different religion by 36% to 0.38. Among different caste groups, exceptionally, the correlation coefficient increases for Brahmin group, implying there are more assortativeness in higher educational categories than at zero education level. For all other caste groups, there is a drop in coefficient. The coefficients drop from Brahmins towards ST in the table. The AM is lower in SC, ST, Muslim at 0.38 which increases in OBC to 0.45 and in FC to 0.6. The decreased coefficients in ICM and non-ICM couples is almost same even after excluding zero education. This supports the claim made in theoretical paper Banerjee et al. (2013) that there is almost no cost attached to marry within caste.

Evolution of Own Education Assortative Matching Fig 25 shows an increasing trend in the coefficient with time. There is an increasing trend in Spearman coefficients for total sample and both Spearman and Polychoric, after excluding zero education cases. In the marriages before 1980, the education level correlation was minimal 0.35-0.4 which increased to 0.6 in marriages after 30 years. There is an increasing observed AM in the society along education. We can't comment on the preferences from this graph. Some of the increase in the AM is simply because the educational level in the country has increased. AM along different education level will be a better parameter to look at.

⁷²The poor education and economic conditions have been extensively enumerated in government set committee reports- Sachar Committee Report, Rangnath Mishra Committee report.

⁷³I use the term AM instead of correlation coefficients.

Figure 25: Evolution of Education Assortative Matching



Note: Author's calculation using IHDS 2011 datasets. The different marriage cohorts are made to keep enough samples in each one of them. The underlying population is almost constant in different cohort groups.

Parents Education- Table 48 and 49 provides for assortative matching at parental level of education- fathers' and mothers' respectively. The removal of zero-education (i.e. where both fathers'/mothers' education is zero) cases reduce sample size to almost half in case of fathers' education and one-third in case of mothers' education. The correlation coefficient is of the order 0.11-0.13 between fathers' education for the overall population. In Urban areas, the correlation is higher at 0.25, whereas in rural area there is no correlation. Among different caste groups, Forward castes (Brahmins and non-Brahmins) have almost similar coefficient at 0.27. There is almost no correlation in mothers' education in overall population. In Rural areas it is negative at 0.16-0.19. The low level of AM between couple's parents shows that they do not care much about their own education. Probably household wealth is more important at parental level. A high level of household wealth signals prosperity after marriage for bride which is the main concern of bride's parents. It also signals potential to give more dowry, which is an important factor for groom's parents. Unfortunately the dataset only contains present household level income and not the time when they were married.

Wife Education and Husband's Father: Since most of the marriages are arranged marriage, i.e. fixed by parents, I check the correlation between education of wife and her Father in law (or husband's father). Table 50 provides the correlation coefficients. Since a large proportion of women and husband's father have zero education, the correlation coefficient after removing the zero education cases is more informative. At all-India level the correlation is at 0.23 with a marked difference in rural and urban areas. In Urban area the correlation coefficient is at 0.36. In Rural area it is almost zero. Looking at different caste groups, the highest correlation is in Forward caste at 0.37-0.4. There is no correlation in wife and husband's father among Dalit (SC), Adivasi (ST) and Muslim communities. The decline doesn't necessarily explain the preferences because the availability of educated women also declines with SC, ST and Muslim as seen above.

Husband Education and Wife's Father: In Table 51, I check the correlation between husband and his father-in-law (i.e. wife's father). Post exclusion of zero education cases (i.e. where husband educ=wife's father educ=0), the correlation stands at 0.33 which is higher than previous the correlation we saw before. Since usually it is the girl's father who approaches the groom's family in arranged marriage he has the first mover advantage and probably this is leading to higher correlation. The preference of higher educated groom will be higher due to strong correlation in education and income/wealth. The correlation is quite high in Forward Caste at 0.4 and it declines to 0.27 in OBC to 0.19 in SC and 0.14 in ST.

Occupation type:

Table 53 provides the estimates of AM on partner's own occupation. The Spearman coefficients are smaller than Polychoric. At all-India level, the coefficient varies from 0.35-0.434. It is higher in Urban area by 38% than in rural area. The Spearman coefficient is higher by 20% in ICM couples compared to non-ICM couples. The higher level of correlation in Adivasi is due to a large population employed in agriculture related occupation.

Annual Wages:

One of the survey question asks women to compare the economic status of natal family to husband's family and 74% women feel that they marry to same economic status family. 16.5% feel that their natal family was economically better off and rest 9.4% feel that their husband's family is better. With this in mind we look at the correlation coefficient in

the data.

Gross Annual Wage Earnings: Table 52 provides the estimates of economic assortative matching. The first part is for the cases where both partners in couples have non-zero wages. The Pearson correlation is high at 0.7 and Spearman at 0.52. Expectedly in urban area the correlation is higher than in rural area. There is almost no difference between ICM and non-ICM couples, which again shows the no-cost of marrying within caste (Banerjee et al., 2013). Muslims have lowest coefficient among all religions (16% less compared to Hindu by Pearson coefficient and 32% less by Spearman coeff) and caste groups (21% less compared to Adivasi (ST) by Pearson coefficient and 45.7% less by Spearman coefficient).

Imputed Gross Annual Wage Earnings: The correlation (0.44-0.50) drops compared to before, but we have larger sample here. Urban area still has higher correlation. ICM couples (0.47) have higher Pearson correlation coefficients than non-ICM couples (0.44). Here we see almost similar level of AM in different castes except Muslims community with lower coefficient (0.34 Pearson coefficient).

7.1.4 India France Comparison

Table 18: India France Comparison: Assortative Matching

	France		Highest Degree India (including zero educ)		India(excluding both zero educ)	
	Polychoric	Spearman	Polychoric	Spearman	Polychoric	Spearman
Own Education	0.593	0.559	0.64	0.63	0.51	0.503
Father's Education	0.506	0.437	0.611	0.487	0.133	0.113
Mother's Education	0.476	0.401	0.695	0.512	0.001	0.005
	Gross Wage Earnings					
	France	India (Survey non-zero earnings)	France	India (Imputed Earnings)	France	India (Imputed Earnings)
	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman
Own Earnings (including 0)	0.175	0.175				
Own Earnings (excluding 0)	0.31	0.269	0.64	0.63	0.439	0.502
	Occupation					
	France	India (Occupation)	France	India (Primary Activity Status)	France	India (Primary Activity Status)
	Polychoric	Spearman	Polychoric	Spearman	Polychoric	Spearman
Own Occupation	0.531	0.453	0.434	0.352	0.728	0.649

Source:The France data comes from Frmeaux and Lefranc (2017).

I compare the educational and earnings level assortative matching between India and France. For France, I use the values estimated in (Frmeaux and Lefranc 2017). We see the level of assortative matching in India is quite comparable to France in own education. In fact including all the couples the correlation is higher in India (.64) compared to .59 for France. Since 22% of men and 40% of women have zero education in the data, this is driving the correlation coefficient high. The correlation coefficient in India (.51) is lower than France if excluding the zero education population from Indian data. The AM level at social origin (parents' education) is higher in India, if we make comparison by keeping all the samples in India. However, excluding zero education cases from Indian data, the coefficients are much lower in India (.113) than in France 0.51. Further the correlation in annual wage earning is higher in India among the couples where both earn. It could be attributed to the fact that income hypogamy (act of marrying lower) for women is socially looked down upon.

8 Conclusion

Caste still plays important role in Indian society. The repercussions of past injustice towards lower castes are now becoming more visible with more information coming out in public. Economic outcomes do not see a converging trend. The relative growth of lower castes is either stable or declining. Probably the most worrying aspect for any Indian policy-makers is poor educational outcomes for lower caste population. This means that in coming future the condition will not improve either. The privatization of education is increasing the cost of education and the requirement of skills in employment is ever increasing. This research work has put together multiple nationally represented datasets to highlight some of the important issues of Indian form of inequality which has its origin in social structure of caste. Economic inequality is working as a skin for caste inequality. The ordering of averages based on any economic or educational parameters of different castes follow the caste hierarchy. The measures of positive discrimination is not working on average. There are some gains as the level of wealth and consumption inequality within lower castes has increased.

References

- Aggarwal, Ashish (2012). "Implementation of Forest Rights Act, changing forest landscape, and politics of REDD+ in India", p. 18.
- Alvaredo, Facundo, Anthony B Atkinson, and Salvatore Morelli (Dec. 2016). "Top wealth shares in the UK over more than a century". *Working Paper*. WID.World 2017/2, p. 111.
- Anand, Ishan and Anjana Thampi (Dec. 10, 2016). "Recent Trends in Wealth Inequality in India". *Economic and Political Weekly* 51.50.
- Anderson, Siwan (2003). "Why Dowry Payments Declined with Modernization in Europe but Are Rising in India". *Journal of Political Economy* 111.2, pp. 269–310.
- Banerjee, A. (May 18, 2005). "Top Indian Incomes, 1922-2000". *The World Bank Economic Review* 19.1, pp. 1–20.
- Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune (May 2013). "Marry for What? Caste and Mate Selection in Modern India". *American Economic Journal: Microeconomics* 5.2, pp. 33–72.
- Basu, Alaka Malwade (1999). "Fertility Decline and Increasing Gender Imbalance in India, Including a Possible South Indian Turnaround". *Development and Change* 30.2, pp. 237–263.
- Blanchet, Thomas (2016). "Wealth inequality in Europe and in the United States: estimations from surveys, national accounts and wealth rankings", p. 96.
- Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (Sept. 11, 2017). *blanchet-t-fournier-j-piketty-t-generalized-pareto-curves-theory-applications-2017.pdf*. WID. URL: <http://wid.world/document/blanchet-t-fournier-j-piketty-t-generalized-pareto-curves-theory-applications-2017/> (visited on 05/30/2018).
- Borooah, Vani K. (Aug. 1, 2005). "Caste, Inequality, and Poverty in India". *Review of Development Economics* 9.3, pp. 399–414.
- Cancian, Maria and Deborah Reed (1999). "The Impact of Wives' Earnings on Income Inequality: Issues and Estimates". *Demography* 36.2, pp. 173–184.
- Chancel, Lucas and Thomas Piketty (July 1, 2017). "Indian income inequality, 1922-2015: From British Raj to Billionaire Raj?" *wid.world* 2017/11, p. 71.
- Chavan, Pallavi (2008). *RAS | Debt of Rural Households in India: A Note on the All-India Debt and Investment Survey*. URL: http://ras.org.in/debt_of_rural_households_in_india (visited on 04/09/2018).
- Chiplunkar, Gaurav and Jeff Weaver (2018). "Marriage Markets and the Rise of Dowry in India", p. 5.
- Das, Kumudin, K C Das, T K Roy, and P K Tripathy (2011). "Dynamics of inter-religious and inter-caste marriages in India". *Princeton Papers*, p. 14.
- Datt, Gaurav (Aug. 1, 1997). *Poverty in India 1951-1944 : trends and decompositions*. 90036. The World Bank, pp. 1–77.
- Deaton, A. (Aug. 11, 2005). "Data and Dogma: The Great Indian Poverty Debate". *The World Bank Research Observer* 20.2, pp. 177–199.
- Deaton, Angus. "Intertemporal Choice and Inequality". *JOURNAL OF POLITICAL ECONOMY* (), p. 31.
- Deaton, Angus, Paul Collier, Jean Drze, Valerie Kozel, Peter Lanjouw, and Karen Mason (2001). "Counting the worlds poor: problems and possible solutions". *World Bank Research Observer*, pp. 125–147.
- Desai, Sonalde, Vanneman Reeve, and National Council of Applied Economic Researc (Dec. 2011). *IHDS Datasets*.
- Deshpande, Ashwini (Feb. 1, 2001). "Caste at Birth? Redefining Disparity in India". *Review of Development Economics* 5.1, pp. 130–144.
- Deshpande, Ashwini and Rajesh Ramachandran (Nov. 2014). "How Backward are the Other Backward Classes? Changing Contours of Caste Disadvantage in India". *Centre for Development Economics* 233 (Working Paper), p. 50.
- Eika, Lasse, Magne Mogstad, and Basit Zafar (July 2014). *Educational Assortative Mating and Household Income Inequality*. Working Paper 20271. National Bureau of Economic Research.
- Frmeaux, Nicolas and Arnaud Lefranc (Oct. 2017). "Assortative Mating and Earnings Inequality in France". *IZA* (IZA DP No. 11084), p. 47.
- Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty (Jan. 17, 2017). "Accounting for Wealth Inequality Dynamics: Methods, Estimates and Simulations for France (1800-2014)", p. 68.
- Ghara, Tushar Kanti (2016). "Status of Indian Women in Higher Education". *Journal of Education and Practice*, p. 7.
- Gothoskar, S. P. (Dec. 1988). "On Some Estimates of Rural Indebtedness". *Reserve Bank of India Occasional Papers* 9.4, pp. 299–325.
- Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos (2014). "Marry Your Like: Assortative Mating and Income Inequality". *NBER Working Paper*, p. 26.
- GUPTA, CHARU (2009). "Hindu Women, Muslim Men: Love Jihad and Conversions". *Economic and Political Weekly* 44.51, pp. 13–15.

- Hershberger, Scott L. (2005). "Polychoric Correlation". *Encyclopedia of Statistics in Behavioral Science*. American Cancer Society.
- Jayadev, Arjun, Sripad Motiram, and Vamsi Vakulabharanam (2007). "Patterns of Wealth Disparities in India during the Liberalisation Era". *Economic and Political Weekly* 42.38, pp. 3853–3863.
- Jayadev, Arjun and Sanjay G. Reddy (2011). "Inequalities and Identities". *SSRN Electronic Journal*.
- Joshi, P D. "CONCEPTUALISATION, MEASUREMENT AND DIMENSIONAL ASPECTS OF POVERTY IN INDIA" (), p. 27.
- Kannabiran, Kalpana (2009). *sociology of caste and the crooked Mirror: recovering B r Ambedkars Legacy*.
- Kessler, Denis and Edward N. Wolff (Sept. 1991). "A COMPARATIVE ANALYSIS OF HOUSEHOLD WEALTH PATTERNS IN FRANCE AND THE UNITED STATES". *Review of Income and Wealth* 37.3, pp. 249–266.
- Kozel, Valerie and Barbara Parker (2003). "A Profile and Diagnostic of Poverty in Uttar Pradesh". *Economic and Political Weekly* 38.4, pp. 385–403.
- Kremer, Michael (May 1996). *How Much Does Sorting Increase Inequality?* Working Paper 5566. National Bureau of Economic Research.
- Kulkarni, P M (2012). "Demographic Transition in India", p. 52.
- Kumar, Rishabh (2016). "Capital and the Hindu Rate of Growth: Top Indian Wealth Holders 1961-1986". *SSRN Electronic Journal*.
- Libraries (CRL), Center for Research and Great Britain India Office. *Digital South Asia Library*. URL: <http://dsal.uchicago.edu/statistics/index.html> (visited on 04/09/2018).
- Luttmer, Erzo F. P. (Aug. 2004). *Neighbors as Negatives: Relative Earnings and Well-Being*. Working Paper 10667. National Bureau of Economic Research.
- Mistri, Avijit. "Estimation of Internal Migration in India, 2011 Census based on Life Table Survival Ratio (LTSR) Method" (), p. 17.
- Motiram, Sripad and Nayantara Sarma (July 3, 2014). "Polarization, Inequality, and Growth: The Indian Experience". *Oxford Development Studies* 42.3, pp. 297–318.
- Narayanan, M. P. (1988). "Debt Versus Equity under Asymmetric Information". *The Journal of Financial and Quantitative Analysis* 23.1, pp. 39–51.
- Olivo-Villabrille, Miguel (Oct. 2017). "Assortative Marriages and Household Income Inequality", p. 57.
- Piketty, Thomas (2014). *Capital in the twenty-first century*. Cambridge Massachusetts: The Belknap Press of Harvard University Press. 685 pp.
- Piketty, Thomas, Li Yang, and Gabriel Zucman (Apr. 2017). *Capital Accumulation, Private Property and Rising Inequality in China, 1978-2015*. w23368. Cambridge, MA: National Bureau of Economic Research.
- Polychoric Correlation - Hershberger - - Major Reference Works - Wiley Online Library*. URL: <https://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa486> (visited on 06/03/2018).
- Population Sciences (IIPS), International Institute for and ICF (2005). *NFHS- III Datasets*.
- Qian, Zhenchao (1997). "Breaking the Racial Barriers: Variations in Interracial Marriage Between 1980 and 1990". *Demography* 34.2, pp. 263–276.
- Rao, Vijender (2006). *The Economics of Dowries in India*. URL: <http://vijayendraraao.org/papers/dowryecon.pdf> (visited on 06/02/2018).
- Ray, Tridip, Arka Roy Chaudhuri, and Komal Sahai (Sept. 18, 2017). "WHOSE EDUCATION MATTERS? AN ANALYSIS OF INTER CASTE MARRIAGES IN INDIA". *Discussion Papers in Economics*, p. 49.
- Sardana, M M K (2012). "LAND ACQUISITION ISSUES", p. 8.
- Sarma, Mandira, Partha Saha, and Nandini Jayakumar. "Asset Inequality in India" (), p. 22.
- SCHWARTZ, CHRISTINE R. (Aug. 2010). "Pathways to Educational Homogamy in Marital and Cohabiting Unions". *Demography* 47.3, pp. 735–753.
- Schwartz, Christine R. (2013). "Trends and Variation in Assortative Mating: Causes and Consequences". *Annual Review of Sociology* 39.1, pp. 451–470.
- Senik, Claudia (Oct. 1, 2009). "Direct evidence on income comparisons and their welfare effects". *Journal of Economic Behavior & Organization* 72.1, pp. 408–424.
- Sharma, H R (1994). "Distribution of Landholdings in Rural India, 1953-54 to 1981-82", p. 14.
- Srinivas, M N (Oct. 15, 1955). "Castes: Can They Exist in India of Tomorrow?" *The Economic Weekly*, p. 3.
- Subramanian, S. and D. Jayaraj (Oct. 16, 2008). "The Distribution of Household Wealth in India". *Personal Wealth from a Global Perspective*. Oxford University Press, pp. 112–133.

- Sundaram, K. and Suresh D. Tendulkar (2001). "NAS-NSS Estimates of Private Consumption for Poverty Estimation: A Disaggregated Comparison for 1993-94". *Economic and Political Weekly* 36.2, pp. 119–129.
- Tarozzi, Alessandro (May 2002). *Estimating Comparable Poverty Counts from Incomparable Surveys: Measuring Poverty in India*. 186. Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies.
- Thorat, Sukhadeo (2002). "Oppression and Denial: Dalit Discrimination in the 1990s". *Economic and Political Weekly* 37.6, pp. 572–578.
- Vermeulen, Philip (May 2016). "Estimating the Top Tail of the Wealth Distribution". *American Economic Review* 106.5, pp. 646–650.
- Waheed, Abdul (Feb. 2009). "Dowry among Indian Muslims: Ideals and Practices". *Indian Journal of Gender Studies* 16.1, pp. 47–75.
- Wazal, Dr Makarand S and Sudesh Kumar Sharma. "Participation of Retail Investors in Indian Equity Market" (), p. 3.
- Zacharias, Ajit and Vamsi Vakulabharanam (Oct. 1, 2011). "Caste Stratification and Wealth Inequality in India". *World Development* 39.10, pp. 1820–1833.
- Zhu, Junyi and Etienne Pasteau (2017). *Love and money with inheritance: marital sorting between labor income and inherited wealth in the modern partnership*. Annual Conference 2017 (Vienna): Alternative Structures for Money and Banking. Verein für Socialpolitik / German Economic Association.

9 Appendix

9.1 A.1 Figures

Figure 26: Population Share from NSS

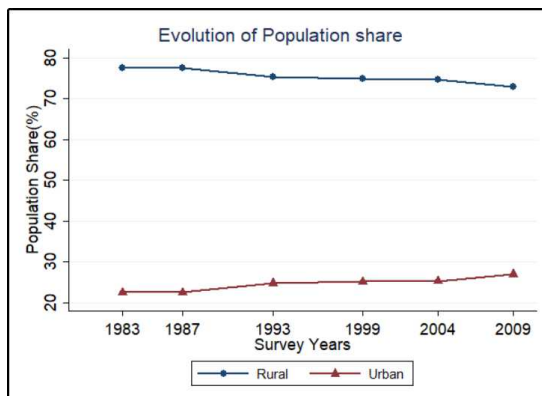
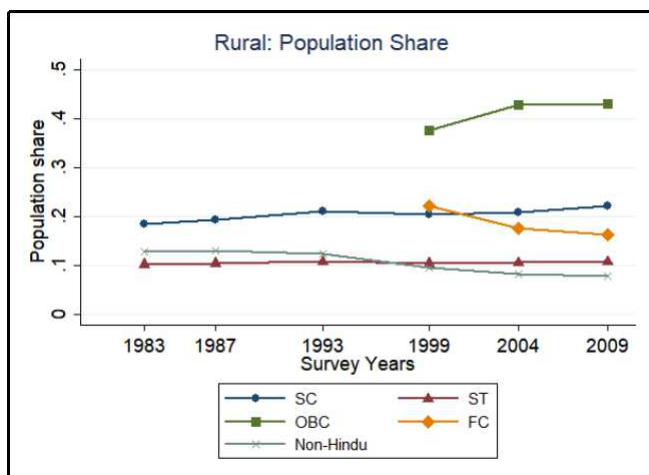


Figure 27: Population Share from NSS-Rural Urban

(a) Proportion of SC/ST/OBC/FC



(b) Proportion of SC/ST/OBC/FC

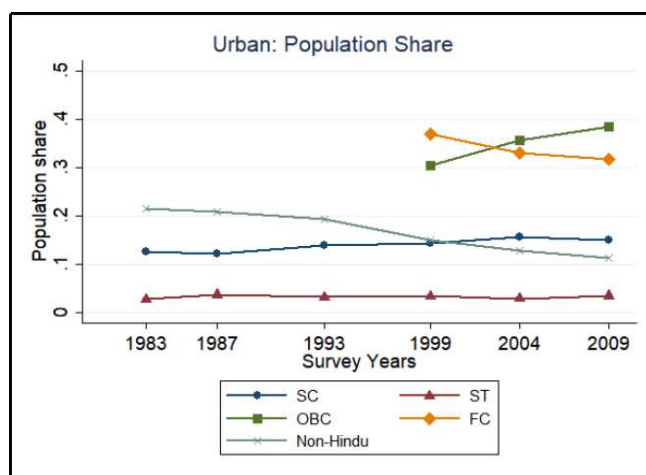


Figure 28: Caste Share in Different Religion from NSS-1883 1887 1993

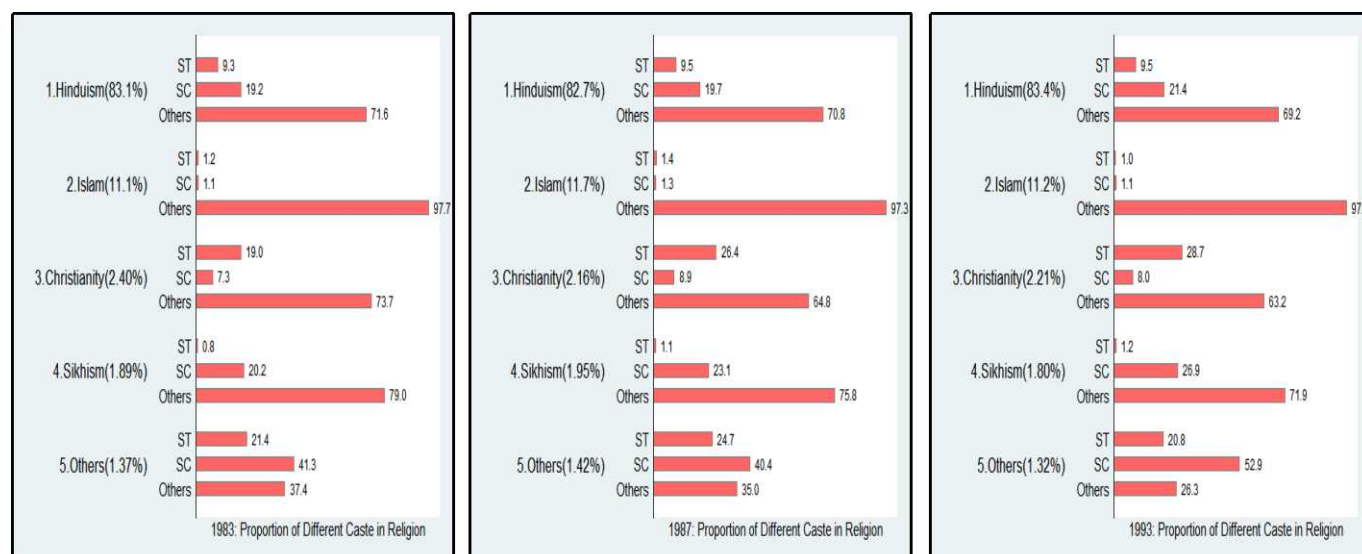


Figure 29: Caste Share in Different Religion from NSS-1999, 2004, 2009

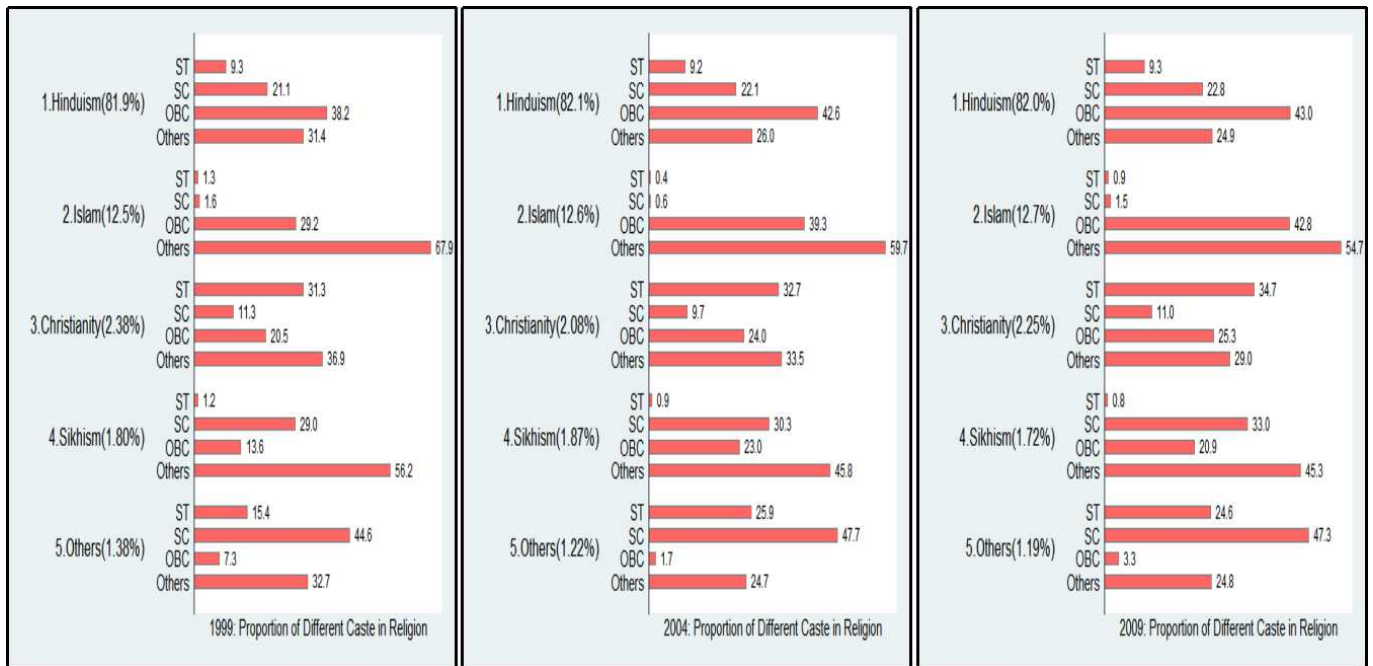


Figure 30: Wealth Share: Top .1% and Top .01%

(a) Top 0.1%

(b) Top 0.01%

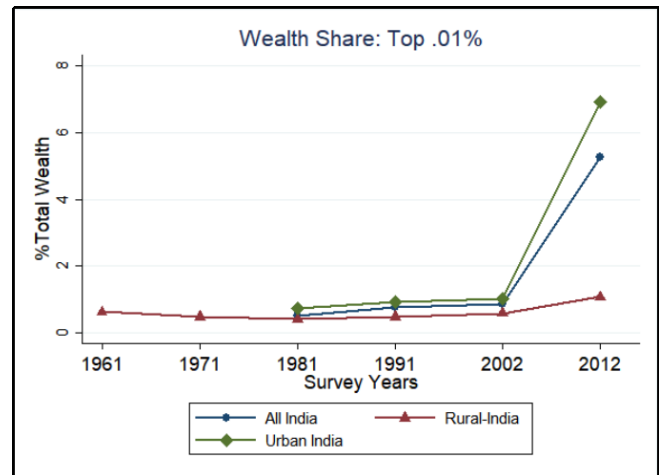
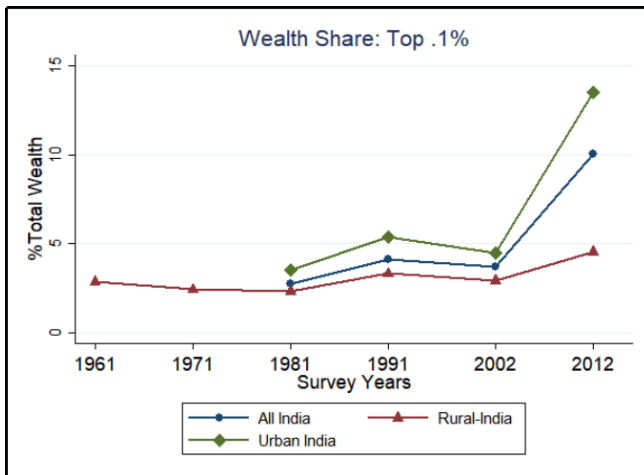
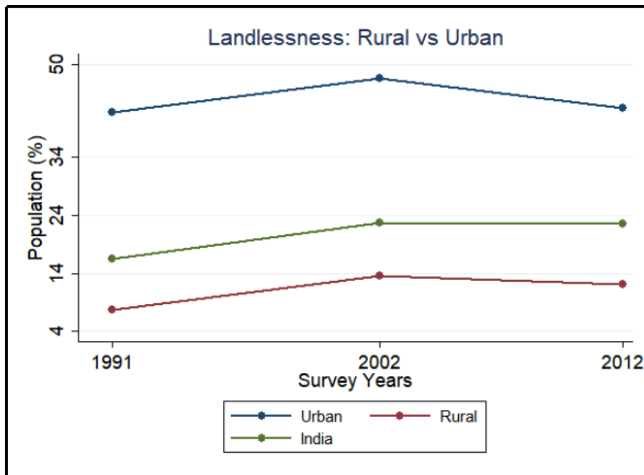
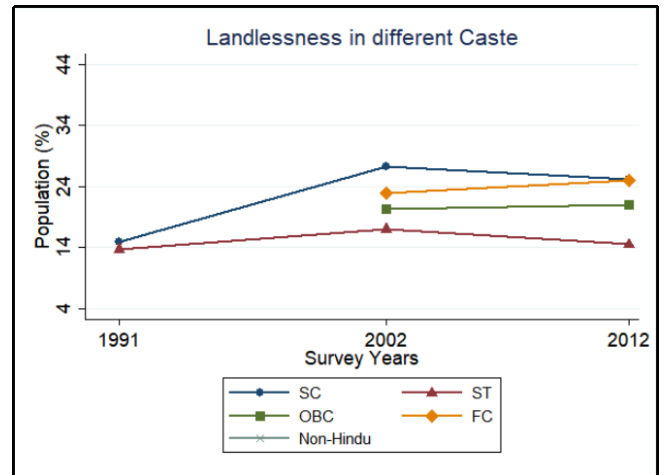


Figure 31: Landlessness by Land Ownership

(a) Rural Urban Comparison



(b) Landlessness in different Caste:All India Level



9.2 A.2 Tables

Table 19: Population Share (Brahmins) - IHDS

	Percentage(%) of Population across Caste group							
	SC	ST	OBC	FC(Brahmin)	FC(Non-Brahmin)	Muslim	Others	Total
Total	21.80	8.00	42.78	4.86	14.90	6.19	1.47	100
0-14 Boys	23.12	8.21	43.95	4.27	12.41	6.96	1.09	100
0-14 Girls	23.44	8.59	44.13	4.08	11.56	7.28	0.92	100
15-20 Boys	22.60	7.75	43.10	4.72	13.93	6.55	1.36	100
15-20 Girls	23.42	7.62	43.18	4.59	13.00	7.09	1.09	100
>21 yrs Men	20.91	8.01	41.79	5.18	16.71	5.69	1.71	100
>21 yrs Women	20.82	7.74	42.43	5.25	16.44	5.55	1.77	100
>60yrs Men	18.93	6.89	43.17	6.06	17.99	4.80	2.16	100
>60yrs Women	19.71	5.92	43.76	6.22	17.87	4.27	2.23	100

Source: Author's calculation using IHDS 2011 datasets.

Table 20: Population Share FC split - NFHS

Population %													
Caste Group	Overall	Young (0-14yrs) %			Adolescent(15-20yrs) %			Adult(>21 yrs) %			Old (>60yrs) %		
		M	F	T	M	F	T	M	F	T	M	F	T
ST	8.397	9.07	9.44	9.25	8.38	8.81	8.6	7.86	7.73	7.79	6.46	6.74	6.6
SC	19.08	19.87	20.39	20.12	19.96	19.79	19.87	18.36	18.1	18.23	16.83	16.3	16.57
OBC	40.13	41.55	41.09	41.33	39.31	40.21	39.78	39.05	39.8	39.43	40.06	39.79	39.93
FC(Brahman)	4.648	3.87	3.67	3.77	4.68	4.07	4.36	5.37	5.2	5.28	6.93	6.59	6.76
FC(Rajput)	4.902	4.37	4.07	4.23	5.03	4.68	4.85	5.39	5.32	5.36	5.83	5.48	5.66
FC(Bania)	2.008	1.77	1.75	1.76	1.96	1.8	1.88	2.25	2.16	2.2	2.35	2.41	2.38
FC(Kayasth)	0.6344	0.41	0.41	0.41	0.46	0.5	0.48	0.81	0.82	0.82	0.97	1.18	1.07
FC(Other)	9.315	7.4	7.59	7.49	8.86	8.1	8.46	10.82	10.59	10.7	11.13	11.7	11.41
Muslim	8.691	9.98	10.07	10.03	9.25	10.31	9.81	7.49	7.63	7.57	6.18	6.28	6.23
Other	2.187	1.71	1.5	1.61	2.12	1.75	1.92	2.62	2.64	2.63	3.26	3.53	3.39

Source: Author's calculation using NFHS 2005 datasets.

Table 21: Wealth, Income and Consumption Standard Deviation-IHDS 2011

	SC	ST	OBC	FC(Brahmin)	FC(Non-Brahmin)	M
Annual Income of HH (in Rs)	190,165	106,633	156,237	211,666	277,773	1
Per Capita Annual Income (in Rs)	40,503	23,252	32,338	44,742	60,841	3
Annual Consumption of HH (in Rs)	77,577	73,477	108,385	139,527	136,120	9
Per Capita Annual Consumption (in Rs)	16,523	16,022	22,433	29,493	29,815	1
ASSETS	6.1	5.7	6.4	6.4	6.2	
Highest Adult Education	5.7	5.3	6.0	5.9	5.7	
Highest Male Education	5.1	5.0	5.1	4.1	4.7	
Highest Female Education	5.0	4.9	5.0	4.0	4.6	

Source: Author's calculation using NFHS IHDS 2011 datasets. Design weights are used to estimate these values.

Table 22: Wealth Index-NFHS 2005

	Wealth Index				
	Poorest	Poorer	Middle	Richer	Richest
Overall	20.63	19.82	19.86	19.6	20.09
ST	51	23.09	12.87	7.78	5.26
SC	28.47	24.8	21.19	16.08	9.46
OBC	18.87	21.66	22.94	20.7	15.83
FC(Brahman)	4.62	9.7	13.86	21.9	49.91
FC(Rajput)	7.27	13.78	21.9	25.89	31.15
FC(Bania)	5.8	11.86	16.52	22.17	43.66
FC(Kayasth)	2.17	5.25	10.89	24.67	57.02
FC(Other)	9.75	13.42	17.13	24.45	35.26
Muslim	20.91	21.19	19.11	21.8	16.99
Other	2.45	4.08	9.45	22.2	61.81

Table 23: Highest Adult Education Level-NFHS 2005

	Highest Adult Education in HH		
Caste Group	Overall	Male	Female
ST	4.34	3.92	1.8
SC	5.69	5.11	2.61
OBC	6.81	6.1	3.62
FC(Brahman)	11.88	10.87	8.28
FC(Rajput)	9.05	8.23	5.71
FC(Bania)	10.33	9.57	6.81
FC(Kayasth)	12.33	11.04	9.86
FC(Other)	9.16	8.15	6.35
Muslim	5.84	4.97	3.35
Other	10.83	9.22	8.51

Table 24: Composition of Population in NSS AIDIS survey

Population Share of different Social groups																		
	Total Population Share									Adult Population Share (>20yrs)								
	India			Rural			Urban			India			Rural			Urban		
caste_code	1991	2002	2012	1991	2002	2012	1991	2002	2012	1991	2002	2012	1991	2002	2012	1991	2002	2012
ST	8.74	8.01	9.29	10.61	9.78	12.01	2.97	2.81	3.5	8.42	7.6	8.78	10.35	9.46	11.63	2.79	2.58	3.33
SC	18.4	19.71	18.78	20.25	21.27	20.61	12.69	15.13	14.88	18.01	18.81	17.89	20.17	20.6	19.76	11.7	14	14.31
OBC		40.28	43.57		41.97	44.54		35.35	41.5		39.72	42.85		41.76	44.2		34.26	40.27
FC		22.56	20.48		18.91	15.81		33.25	30.43		24.87	22.8		20.55	17.56		36.5	32.8
Muslim		7.56	6.26		6.62	5.7		10.32	7.45		6.75	5.77		5.87	5.24		9.11	6.77
Non-Hindu-Muslim		1.89	1.62		1.46	1.33		3.15	2.24		2.24	1.91		1.76	1.6		3.54	2.51
Others	72.86			69.14			84.34			73.57			69.48			85.52		
Total Pop (mill)	810	986	1,058	612	736	721	198	251	338	427	546	638	318	398	419	109	148	220

Note: This table presents the total population and total adult population share in NSS AIDIS datasets, by caste groups in Rural, Urban and all-India level separately.

Table 25: Ratio of Threshold to Population average

	Year	Average Wealth	Percentiles				
			.50	.10	.01	0.001	0.0001
India	1981	8,805	0.53	2.15	7.55	19.13	38.85
	1991	40,796	0.47	2.24	8.19	26.28	60.82
	2002	49,257	0.47	2.27	8.34	22.87	56.22
	2012	132,826	0.37	2.06	9.4	31.69	138.53
Rural	1961	1,373	0.55	2.07	7.2	18.32	42.3
	1971	2,747	0.57	2.13	6.99	16.69	34.68
	1981	10,898	0.57	2.11	7.13	17.12	32.45
	1991	38,504	0.51	2.24	7.58	24.25	44.6
	2002	44,523	0.51	2.25	7.85	19.84	41.97
	2012	96,059	0.45	2.1	9.01	27.51	69.42
Urban	1981	15,282	0.43	2.25	8.64	23.6	51.83
	1991	47,885	0.38	2.31	9.47	35.66	77.2
	2002	63,140	0.4	2.42	8.31	27.22	68.02
	2012	211,261	0.33	2.07	8.89	37.93	204.96

Table 26: Representational Wealth Inequality

Difference between wealth share and population share									
	India			Rural			Urban		
caste_code	1991	2002	2012	1991	2002	2012	1991	2002	2012
ST	-2.83	-0.18	-1.95	-3.47	0.37	-1.91	1.12	2.24	1.72
SC	-11.23	-10.3	-10.84	-11.76	-10.14	-10.48	-7.08	-7.59	-8.63
OBC		-7.88	-10.21		-5.37	-4.39		-7.33	-12.02
FC		14.02	18.17		9.42	10.15		12.01	16.29
Muslim		-1.84	-1.75		-0.93	-1.78		-4.24	-2.41
Non-Hindu-Muslim		6.2	6.59		6.65	8.42		4.92	5.05

Table 27: Representational Land Inequality

Difference between wealth share and population share																		
caste_code	Land Area									Land Value								
	India			Rural			Urban			India			Rural			Urban		
	1991	2002	2012	1991	2002	2012	1991	2002	2012	1991	2002	2012	1991	2002	2012	1991	2002	2012
ST	3.6	7.04	10.81	2.33	6.23	9.6	3.2	5.22	10.36	-2.81	0.11	-2	-4.16	0.04	-3.32	1.51	2.36	2.11
SC	-9.83	-9.94	-9.4	-11.23	-10.68	-5.61	-8.35	-7.94	-11.07	-10.84	-10.96	-12.31	-11.36	-11.58	-6.83	-8	-8.47	
OBC		-1.5	-4.21		-2.85	-4.48		-0.5	-5.49		-6.2	-7.93		-5.66	-5.21		-4.74	-8.89
FC		5.3	3.91		8.04	6.24		3.86	4.39		11.4	14.74		9.69	11.91		9.13	11.75
Muslim		-2.25	-2.74		-1.68	-2.47		-2.56	-2.8		-2.04	-1.82		-0.94	-1.97		-4.74	-2.24
Non-Hindu-Muslim		1.37	1.63		1.5	1.8		2.34	1.49		7.57	7.98		8.23	10.17		6	5.74

Table 28: Threshold Wealth Value within caste decile

Threshold Wealth Value within caste deciles											
Caste/Deciles		1	2	3	4	5	6	7	8	9	10
1991	ST	0	1,862	4,197	6,575	9,405	12,345	16,115	21,475	27,070	42,050
	SC	0	1,529	3,207	5,000	7,078	9,940	13,249	18,638	26,333	41,758
	Others-Mix	0	2,013	6,065	10,700	16,138	22,617	31,506	43,425	62,556	105,647
2002	ST	0	5,625	11,323	17,362	23,978	31,697	40,663	52,273	72,572	118,060
	SC	0	5,140	10,525	16,025	22,197	29,226	38,483	51,888	72,680	113,953
	OBC	0	6,925	17,110	27,190	38,303	51,903	69,503	94,205	132,225	221,000
	FC	0	8,260	27,027	44,580	64,167	88,513	117,767	160,781	232,750	375,700
	Muslim	0	3,075	10,100	16,600	25,075	36,150	50,825	71,667	105,125	175,250
2012	ST	0	22,865	40,817	58,335	78,500	105,800	141,900	195,238	286,860	502,165
	SC	0	21,575	43,185	65,750	91,700	123,525	162,800	222,013	312,925	509,000
	OBC	0	34,164	70,595	108,314	152,350	211,163	286,173	400,650	595,578	1,019,500
	FC	0	36,025	92,875	161,817	237,550	348,315	488,742	695,458	1,100,000	2,006,000
	Muslim	0	20,400	43,790	74,140	108,825	156,077	221,167	309,150	455,386	898,667

Table 29: Cross-Ownership

Cross Ownership(Rural-Urban)				
2012	HH's Residence	Location of the Land	%age of HH's Owning Land	%age of Cross-Ownership
Land Value	Rural	Land_Value_Urban	0.81	1.18
	Urban	Land_Value_Urban	56.03	98.82
	Rural	Land_Value_Rural	93.96	85.25
	Urban	Land_Value_Rural	14.84	14.75
Land Area	Rural	Land_Area_Urban	0.81	4.29
	Urban	Land_Area_Urban	57.36	95.71
	Rural	Land_Area_Rural	94.02	87.67
	Urban	Land_Area_Rural	14.85	12.33

Table 30: Employment Type in different Caste(2012)

Employment Type (2012)												
caste_code	Rural							Urban				
	Self-Employed in Agri	Self-Employed in Non-Agri	Agri Labour	Regular Wage/Salary	Casual Labour	Other	Caste Total	Self-Employed	Regular Wage/Salary	Casual Labour	Other	Caste Total
ST	51.11	4.59	17.46	8.91	15.28	2.65	100.00	20.03	45.45	23.47	11.04	100.00
SC	28.90	10.41	24.76	10.29	20.21	5.43	100.00	23.44	43.80	25.51	7.26	100.00
OBC	44.76	12.56	15.05	9.43	12.43	5.76	100.00	36.17	35.86	17.65	10.32	100.00
FC	51.82	11.25	9.21	13.17	6.25	8.31	100.00	30.92	49.06	6.89	13.13	100.00
Muslim	31.84	21.04	13.89	8.85	16.44	7.94	100.00	42.46	32.02	19.24	6.28	100.00
Non-Hindu-Muslim	49.29	13.24	7.49	15.78	6.10	8.08	100.00	33.55	44.15	8.11	14.19	100.00
Employment Total	42.76	11.45	16.19	10.23	13.47	5.91	100.00	32.44	41.59	15.28	10.68	100.00

Table 31: Percentage of Employment Type in Castes(2002)

Employment Type (2002)											
caste_code	Rural					Caste Total	Urban				
	Self-Employed in Agri	Self-Employed in Non-Agri	Agri Labour	Casual Labour	Other		Self-Employed	Regular Wage/Salary	Casual Labour	Other	Caste Total
ST	42.75	6.16	33.95	11.56	5.58	100.00	21.86	45.62	21.44	11.08	100.00
SC	22.26	12.53	42.40	13.96	8.85	100.00	29.50	42.25	23.88	4.37	100.00
OBC	39.40	16.55	23.03	10.80	10.23	100.00	40.43	37.01	14.68	7.88	100.00
FC	50.23	13.44	13.11	6.65	16.57	100.00	34.34	49.62	4.98	11.05	100.00
Muslim	30.48	24.01	23.67	12.12	9.72	100.00	47.86	34.07	12.57	5.50	100.00
Non-Hindu-Muslim	49.81	13.59	9.17	10.69	16.74	100.00	43.26	38.76	4.58	13.40	100.00
Employment Total	37.63	14.46	26.31	10.86	10.75	100.00	36.87	42.34	12.21	8.58	100.00

Table 32: Proportion of caste in Employment types(2012)

Employment Type (2012)												
caste_code	Rural							Urban				
	Self-Employed in Agri	Self-Employed in Non-Agri	Agri Labour	Regular Wage/Salary	Casual Labour	Other	Caste Total	Self-Employed	Regular Wage/Salary	Casual Labour	Other	Caste Total
ST	13.90	4.66	12.55	10.13	13.20	5.22	11.63	2.11	3.73	5.24	3.53	3.41
SC	13.95	18.78	31.58	20.77	30.98	18.96	20.64	10.12	14.75	23.37	9.52	14.00
OBC	46.15	48.38	40.99	40.63	40.70	42.99	44.08	45.77	35.40	47.40	39.69	41.06
FC	20.27	16.44	9.51	21.53	7.76	23.51	16.73	31.28	38.71	14.81	40.34	32.82
Muslim	4.10	10.11	4.72	4.76	6.72	7.39	5.50	8.20	4.82	7.89	3.69	6.26
Non-Hindu-Muslim	1.63	1.64	0.65	2.18	0.64	1.93	1.41	2.53	2.59	1.30	3.25	2.44
Employment Total	100	100	100	100	100	100	100	100	100	100	100	100

Table 33: Proportion of caste in Employment types(2002)

Proportion of caste in Employment types (2002)											
caste_code	Rural						Urban				
	Self-Employed in Agri	Self-Employed in Non-Agri	Agri Labour	Casual Labour	Other	Caste Total	Self-Employed	Regular Wage/Salary	Casual Labour	Other	Caste Total
ST	13.90	4.66	12.55	10.13	13.20	5.22	11.63	2.11	3.73	5.24	3.53
SC	13.95	18.78	31.58	20.77	30.98	18.96	20.64	10.12	14.75	23.37	9.52
OBC	46.15	48.38	40.99	40.63	40.70	42.99	44.08	45.77	35.40	47.40	39.69
FC	20.27	16.44	9.51	21.53	7.76	23.51	16.73	31.28	38.71	14.81	40.34
Muslim	4.10	10.11	4.72	4.76	6.72	7.39	5.50	8.20	4.82	7.89	3.69
Non-Hindu-Muslim	1.63	1.64	0.65	2.18	0.64	1.93	1.41	2.53	2.59	1.30	3.25
Employment Total	100	100	100	100	100	100	100	100	100	100	100

Table 34: Rural Top Decile: Employment type and Caste

2002												
caste_code	Rural							2012				
	Self-Employed in Agri	Self-Employed in Non-Agri	Agri Labour	Casual Labour	Other	Caste Share (%)	Self-Employed in Agri	Self-Employed in Non-Agri	Regular Wage/Salary	Agri Labour	Casual Labour	Other
ST	3.03	2.17	4.59	2.46	3.50	3.01	3.75	1.63	7.09	5.03	6.33	0.85
SC	4.47	5.45	16.64	7.56	6.26	4.97	5.29	4.88	10.50	6.89	9.45	4.89
OBC	41.60	46.72	46.03	38.59	41.80	41.80	46.32	59.56	42.53	49.77	61.62	54.64
FC	39.43	30.51	22.43	29.08	39.34	38.18	34.20	29.48	26.56	15.70	11.23	29.43
Muslim	3.96	6.83	7.03	7.11	3.27	4.25	2.88	6.53	5.09	12.07	5.62	0.49
Non-Hindu-Muslim	7.51	8.32	8.73	7.76	9.05	7.79	7.56	6.92	8.23	10.54	5.74	9.70
Column Sum	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Employment % share	7.55	0.91	0.11	0.23	1.20	9.99	7.07	1.07	1.08	0.13	0.27	0.39

Table 35: Urban Top Decile: Employment type and Caste

caste_code	2002					2012				
	Self-Employed	Regular Wage/Salary	Casual Labour	Other	Caste Share (%)	Self-Employed	Regular Wage/Salary	Casual Labour	Other	Caste Share (%)
ST	1.10	0.79	0.19	2.06	1.10	0.87	2.33	6.51	2.83	1.86
SC	1.84	5.48	7.28	1.37	3.16	1.84	4.80	7.27	2.93	3.34
OBC	26.93	17.55	44.30	19.99	22.76	34.07	23.94	47.93	31.67	29.77
FC	53.01	65.11	12.12	69.48	59.19	50.45	58.97	30.50	46.17	53.07
Muslim	5.07	2.35	16.92	2.14	3.80	5.30	2.21	5.36	6.44	4.14
Non-Hindu-Muslim	12.05	8.72	19.19	4.95	10.01	7.47	7.75	2.43	9.96	7.81
Column Sum	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Employment % share	5.04	3.67	0.07	1.22	10.00	4.27	4.22	0.21	1.30	10.00

Table 36: Age Profile in Wealth Deciles: Rural

Age Profile in Wealth Deciles (Rural)												
	Deciles/Year	Life Expectancy	1	2	3	4	5	6	7	8	9	10
Mean Age (yrs)	1991	58.9	24.22	23.78	23.97	24.4	24.49	24.47	24.52	24.27	24.84	25.03
	2002	62.3	24.89	24.84	25.21	25.02	25.43	25.3	25.73	25.75	26.25	26.97
	2012	64.9	25.94	26.69	27.4	27.13	27.09	27.85	27.88	28.28	28.51	30.2
Median Age(yrs)	1991	58.9	22	21	21	20	21	20	20	20	20	19
	2002	62.3	22	22	22	22	22	21	22	21	21	21
	2012	64.9	23	24	24	24	24	24	24	24	25	27

Table 37: Age Profile in Wealth Deciles: Urban

Age Profile in Wealth Deciles (Urban)												
	Deciles/Year	Life Expectancy	1	2	3	4	5	6	7	8	9	10
Mean Age (yrs)	1991	63	23.6	23.83	24.32	24.38	24.92	25.34	25.29	26.26	26.36	28.05
	2002	66.1	24.08	25.18	25.55	25.28	25.93	26.85	27.38	28.64	29.73	31.52
	2012	68.1	25.22	27.25	27.7	28.45	28.53	30.03	30.9	31.45	32.67	34.8
Median Age(yrs)	1991	63	22	22	22	22	22	22	21	22	22	24
	2002	66.1	22	23	23	22	23	24	24	25	27	30
	2012	68.1	22	25	25	25	26	27	29	30	32	35

Table 38: Education Profile in Wealth Deciles (Rural)

Educational Profile- Decile-wise Percentage (Rural)											
year	Deciles	1	2	3	4	5	6	7	8	9	10
1991	Illiterate	69.18	72.23	67.58	64.28	60.68	59.02	57.24	53.44	48.7	43
	Below Primary	14.38	14.34	15.56	16.28	17.27	17.65	17.32	18.21	19.16	19.51
	Below Secondary	13.07	10.55	13.7	15.04	17.51	18.17	18.64	20.81	22.5	24.47
	Secondary/Higher Secondary	2.94	2.39	2.6	3.51	3.71	4.28	5.24	5.75	7.4	9.75
	Above Higher Secondary	0.4	0.49	0.51	0.88	0.83	0.88	1.56	1.76	2.23	3.26
2002	Illiterate	57.55	59.48	55.78	52.96	50.76	47.76	46.15	42.76	37.52	31.91
	Below Primary	15.37	16.59	16.37	16.9	16.92	17.04	17.11	16.88	17.24	15.57
	Below Secondary	20.86	19.58	22.42	23.41	25.18	26.2	26.97	28.4	29.61	31.91
	Secondary/Higher Secondary	5.42	3.85	4.71	5.9	6.32	7.75	8.31	9.9	12.95	16.6
	Above Higher Secondary	0.81	0.49	0.72	0.82	0.82	1.25	1.45	2.05	2.66	4.02
2012	Illiterate	48.8	49.16	43.34	43.5	41.33	39.23	35.72	33.47	30.65	26.42
	Below Primary	19.82	19.96	18.79	19.52	18.61	17.8	17.83	19.36	18.3	15.62
	Below Secondary	22.44	22.64	26.57	24.35	25.5	26.52	27.63	26.95	27.4	28.61
	Secondary/Higher Secondary	7.92	7.17	9.52	11.14	12.54	14.02	15.6	16.39	19.57	23.49
	Above Higher Secondary	1.01	1.08	1.77	1.5	2.02	2.45	3.22	3.83	4.09	5.86

Table 39: Education Profile in Wealth Deciles (Urban)

Educational Profile- Decile-wise Percentage (Urban)											
year	Deciles	1	2	3	4	5	6	7	8	9	10
1991	Illeterate	41.27	35.71	42.87	41.75	41.25	34.01	32.45	25.46	22.6	16.13
	Below Primary	18.61	19.72	19.51	18.32	18.73	17.58	18.09	18.09	17.77	16.26
	Below Secondary	27.95	27.73	22.7	22.65	24.99	29.33	29.8	29.59	28.35	24.7
	Secondary/Higher Secondary	9.49	12.19	9.18	11.19	10.45	13.78	14.19	17.05	19.38	23.11
	Above Higher Secondary	2.68	4.65	5.67	5.93	4.58	5.27	5.43	9.8	11.76	19.78
2002	Illeterate	35.63	29.89	37.21	36.02	32	25.69	23.65	19.27	13.98	10.07
	Below Primary	16.92	15.17	16.39	17.7	16.13	15.8	14.42	14.19	12.38	10.59
	Below Secondary	31.94	31.34	28.68	29.3	32.21	32.95	32.48	29.62	28.14	23.11
	Secondary/Higher Secondary	12.62	18.21	12.56	12.94	14.81	18.92	21.06	24.62	27.83	29.11
	Above Higher Secondary	2.89	5.39	5.17	4.03	4.86	6.64	8.4	12.3	17.66	27.13
2012	Illeterate	30.51	27.29	29.35	28.55	25.08	21.67	18.08	15.09	11.67	9.85
	Below Primary	15.06	15.58	16.05	15.93	16.08	13.96	13.09	12.7	11.62	9.79
	Below Secondary	29.29	26.96	27.48	27.41	26.99	26.1	25.61	22.99	21	16.97
	Secondary/Higher Secondary	20.74	22.92	19.38	20.35	22.39	26.6	29.19	30.48	31.61	27.91
	Above Higher Secondary	4.4	7.24	7.73	7.76	9.45	11.68	14.03	18.75	24.1	35.49

Table 40: Relative position of different Caste groups

	year/caste_group	MPCE							Percentage difference with Overall Consumption					
		ST	SC	OBC	FC	Hindu	Muslim	Overall	ST	SC	OBC	FC	Hindu	Muslim
RURAL	1983	489	526			671.0092	597	628	-22.2	-16.22			6.85	-4.91
	1987	580	621			786.0208	710	736	-21.22	-15.58			6.8	-3.59
	1993	617	629			795.9214	703	741	-16.69	-15.08			7.41	-5.1
	1999	650	701	794	1001		754	814	-20.2	-13.83	-2.48	22.96		-7.43
	2004	643	716	840	1068		802	843	-23.76	-15.08	-0.41	26.71		-4.86
	2009	765	790	924	1174		841	928	-17.58	-14.88	-0.42	26.47		-9.39
URBAN	1983	920	872			1194.496	857	1108	-16.96	-21.27			7.81	-22.67
	1987	1022	937			1359.602	966	1245	-17.91	-24.73			9.2	-22.42
	1993	1113	1001			1457.907	1023	1339	-16.87	-25.25			8.88	-23.57
	1999	1217	1073	1295	1898		1184	1507	-19.27	-28.78	-14.04	25.94		-21.41
	2004	1244	1100	1264	2038		1209	1527	-18.52	-27.94	-17.24	33.44		-20.82
	2009	1610	1275	1488	2388		1424	1786	-9.86	-28.58	-16.71	33.68		-20.28

Table 41: Evolution of Population share : NSS Consumption Dataset

NSS Consumption Population Share								
	Year/Caste	ST	SC	OBC	FC	Hindu	Muslim	Non-Hindu
	1983	8.59	17.2			59.5	10.93	3.79
India	1987	8.97	17.66			58.57	11.43	3.37
	1993	8.94	19.3			57.72	11	3.05
	1999	8.71	18.9	35.73	25.78		8.49	2.4
	2004	8.63	19.59	40.94	21.42		7.55	1.87
	2009	8.8	20.29	41.75	20.43		7	1.73
Rural	1983	10.28	18.5			58.44	9.47	3.3
	1987	10.48	19.24			57.21	10.13	2.94
	1993	10.82	21.1			55.78	9.67	2.63
	1999	10.49	20.42	37.52	22.04		7.51	2.02
	2004	10.57	20.92	42.75	17.51		6.74	1.5
	2009	10.79	22.23	42.97	16.23		6.44	1.35
Urban	1983	2.76	12.7			63.13	15.95	5.47
	1987	3.77	12.2			63.25	15.91	4.87
	1993	3.21	13.85			63.6	15.04	4.3
	1999	3.39	14.35	30.38	36.93		11.41	3.53
	2004	2.92	15.64	35.6	32.94		9.95	2.94
	2009	3.45	15.06	38.47	31.77		8.5	2.75

Table 42: Representational Inequality to compare Rural and Urban

Year (y)	Caste (c)	RI ^{ycd}		
		Bottom 50%	Middle 40%	Top 10%
1983	ST	0.12	-0.1	-0.19
	SC	-0.02	0	0.11
	Hindu	0.01	0	-0.03
	Muslim	-0.3	0.28	0.4
	Non-Hindu	-0.18	-0.03	1.03
1987	ST	0.11	-0.12	-0.05
	SC	-0.11	0.09	0.15
	Hindu	0.02	-0.02	-0.02
	Muslim	-0.31	0.31	0.34
	Non-Hindu	-0.2	-0.04	1.15
1993	ST	0.07	-0.05	-0.15
	SC	-0.14	0.14	0.15
	Hindu	0.02	-0.01	0
	Muslim	-0.28	0.26	0.36
	Non-Hindu	-0.1	-0.15	1.1
1999	ST	0.08	-0.07	-0.15
	SC	-0.14	0.12	0.18
	OBC	-0.19	0.15	0.34
	FC	-0.02	-0.01	0.13
	Muslim	-0.24	0.26	0.17
2004	Non-Hindu	-0.01	-0.35	1.52
	ST	0.22	-0.19	-0.32
	SC	-0.12	0.1	0.18
	OBC	-0.23	0.18	0.42
	FC	0.04	-0.05	0.03
2009	Muslim	-0.27	0.28	0.24
	Non-Hindu	-0.19	-0.28	2.05
	ST	0.18	-0.16	-0.28
	SC	-0.17	0.17	0.17
	OBC	-0.19	0.16	0.33
	FC	0.04	-0.07	0.12
	Muslim	-0.19	0.18	0.19
	Non-Hindu	-0.15	-0.47	2.63

Table 43: Parallel between Occupational categories

Ordered Classification - FRANCE	INSEE (2003) Socio-Professional Category	Ordered Classification - INDIA	NCO 1968 Categorization
Higher grade professionals	Senior Managers and Professional Occupation	Professional	Professional, Technical and Related
Lower-grade Professionals	Intermediate Professions	Admin,Exec,Managers	Administrative, Executive and Managerial
Artisans	Craftsmen,Traders and Entrepreneurs	Clerical	Clerical
Non- Manual employees	Employees	Sales	Sales Workers
Farmers + Workers	Farmers	Service Providers	Service Workers
	Workers	Farmers,Cultivators	Farmers,Fishermen and related workers
		Labourers (Non-Agri)	Production and related workers, Labourers(Non-Agriculture)

Table 44: Parallel between Educational categories

Ordered Ranking	France	India
1	No degree	No Education
2	General Lower secondary degree	Less than Primary
3	Vocational Lower degree	5th Pass
4	Vocational Upper Secondary Degree	8th Pass
5	General Upper Secondary degree	Secondary
6	College (Bachelor/Technical degree)	H.Secondary and Diploma (<3yrs)
7	Master's degree	Bachelors(BA,Bsc,Diploma 3+)
8	PhD/Elite schools	BTech,MBBS,MD,CA,PhD

Table 45: ICM vs non-ICM Couples

Comparison between ICM vs non-ICM Couples				
	Inter-Caste Marriage		Within Caste Marriage	
	Men	Women	Men	Women
% of total sample	4.83	4.83	93.12	32324
% of Total Couples	5.3	5.3	94.09	94.09
Age	39.62	34.7	40.11	35.11
Education				
No Education 0	22.99	41.47	21.86	39.5
Less than Primary	8.66	6.1	9.1	7.28
5th Pass	9.03	8.61	9.29	9.01
8th Pass	27.04	21.64	26.7	22.7
Secondary	15.07	10.3	14.31	9.88
H.Secondary and Diploma (<3yrs)	6.29	5.64	9.25	6.4
Bachelors(BA,Bsc,Diploma 3+)	7.58	4.27	6.85	3.86
BTech,MBBS,MD,CA,PhD	3.33	1.97	2.64	1.37
Primary Occupation Status				
%age of Population Covered	94.65	20.46	92.57	22.77
Salaried/Professional	21.12	22.53	21.91	14.96
Small Business/Artisan	14.78	5.58	15.77	7.18
Cultivators	24.34	22.27	24.27	20.01
Non Agri Wage Labour	27.72	22.09	26.43	23.34
Agri Wage Labor	12.04	27.52	11.63	34.51
Occupation Types				
%age of Population Covered	99.93	27.71	99.93	31.97
Professional	5.14	5.78	5.07	5.1
Admin,Exec,Managers	0.77	0.06	1.27	0.14
Clerical	6.61	4.65	5.96	2.45
Sales	13.86	4.85	14.92	4.37
Service Providers	3.15	5.71	3.99	4.8
Farmers,Cultivators	29.86	53.69	29.28	51.64
Labourers (Non-Agri)	40.62	25.26	39.5	31.5
Earnings				
%age with positive income	70.19	21.73	71.36	27.12
Average Annual Income	63923	26930	66209	21956
Imputed Earnings				
%age with positive income	99.17	99.55	99.08	99.88
Average Annual Income	53172	21078	55197	19612

Source: Author's calculation using IHDS 2011 datasets. Design weights are used to estimate these values. This table outlines the difference between ICM and non-ICM couples.

Table 46: Comparison Among Different Castes

	Comparison Across Different Castes											
	Brahmins		FC (except Brahmins)		OBC		Dalits(SC)		Adivasi(ST)		Muslims	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
% of total samples (N)	4.91	4.91	16.54	16.54	33.16	33.16	21.2	21.2	8.14	8.14	11.8	11.8
% of total Couples	4.69	4.69	15.53	15.53	35.38	35.38	22.39	22.39	7.82	7.82	12.12	12.12
Age	41.54	36.77	41.43	36.13	40.19	35.28	39.28	34.33	38.99	34.34	39.39	34.07
Education												
No Education 0	3.47	15.37	8.52	19.44	19.94	40.92	29.57	50.56	36.77	57.5	31.56	44.44
Less than Primary	2.78	3.23	6.09	6.27	8.9	7	11.05	7.78	11.13	8.75	11.63	9.1
5th Pass	4.81	12.86	6.61	7.79	9.33	9.4	10.79	9.18	10.94	7.26	11.16	8.75
8th Pass	22.52	27.07	26.05	27.6	29	22.56	26.24	19.66	23.35	17.16	24.79	23.81
Secondary	20.65	14.34	19.64	17.02	15.44	9.6	11.18	7.05	8.04	4.38	9.97	6.91
H.Secondary and Diploma (<3yrs)	17.42	11.14	14.52	10.53	9.14	6.12	6.29	3.85	5.37	3.79	5.15	4.39
Bachelors(BA,Bsc,Diploma 3+)	18.69	11.48	13.64	8.28	5.97	3.4	3.53	1.22	3.45	0.86	4.2	2.03
BTech,MBBS,MD,CA,PhD	9.66	4.52	4.93	3.07	2.28	1	1.36	0.7	0.95	0.31	1.54	0.57
Primary Occupation Status												
%age of Population Covered	89.81	10.46	94.21	17.04	91.87	23.21	93.3	29.69	96.88	34.98	89.88	13.05
Salaried/Professional	43.84	55.19	34.82	26.64	19.03	12.99	18.11	11.25	13.32	8.63	15.49	13.28
Small Business/Artisan	20.91	13.88	18.76	8.85	15.82	7.14	10.14	5.09	5.07	2.53	27.08	16.82
Cultivators	26.18	21.18	29.58	34.99	29.86	25.25	13.27	8.85	34.22	28.1	13.77	5.18
Non Agri Wage Labour	8.18	5.26	12.39	15.4	25.56	21.58	37.57	27.08	29.18	22.06	33.12	38.34
Agri Wage Labor	0.89	4.49	4.45	14.13	9.73	33.04	20.9	47.74	18.21	38.69	10.54	26.38
Occupation Types												
%age of Population Covered	99.93	13.8	99.96	20.97	99.92	32.95	99.93	42.19	99.92	52.46	99.92	17.81
Professional	12.3	26.21	7.99	12.05	4.63	4.18	2.95	2.39	2.92	1.89	4.81	7.11
Admin,Exec,Managers	4.02	0.5	2.54	0.37	1	0.12	0.68	0.04	0.61	0.12	0.58	0.23
Clerical	14.29	11.51	11.01	3.87	4.95	2.47	4.69	1.7	2.93	1.72	3.24	1.51
Sales	20.29	10.47	19.5	6.94	14.85	4.38	8.75	2.98	4.23	1.24	24.36	9.6
Service Providers	5.88	5.51	3.83	6.21	3.26	3.91	5.5	6.67	3.96	2	2.53	4.56
Farmers,Cultivators	20.65	21.61	28.84	43.81	31.42	54.65	29.76	55.49	41.69	63.46	18.85	31.26
Labourers (Non-Agri)	22.56	24.19	26.3	26.75	39.88	30.29	47.66	30.74	43.66	29.57	45.63	45.74
Earnings												
%age with positive income	59.41	10.21	59.55	14.53	68.27	27.18	85.16	38.76	82.34	45.98	69.3	14.98
Average Annual Income	127496	65939	102873	42908	59964	19810	54691	16996	42692	16358	58904	22186
Imputed Earnings												
%age with positive income	98.75	100	99.14	100	99.12	99.96	99.04	99.96	99.12	99.99	99.16	99.34
Average Annual Income	90431	32623	74418	27703	48669	17872	50429	16264	38232	14219	49407	15280

Source: Author's calculation using IHDS 2011 datasets. Design weights are used to estimate these values. This table outlines the differences across caste groups.

Table 50: Wife and Husband's Father Education Correlation

	Overall	Place		Inter-Caste Marriage		Religion				Caste_Group					
		Rural	Urban	ICM	Non-ICM	Hindu 1	Muslim 2	Christian 3	Sikh 4	Brahmin 1	Forward caste 2	OBC 3	Dalit 4	Adivasi 5	Muslim 6
Wife's and Father-in-Law's Education															
Polychoric	0.574	0.511	0.562	0.624	0.57	0.577	0.497	0.531	0.646	0.469	0.569	0.515	0.48	0.518	0.495
Confidence Interval (C.I)	[.569,.578]	[.504,.517]	[.555,.568]	[.606,.642]	[.566,.575]	[.572,.582]	[.482,.512]	[.504,.558]	[.622,.669]	[.45,.489]	[.560,.579]	[.507,.523]	[.468,.492]	[.498,.539]	[.48,.51]
Spearman	0.473	0.405	0.502	0.51	0.471	0.478	0.38	0.466	0.549	0.457	0.523	0.415	0.366	0.366	0.38
C.I.	[.465,.481]	[.394,.416]	[.488,.516]	[.474,.545]	[.462,.479]	[.469,.487]	[.354,.406]	[.408,.519]	[.499,.595]	[.419,.494]	[.504,.542]	[.4,.43]	[.346,.386]	[.334,.397]	[.353,.406]
Obs	34174	22766	11408	1677	32324	27963	4169	764	814	1704	5741	11511	7359	2826	4095
Wife's and Father-in-Law's Education (excluding cases wife.educ=husb.fathers.educ=0)															
Polychoric	0.225	0.066	0.361	0.277	0.221	0.227	0.083	0.397	0.416	0.328	0.382	0.138	0.015	0.003	0.082
C.I.	[.218,.231]	[.057,.075]	[.352,.371]	[.248,.307]	[.214,.228]	[.22,.234]	[.061,.104]	[.363,.43]	[.381,.451]	[.305,.351]	[.37,.395]	[.126,.15]	[-.002,.031]	[-.026,.032]	[.06,.103]
Spearman	0.233	0.083	0.36	0.287	0.23	0.24	0.049	0.35	0.416	0.37	0.396	0.145	0.044	-0.014	0.048
C.I.	[.221,.245]	[.067,.1]	[.343,.378]	[.233,.339]	[.218,.243]	[.227,.253]	[.011,.087]	[.282,.414]	[.349,.479]	[.327,.412]	[.372,.419]	[.123,.166]	[.015,.074]	[-.066,.038]	[.01,.086]
Obs	10186	4962	5224	546	9584	8190	1075	473	298	959	2710	3192	1287	374	1056

Table 51: Husband and Wife's Father Education Correlation

	Overall	Place		Inter-Caste Marriage		Religion				Caste_Group					
		Rural	Urban	ICM	Non-ICM	Hindu 1	Muslim 2	Christian 3	Sikh 4	Brahmin 1	Forward caste 2	OBC 3	Dalit 4	Adivasi 5	Muslim 6
Husband's and Father-in-Law's Education															
Polychoric	0.548	0.503	0.532	0.568	0.548	0.555	0.459	0.547	0.526	0.463	0.511	0.484	0.482	0.503	0.459
Confidence Interval (C.I)	[.544,.552]	[.497,.509]	[.525,.539]	[.549,.587]	[.543,.552]	[.551,.560]	[.444,.474]	[.521,.573]	[.498,.555]	[.444,.483]	[.5,.521]	[.475,.492]	[.471,.494]	[.483,.523]	[.444,.473]
Spearman	0.457	0.395	0.488	0.458	0.457	0.469	0.355	0.453	0.397	0.458	0.46	0.402	0.36	0.357	0.353
C.I.	[.448,.465]	[.384,.406]	[.474,.502]	[.42,.495]	[.448,.465]	[.459,.478]	[.328,.381]	[.395,.508]	[.337,.453]	[.42,.495]	[.44,.48]	[.387,.418]	[.34,.38]	[.324,.389]	[.326,.379]
Obs	34174	22766	11408	1677	32324	27963	4169	764	814	1704	5741	11511	7359	2826	4095
Wife's and Father-in-Law's Education (excluding cases husb.educ==wife.fathers.educ==0)															
Polychoric	0.332	0.235	0.41	0.343	0.333	0.35	0.141	0.448	0.355	0.409	0.408	0.271	0.194	0.178	0.141
C.I.	[.327,.338]	[.227,.243]	[.402,.419]	[.317,.368]	[.327,.339]	[.344,.356]	[.122,.161]	[.417,.479]	[.319,.391]	[.388,.429]	[.396,.42]	[.261,.282]	[.18,.209]	[.152,.204]	[.121,.16]
Spearman	0.322	0.22	0.412	0.311	0.323	0.342	0.124	0.39	0.275	0.431	0.394	0.267	0.179	0.178	0.121
C.I.	[.312,.332]	[.206,.234]	[.396,.428]	[.262,.358]	[.312,.333]	[.331,.354]	[.09,.159]	[.326,.45]	[.205,.343]	[.391,.469]	[.371,.416]	[.248,.285]	[.153,.204]	[.134,.221]	[.085,.155]
Obs	28253	17815	10438	1379	26738	23336	3108	716	698	1666	5364	9755	5614	1914	3050

Table 52: Wage Earnings Correlation

	Overall	Place		Inter-Caste Marriage		Religion				Caste_Group						
		Rural	Urban	ICM	Non-ICM	Hindu 1	Muslim 2	Christian 3	Sikh 4	Brahmin 1	Forward caste 2	OBC 3	Dalit 4	Adivasi 5	Muslim 6	
Gross Annual Imputed Wage Earnings																
Polychoric	0.704	0.627	0.693	0.698	0.708	[.694,.71]	[.578,.596]	[.678,.694]	[.795,.808]	[.755,.769]	[.77,.784]	0.777	0.599	0.552	0.745	0.588
Confidence Interval (C.I)	[.696,.712]	[.618,.635]	[.685,.702]	[.690,.706]	[.700,.716]	[.694,.71]	[.578,.596]	[.678,.694]	[.795,.808]	[.755,.769]	[.77,.784]	[.77,.784]	[.591,.608]	[.542,.561]	[.737,.752]	[.579,.597]
Spearman	0.521	0.453	0.546	0.524	0.521	0.517	0.348	0.604	0.438	0.761	0.559	0.522	0.522	0.416	0.63	0.342
C.I.	[.505,.537]	[.433,.472]	[.509,.581]	[.437,.6]	[.504,.537]	[.5,.534]	[.265,.427]	[.508,.686]	[.285,.568]	[.685,.820]	[.505,.608]	[.505,.608]	[.495,.549]	[.383,.447]	[.596,.662]	[.257,.421]
Obs	8006	6555	1451	307	7660	7096	451	198	127	153	685	2772	2772	2545	1275	442
Gross Annual Imputed Wage Earnings																
Polychoric	0.439	0.319	0.377	0.467	0.439	0.446	0.339	0.394	0.482	0.477	0.463	0.4	0.4	0.394	0.449	0.34
C.I.	[.434,.444]	[.314,.324]	[.372,.382]	[.462,.471]	[.434,.444]	[.441,.451]	[.334,.344]	[.389,.399]	[.477,.487]	[.472,.481]	[.459,.468]	[.395,.405]	[.395,.405]	[.389,.399]	[.444,.454]	[.335,.345]
Spearman	0.502	0.347	0.396	0.492	0.502	0.515	0.399	0.396	0.261	0.551	0.522	0.516	0.516	0.446	0.542	0.398
C.I.	[.494,.51]	[.335,.358]	[.38,.411]	[.454,.528]	[.494,.51]	[.506,.524]	[.373,.425]	[.335,.455]	[.195,.324]	[.517,.584]	[.503,.541]	[.502,.529]	[.502,.529]	[.428,.465]	[.515,.567]	[.372,.424]
Obs	33828	22541	11287	1646	32010	27714	4100	756	803	1682	5699	11411	11411	7292	2802	4027

Table 53: Occupation Correlation

	Overall	Place		Inter-Caste Marriage		Religion						Caste_Group			
		Rural	Urban	ICM	Non-ICM	Hindu 1	Muslim 2	Christian 3	Sikh 4	Brahmin 1	Forward caste 2	OBC 3	Dalit 4	Adivasi 5	Muslim 6
						OCCUPATION TYPE									
Polychoric	0.434	0.39	0.54	0.429	0.436	0.44	0.387	0.428	0.477	0.458	0.491	0.436	0.323	0.493	0.39
Confidence Interval (C.I)	[.425,.443]	[.378,.401]	[.522,.558]	[.383,.474]	[.427,.446]	[.43,.45]	[.346,.428]	[.368,.487]	[.405,.548]	[.403,.513]	[.467,.516]	[.42,.451]	[.303,.343]	[.47,.516]	[.349,.431]
Spearman	0.352	0.328	0.447	0.424	0.349	0.354	0.345	0.387	0.332	0.447	0.405	0.333	0.287	0.405	0.346
C.I.	[.336,.369]	[.309,.346]	[.413,.481]	[.345,.496]	[.332,.366]	[.336,.371]	[.277,.41]	[.283,.483]	[.19,.46]	[.344,.54]	[.357,.451]	[.305,.361]	[.254,.319]	[.363,.445]	[.277,.412]
Obs	10843	8732	2111	452	10336	9548	669	281	169	258	1208	3911	3023	1592	655
						PRIMARY ACTIVITY STATUS									
Polychoric	0.728	0.676	0.734	0.789	0.724	0.732	0.663	0.731	0.69	0.605	0.735	0.692	0.698	0.714	0.661
C.I.	[.722,.734]	[.668,.684]	[.72,.748]	[.765,.812]	[.718,.73]	[.726,.739]	[.634,.693]	[.694,.769]	[.632,.748]	[.544,.666]	[.719,.752]	[.681,.703]	[.684,.711]	[.698,.731]	[.631,.691]
Spearman	0.649	0.606	0.622	0.709	0.645	0.653	0.586	0.54	0.514	0.488	0.673	0.609	0.615	0.649	0.586
C.I.	[.635,.662]	[.589,.623]	[.592,.651]	[.649,.761]	[.631,.658]	[.638,.667]	[.521,.645]	[.439,.628]	[.371,.633]	[.366,.593]	[.635,.708]	[.583,.633]	[.587,.642]	[.612,.683]	[.52,.646]
Obs	7100	5416	1684	308	6751	6185	431	218	124	175	872	2481	1968	1031	423

9.3 A.3 Methodologies

9.3.1 Stratification in NSS AIDIS

A two stage-stratified sample design with villages and urban blocks as First stage units (FSU) and households as Second-stage units (SSU) is adopted. The list of villages are drawn from the recent most available census information. Since 1971-72 there was an increase in the size of sample to get the estimates not only at all-India and individual states but also at broader regions within the states. The rural/urban area was divided into agricultural regions⁷⁴ having similar crop patterns and population density. Each region was further divided into strata by grouping geographically contiguous tehsils/districts having similar characteristics based on altitude, transport and communication facilities. The effort was to keep the population dependent on agriculture same across all strata⁷⁵. Each stratum was grouped into two independent sub-samples of 12 villages each⁷⁶, on the basis of circular systematically with equal probability. All the households in sample villages were divided into 4 sub-strata (non-cultivators, small, medium and large classes of cultivator) on the basis of land area operated. The non-cultivator sub-stratum was further divided into agricultural labourers, artisans and other non-cultivators based on their means of livelihood. The cut-off points for small, medium and large cultivators was decided on regional basis to equalise the total land area operated in each of the sub-strata at regional level. From each sub-stratum in a village on average 3 households were selected linear systematically with specified intervals and random starts, which were pre-determined to design self-weighting at the regional level for each sub-stratum. The estimate of the state total Y for a given y_{sitj} value of characteristic, where j denotes household, t for sub-stratum, i for village and s- rural stratum is based on following formula:

$$Y = \sum_{t=1}^4 \sum_{s=1}^k M_{st} \sum_{i=1}^{n_s} \sum_{j=1}^{h_{sit}} Y_{sitj} \quad (3)$$

where M^{st} is the multiplier for a given stratum s and sub-stratum t. It varies from one sub-stratum to other but remains constant over the strata in a region. k is the number of rural strata in state, n_s is the number of sample villages from a given stratum s and h_{sit} is the number of sample household from a stratum s, substratum t and village i.

In 1991-92 the major change in sampling method was that land possession and indebtedness status both were used. The stratum within each agro-economic regions were decided based on the population. In rural areas, each region with district population less than 1.8 million⁷⁷ formed a single basic stratum. Larger districts were divided into two or more strata, by grouping contiguous tehsils to keep similar population density and crop pattern. In Urban areas each million population city formed a separate stratum by itself. The remaining towns of each region were grouped to form three different strata based on the population criteria - 200k-1 million 50k-200k and less than 50k. The selection of sample villages was done by Probability proportional to size with replacement (PPSWR) with population as the size of variable and the sample blocks were selected by simple random sampling without replacement. Both are in the form of two independent sub-samples.

In Rural sector 7 AIDIS sub-strata were formed based on - "land possession" and "indebtedness status" of the households.⁷⁸ Independently 1,1,1,2,1,1, and 2 households were selected circular systematically from 1-7 sub-strata respectively. In Urban sector 7 AIDIS sub-strata were formed based on "monthly per capita expenditure (MPCE)" and "indebtedness" of households.⁷⁹ Independently 1,1,1,1,2,2, and 1 households were selected circular systematically

⁷⁴66 in 1971-72, 78 in 1991

⁷⁵In 1971-72 there were 379 rural strata

⁷⁶except few smaller states, and if the stratum was very big it was divided into manageable investigation zones

⁷⁷1.5 million for Assam

⁷⁸First, based on land possession 4 groups were formed, namely non-cultivators (.005 acres), small (.005-X), medium (X-Y) and large ($\geq Y$) cultivators with increasing land area. X and Y varied with state to ensure that in all the three cultivators group total land area remained same. The first two groups were sub-divided into "indebted" and "non-indebted" groups to form AIDIS sub-strata 1-4. Medium and Large cultivators were then merged and subdivided into 3 AIDIS sub-strata- "indebted to institutional agencies with or without being indebted to non-institutional agencies", "indebted to non-institutional agencies alone" and "non-indebted".

⁷⁹Households were first grouped into 3 classes - less than x, x-y and greater than y. The cut-off x and y was decided at state level, based on NSS consumption survey to allocate 30%,60% and 10% of the urban population of the state. Similar to in rural sector this three group were divided into 7 AIDIS sub-strata based on indebtedness.

from 1-7 sub-strata respectively. Selection of households in each sub-stratum was done by Simple Random sampling without replacement (SRSWOR).

2002-03: The stratification exercise was similar to previous round of survey with slight modifications. It became multi-stage design with First stage, Intermediate stage and Ultimate stage units. The intermediate stage of sampling was in case of large FSU's. Ultimate stage unit was households, which was called Second stage unit in previous rounds. Two special stratas were formed from all small villages (< 50 population) and big villages (> 15000 population) in a state. Then in the rest of the villages/urban blocks stratum were formed as in 1991-92. Instead of 1.8 million district population criteria for strata formation in rural sector 2 million (as per 1991 census) or 2.5 million (as per 2001 census) was used. Sub-strata formation was similar as in 1991-92, with a slight change in rural sector where the focus changed on deciding X and Y - from keeping the same land area to fixing of certain population percentage. X and Y were determined at state level to ensure that 40% of households possess land area less than X, 40% possess land area between X-Y and 20% possess land area greater than Y. In Urban sector, based on MPCE, 4 groups were formed (instead of 3 in 1991-92)- less than A, A-B, B-C, > C. The cut-off points A, B and C was decided at state level to ensure 30% households fall into first, second and third groups with 10% falling into fourth. The first two groups were further divided into indebted and non-indebted. The last two groups were merged and further sub-divided into institutional indebted, non-institutional indebted and non-indebted. 14 households were selected from each sample village/block.

9.3.2 Converting from Household Level to Individual Level

I use non-parametric approach to estimate the average adult size at HH Level (AVAH) for 1961-62, 1971-72 and 1981-82 survey years from the next round of survey years.⁸⁰ Essentially I am assuming the decadal rate of change of AVAH during 1991-92 and 2002-03 to be same to predict AVAH for 1981-82. The assumption might fail if there is a big change in demography which makes these two decades growth rate entirely different. Since demographic factors usually experience slow changes except in extreme cases, like some natural calamity or disease wiping out a big share of population, war, or strict government policy like One child Policy. India didn't experience such drastic measures to be considered as demographic shock. The wars between 1961-1981⁸¹ were not on a large scale in terms of human capital. Forced sterilization during Emergency (1975-77) resulted in 8 millions extra sterilization (as per Shah Commission Third Report 19789:2007) which was 1.5% of the 1971 census population is a potential threat to the assumption as it can bring negative change in the adult population after 20 years, i.e. 1991-2001 decade. Since household surveys ask for people living in the house, a simple migration can also impact AVAH. Census provides total migrating population with respect to total population which stand at 30.8%, 28.7%, 29.4%, 26.6% and 29.3% in 1961, 1971, 1981, 1991 and 2001 respectively. The decadal variation is not too large to produce concern. Keeping the above potential issues in mind, I proceed with the estimation of AVAH.

The objective is to re-calculate the HH level wealth to individual level wealth. First Step is to estimate AVAH from next round of micro-surveys files. The procedure I followed is:

- 1) Assuming same decadal rate at decile level, predict AVAH and Average HH size (AVH) for 1981. 2) Correct the predicted 1981 AVH and AVAH values using "correction factors". These correction factors are:
 - a) AVH: A factor to make the predicted population level equal to the representative population of Survey.
 - b) AVAH: A factor to make the predicted adult population share same as in census.^{82,83}
- 3) Use the predicted AVH and AVAH for 1981 and repeat 1) and 2) to predict for year 1971.
- 4) Repeat 3) to predict for 1961 level.

The second step involves generating full distribution using Generalized Pareto Interpolation (using gpinter)- first at HH level and then applying the decile wise estimated AVAH to recalibrate the p values, bracket average, threshold values at individual level and re-running gpinter programme at individual level.

Using predicted AVH and AVAH 1981 level, the above process is repeated for 1971 level. And then using predicted 1971 level, 1961 level is predicted.

9.3.3 Sampling Methodology in Surveys

1) Systematic Sampling implies selecting a sample from a given population based on a skip pattern with a sampling interval. So for example, every k^{th} element from a sequential population data is selected with a sampling interval of $k=N/n$ where n- number of selected units, N- Population units. There can be two types of systematic sampling schemes. First Linear systematic sampling which is used when N/n is an integer and k samples of exactly n units in a

⁸⁰The parametric approach imply estimating the adult population per HH using regression analysis with potential dependent variables like -age, education, place of residence, caste etc. But it will require micro-data which is lacking in our case

⁸¹1962 India-China war- resulted in 10,000 (0.002% of total population) human loss; 1965 India-Pakistan war resulted in 3000 casualties; 1971-72 India Pakistan war - 4000 casualties.

⁸²The data for adult population share is taken from the website: <https://www.populationpyramid.net/india/>, I use the same adult share for Rural and Urban area. It will be better to use separate. Lack of good source has resulted in compromising here.

⁸³Different sampling of adults in Census and survey might be a threat. For 1981, I have comparison between survey sampling and census at household level, which i use here. For 1971 and 1961, I use census population, which means survey population was perfect to the census population

systematic manner can be collected with probability $1/k$. However, if N/n is not an integer then it is a problem which leads to the second type of systematic sampling, called Circular systematic sampling scheme. In this case once n is fixed, the skip k is taken as the minimum integer of N/n . There is a possibility of N possible distinct samples instead of k and hence probability $1/N$. The advantages of using systematic sampling over random sampling is that ordering of the population ensures representativeness and even spread of sample. It can be more efficient than SRSWOR when the ordering is related to the characteristics of the variable of interest. One limitation is that it requires the complete list with proper ordering. A bad ordering may lead to inefficient sample.

9.3.4 Imputation of Wage Earnings

The wage information for 30% of men and 80% of women is absent in the IHDS-2011 dataset. This is not surprising since a large section of population is engaged in informal sector with irregular payments. The labour force participation of women is lower and most of them are housewives with no earnings. I impute the wages for men and women separately from using the information from the population for which wage information is available. **Husband:** The regression specification for men

$$\text{Log_Annual_wage}_{ih} = \alpha_0 + \alpha_1 X_i + \epsilon_i \quad (4)$$

where X includes education, age, age'square, occupation, household assets, total number of children, parents' education, parents'-in-law education, state dummies, caste dummies.

Wife: The regression specification for men

$$\text{Log_Annual_wage}_{ih} = \alpha_0 + \alpha_1 X_i + \epsilon_i \quad (5)$$

where X includes education, age, age'square, household assets, total number of children, parents' education, parents'-in-law education, state dummies, caste dummies. The output of the regression is presented below:

Table 54: Regression Output table

X (Regressors)	(1) WIFE	(2) HUSBAND
educ_cat==1	-0.0469 (0.0375)	-0.0636*** (0.0203)
educ_cat==2	-0.113*** (0.0400)	-0.0229 (0.0211)
educ_cat==3	-0.155*** (0.0346)	-0.0644*** (0.0166)
educ_cat==4	0.115** (0.0571)	-0.0256 (0.0214)
educ_cat==5	0.809*** (0.0683)	-0.0308 (0.0257)
educ_cat==6	0.959*** (0.0786)	0.0957*** (0.0312)
educ_cat==7	1.548*** (0.0944)	0.230*** (0.0425)
age	0.0456*** (0.00956)	0.0269*** (0.00457)
age2	-0.000579*** (0.000127)	-0.000351*** (5.52e-05)
Urban residence from census 2011	0.520*** (0.0305)	0.294*** (0.0153)
Total household assets (0-33)	0.0250*** (0.00275)	0.0425*** (0.00148)
occupation_cat==2		0.342*** (0.0483)
occupation_cat==3		0.0111 (0.0329)
occupation_cat==4		-0.383*** (0.0405)
occupation_cat==5		-0.263*** (0.0379)
occupation_cat==6		-0.320*** (0.0348)
occupation_cat==7		-0.281*** (0.0311)
primary_status_cat==2		-1.125*** (0.0308)
primary_status_cat==3		-1.369*** (0.0230)
primary_status_cat==4		-0.352*** (0.0187)
primary_status_cat==5		-0.581*** (0.0257)
Caste_Category==2	-0.0288 (0.0846)	-0.0585* (0.0315)
Caste_Category==3	0.00596 (0.0817)	-0.101*** (0.0308)
Caste_Category==4	0.140* (0.0824)	-0.0258 (0.0316)
Caste_Category==5	0.0398 (0.0864)	-0.120*** (0.0356)
Caste_Category==6	-0.262** (0.113)	-0.222*** (0.0548)
total_children	-0.0238*** (0.00807)	0.00891** (0.00445)
Constant	8.453*** (0.258)	10.49*** (0.114)
Observations	9,135	21,997
R-squared	0.361	0.550
Controls	YES	YES

9.3.5 Correlation Coefficients

Pearson correlation coefficient

Pearson correlation coefficient (also referred to as Pearson product-moment correlation coefficient) is the most commonly used correlation coefficient. This provides the strength of linear association between two variables. It is given by the covariance between two variables divided by standard deviation of each variable, i.e., for variables X and Y ,

$$\rho_{pearson} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Polychoric correlation coefficient

Polychoric correlation is used to estimate Pearson correlation coefficient between two continuous, bivariate-normally distributed variables from categorized versions of those variable (Hershberger, 2005). This coefficient, thus, measures the association between two ordinal variables. Maximum likelihood method is used for calculation. For the special case where two variables are binary, then the coefficient is called Tetrachoric correlation.

Spearman's rank correlation coefficient

The spirit of Spearman's rank correlation is same as polychoric correlation coefficient, i.e. it estimates the correlation between two ordinal (or rank-ordered) variables. However, the intuition is much similar to Pearson correlation coefficient and the estimation is much simpler than polychoric correlation coefficient. While Pearson correlation coefficient measures a linear relationship, it measures a *monotonic* relationship between two variables. No assumptions (e.g. bivariate normality) is required, as compared to polychoric coefficient. For ranked variables X_R and Y_R that have no tied ranks (i.e. no same rank is assigned to individuals more than once), Spearman's rank correlation coefficient is given by $\rho_{spearman} = 1 - \frac{6\sum_i d_i^2}{n(n^2-1)}$, where d_i is the difference in paired ranks and n is the number of observations. If they have tied ranks, then $\rho_{spearman}$ is calculated in the same way as Pearson's correlation coefficient.