Multigenerational Persistence and The Great Gatsby Relation for India**

Chinmayi Srikanth†

Shubhasis Dey§

Abstract

Extant literature on multigenerational mobility in India establishes the existence of a 'grandfather effect'. By using distinct models to measure intergenerational and multigenerational educational persistence, this paper contends that this effect can be direct, indirect, or non-existent depending on the identity group - caste or religion - to which the family belongs. The study also finds evidence for differences in persistence due to regional factors. This paper is also the first to develop an identification strategy for multigenerational mobility in India through the use of the abolition of untouchability in 1955 as an instrument for the father's education. The study is the first to explore *The Great Gatsby* relation for regions within India using caste and religion-based identifiers. This study finds that the relation holds regardless of the identity group considered.

Keywords— Persistence, multigenerational mobility, identity, education, untouchability, India
JEL Classification: I24, J15, J62

*We are thankful for useful comments from François Bourguignon.

**This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of interest: none

†Doctoral Student, Economics Area, Indian Institute of Management Kozhikode, IIMK Campus P.O, Kunnamangalam, Kerala - 673570, India. +91-8056696020 srikanthc12fpm@iimk.ac.in

§Professor, Economics Area, Indian Institute of Management Kozhikode, IIMK Campus P.O, Kunnamangalam, Kerala - 673570, India. +91-495 2809115 s.dey@iimk.ac.in

1 Introduction

The existence of the caste system in India has led to the creation of a stratified society based on an underlying hierarchy that was primarily occupation-based. While this system of occupation-based hierarchy has been diluted by the growing complexity of economies and consequently jobs, the concept of caste and its inherent social hierarchy continues to perpetuate in occupational and educational spheres.

Categorical social identities, such as caste identities, are those identities that people assume at birth. An inheritance of caste identity also means an inheritance of the disadvantages or advantages that come with it. Such inheritance are indeed a multigenerational, rather than an intragenerational or intergenerational phenomenon. Studying this phenomenon is particularly challenging given the paucity or even absence of data at a multigenerational level (Güell, Rodríguez Mora, & Solon, 2018), especially in the Indian context. Kundu and Sen's (2022) study is the only exception. The study finds evidence of a 'grandfather effect' for India, where the effect of the grandfather's outcome on the grandson's outcome is independent of the father's outcome, where outcome refers to education and occupation. The study, however, makes an *à priori* assumption about the model specification for multigenerational mobility. Whether the grandfather affects the grandson directly or indirectly through the father is a proposition that needs to be tested to assess the persistence of categorical affiliations over generations.

The purpose of this study is to test whether the 'grandfather effect' is direct, indirect, or even non-existent for each caste and religious group. This is based on the premise that each identity group has unique cultural influences that affect the future generations in ways that cannot be generalized using a straitjacket model specification. This is explored using data from the India Human Development Survey (IHDS) 2011-12. IHDS is a rich source of data of Indian households on socio-economic variables. The household questionnaire asks questions on the education and occupation of the household head's father, while the individual household roster records the education, age, sex etc, of each household member. While IHDS does collect data on non-resident members, this study considers only resident sons of the household, since it controls for household environment and geographical location. Three generations of males in the same household are mapped in this study namely, the grandfather (G0), father (G1), and son/ grandson (G2).

Through this study, we find that each identity group, whether based on caste or religion, experiences multigenerational or intergenerational transmission of education differently. This difference is also determined by the region to which a household belongs. While the use of grandfather's outcome as an instrument for the father's outcome is widespread in literature (Boserup, Kopczuk, & Kreiner, 2013; Lindahl, Palme, Sandgren-Massih, & Sjögren, 2014), we explore a novel identification strategy- the abolition of untouchability marked by the ratification of 'The Untouchability (Offenses) Act 1955'- that is likely to have affected the G0 generation's decision to educate G1 generation.

From a policy perspective, understanding the relationship between inequality and intergenerational persistence is crucial. *The Great Gatsby* curve establishes that countries with high inequality have high intergenerational persistence and therefore low social mobility (Corak, 2013). This study explores the same relation for each of the zonal councils in India-namely, North, North-East, Central, East, West, and South.

The study, therefore, focuses on three major aspects of categorical identity: its persistence, how differently its transmission operates for different identity groups, and its implications for macroeconomic indicators such as inequality.

This paper is organized as follows: Section 2 offers a brief review of literature on intergenerational and multigeneration mobility, Section 3 discusses the data source and construction of G0-G1-G2 triads. In Section 4, the models for educational transmission are discussed; Section 5 discusses persistence for each identity group based on the methods proposed in the previous section, while Section 6 tests *The Great Gatsby* for regions within India. Section 7 entails a discussion based on the findings in this paper.

2 Review of Literature

Becker and Tomes (1979) is credited with the first model of intergenerational mobility, where the parents' allocation of endowments towards human and non-human capital of the children subject to their budget constraints, affect the latter's outcomes. The underlying premise of the model is that parents' altruism

is what drives the allocation of wealth to their children (Becker, 1974; Becker & Tomes, 1979). This model was tested by Behrman and Taubman (1985) where they find no empirical evidence for the model. Contrary to Becker and Tomes (1979), several papers have discussed non-altruistic motives of parents (Bernheim, Shleifer, & Summers, 1986; Kotlikoff & Spivak, 1981; Pollak, 1988). Recently, parental altruism limited by their economic position or 'limited parental altruism' has been explored with respect to human capital expenditure on children (Das, 2007).

Although much of the methodological advances in social mobility literature has used earnings and wealth as the key variables[1], a large part of extant literature study occupational (Corak, Lindquist, & Mazumder, 2014; Heath & Zhao, 2021; Li & Zhao, 2017; Wu & Treiman, 2007) and educational mobility (Alesina, Hohmann, Michalopoulos, & Papaioannou, 2021; Ferreira et al., 2012; Hertz et al., 2008; Narayan et al., 2018) because of lack of reliable data on earnings (Iversen, Krishna, Sen, et al., 2016).

Mobility literature has largely been confined to within-country analyses, i.e; Chetty, Hendren, Kline, and Saez (2014); Chetty, Hendren, Kline, Saez, and Turner (2014) for the USA, Checchi, Fiorio, and Leonardi (2013) for Italy, Azam (2019); Azam and Bhatt (2015); Kundu and Sen (2022) for India, and Gong, Leigh, and Meng (2012) in China, Lambert, Ravallion, and Van de Walle (2014) for Africa, to name a few. Neidhöfer, Serrano, and Gasparini (2018) is a cross-country study of Latin America.

Studies focusing on mobility in India such as Azam (2019), find that intergenerational mobility follows a hierarchy that mimics one's position in the social (caste) hierarchy, with lower caste groups reflecting high persistence and lower mobility and higher caste groups showing lower persistence and therefore, higher mobility. Azam and Bhatt (2015) discover a decline in educational persistence in India, which can be attributed to public education policies. Dhar, Jain, and Jayachandran's (2019) unconventional study uses intergenerational transmission to study gender attitudes in India.

Predicting long-term intergenerational persistence - one that involves more than two generations - based on data for two generations alone, leads to extrapolation errors that often manifest in the overestimation of educational and occupational mobility (Lindahl, Palme, Massih, & Sjögren, 2015). Multigenerational mobility is, therefore, more desirable, where data is available. The underlying mechanism of generational persistence is also influenced by the model used (Clark, 2012; Clark & Cummins, 2014; Solon, 2013; Stuhler, 2013). Considerations such as the size of the dataset in terms of the number of generations covered and the model specification are crucial in influencing our understanding of multigenerational persistence.

Lindahl et al. (2015) use Swedish data for four generations of educational attainment and three generations of earnings. Their study establishes that long-run intergenerational persistence is underestimated when only two generations are taken into account. As far as developing countries are concerned, studies on multigenerational mobility are limited (Solon, 2018). Kundu and Sen (2022) make use of data for three generations and is the only study for India on multigenerational persistence and establishes the existence of a 'grandfather effect' for India. Multigenerational mobility literature has established specifications that either assumes a direct effect of the grandfather's outcomes on the grandson's outcomes (Solon, 2014) or an indirect effect that operates through the father's outcomes. This study is motivated by the need to explore whether the 'grandfather effect' operates directly or indirectly, and it does so using model specifications that account for differences in caste and religious identity. While the use of *grandfather's outcome* as an instrument for the *father's outcome* is widely used in literature (Boserup et al., 2013; Lindahl et al., 2014), studies have also used identification strategies that are unique to their study setting[2].

By drawing parallels between Borjas's (1992) 'ethnic capital'[3] and the social capital of identity groups, this study treats identity as paramount to understanding differences in multigenerational transmission

[1] Refer Black and Devereux (2010) for a comprehensive review of methodologies based on earnings mobility.

[2] In extant literature on educational mobility, laws on compulsory schooling (Angrist & Krueger, 1991), unemployment rate (Arkes, 2010), distance from school and college fees (Carneiro, Meghir, & Parey, 2013), the incidence of religious conflict (Moreno, 2021) have been used as instruments.

[3] Borjas (1992) used 'ethnic capital' in his study of intergenerational mobility to test the assimilation hypothesis. According to this hypothesis, social, cultural, and economic differences among various ethnic groups which are a direct outcome of migration, vanish after a few generations. Ethnic capital refers to the average quality of the ethnic environment in which one generation makes its investments in the next generation. According to Borjas (1992), the skills of the next generation depend not only on the skills of their parents but also on the 'ethnic capital'. More recently, this concept found importance in studies such as Fernández and Fogli (2009); Giuliano (2007).

within India.

At the macroeconomic level, the relationship between inequality and persistence has been established by *The Great Gatsby Curve*, which finds countries with high income inequality to have low social mobility (high intergenerational persistence), while countries with low income inequality have high social mobility (low intergenerational persistence) (Corak, 2013). We also test whether the relationship holds for zonal councils within India.

3 Data

This study uses household-level data from the India Human Development Survey (IHDS), which is a panel survey conducted by the National Council for Applied Economic Research (NCAER) and the University of Maryland. The first wave of IHDS (2004-05) comprises 41,554 households, while the second wave (2011-12) includes 42,152 households of which 83% of the households were re-interviewed from 2004-05 (Desai & Vanneman, 2010, 2015).

For the purpose of this study, we use education data from the second wave. The household questionnaire collects information on the education of the head of the household, the father of the household head while the individual questionnaire collects information on the education of each member of the household, both residents and non-residents, and their children. This study uses data only for male-headed households and studies three generations of males (G0, G1, and G2) since traditionally in Indian society, women once married are considered a part of their husbands' household and are, therefore, difficult to track based on households. Using male members in three generations of which G1 and G2 are residents[4], makes it easier to compare across households since the within-household environment are similar for all three generations.

We construct our sample based on these triads with an age cut-off of 18 years for G2 as in Kundu and Sen (2022). The final sample entails 11,174 grandfather-father-son (G0-G1-G2) triads. Each household is identified based on the region to which it belongs[5], since it is much more consistent in terms of boundaries than states, for the period spanning three generations.

4 Methodology

4.1 Transmission Models

The use of distinct model specifications to ascertain the nature of persistence for each identity group forms the crux of this study. This also serves the purpose of testing the 'grandfather effect' for each region in India using two specifications - one that assumes a direct and independent effect of the grandfather's educational attainment on the grandson's education, and one that assumes an indirect effect of the grandfather's education on the grandson's education. A third model specification does away with the existence of the 'grandfather effect' which reduces it to an intergenerational model wherein the father's education alone affects the son's education.

The three models are presented in increasing order of complexity. First, is the model (Model I) with a direct effect (intergenerational model), where G2's education depends only on G1's educational attainment:

$$y_{it} = \beta_0 + \beta_1 y_{i,t-1} + \epsilon_{it} \tag{1}$$

where β_1 is the father's effect on the son's education.

To test the 'grandfather effect' for India, we use two regression specifications, wherein, to test the direct effect (Model II), we use the following specification:

[4] In the majority of the households, the male member belonging to G0 generation (grandfather) is not alive, so we impose this restriction only on G1 and G2.

[5] Refer Appendix A.2 for construction of region based on Zonal Councils.

$$y_{it} = a_1 + b_1 y_{i,t-1} + b_2 y_{i,t-2} + u_{it} \tag{2}$$

where b_1 represents intergenerational persistence or the persistence of the father's educational attainment (G1) on the son's educational attainment (G2), while b_2 is the persistence of the grandfather's education (G0) on the son's education (G2).

If the effect of the grandfather's outcome on the grandson's outcome is presumed as indirect, we use an instrumental variable specification where the grandfather's education (G0) affects the grandson's education (G2) only through the father (G1). In other words, G1 is endogenous and is therefore instrumented with G0's education. The resultant model specification (Model III) is as follows:

$$y_{it} = a_2 + b_3 y_{i,t-1} + v_{it} \tag{3}$$

where $y_{i,t-1}$ is instrumented with $y_{i,t-2}$.

To arrive at the best model specification, we first test for endogeneity and correct for it by using the grandfather's education as an instrument for the father's education. In other words, Model III is chosen as the ideal model where endogeneity exists. If the father's education is an exogenous variable, a Hausman specification test is used to determine the choice between Models I and II.

Model III is based on the assumption that the grandfather's education (G0) is a good instrument for the father's education (G1). Although we carry out tests for checking whether the instrument is good, we find that the study requires a good identification strategy that can explain variation in G1's education attainments without affecting G2's educational attainment. G0's education cannot be strictly seen as exogenous to G2's education, since there may be cases where the 'grandfather's effect' is independent and strong.

In the following section, we explore an unprecedented identification strategy that is specific to the Indian context and test for the accuracy of the proposed instrumental variable.

4.2 Identification Strategy: The Abolition of Untouchability

The practice of *untouchability*, whereby individuals of a higher caste status consider it polluting to be in close proximity with a person belonging to the lowest caste, was a social norm that enabled the oppressive occupation-based *Varna* system to be practiced. The most menial of occupations - those that primarily deal with death, blood, and excrement - were carried out by those at the bottom of the social hierarchy or the *Dalits*.[6] Untouchability, therefore had a much greater bearing on occupational choice (or the lack of it) than wage rates or labor laws. This practice also prevented access to education.

In 1955, The Untouchability (Offenses) Act abolished the practice of untouchability. Affirmative action policies in education, public employment, and political spheres ensued after abolition. The abolition is likely to have caused the eventual breakdown of the traditional caste-based occupations by enabling greater access to education and employment opportunities. The abolition of untouchability, therefore, created an external variation in a lower-caste household's decision to send their children to school. Children who were born in the 1950s or later would have had better opportunities for education than those born before 1950. The latter group's choice of getting educated by their parents would have largely been determined by their household's socio-economic background and the existing social norms guiding mobility.

To formalize this argument, assume that there are three generations G0, G1, and G2. G0 represents the grandfather's generation, G1 is the father's generation, and G2 pertains to the son's generation. We posit that the abolition of untouchability affected G0 generation's decision to send G1 to school, especially if they belonged to a former untouchable caste. Here, abolition of untouchability becomes an

[6] While *Dalits* constitute the 'Scheduled Castes' (SC) category under the Indian constitution, there is reason to believe that they weren't the only groups that were considered untouchables. The 'Scheduled Tribes' (ST) category, although outside the purview of the caste system - also known as *Avarnas*, have also been subjected to untouchability.

instrument for the educational attainment of G1 that further affects the next generation G2's education as shown in Figure 1. The variation in G1's education arises through its caste affiliation and if he belongs to a former untouchable group, the effect on education choice is larger.

$$Ed_{i1} = \alpha_0 + \alpha_1 Untouchability * Caste + \alpha_2 Untouchability + \alpha_3 Caste + \epsilon_{1it} \tag{4}$$

$$Ed_{i2} = \alpha_2 + \alpha_3 \hat{Ed}_{i1} + e_{2it} \tag{5}$$

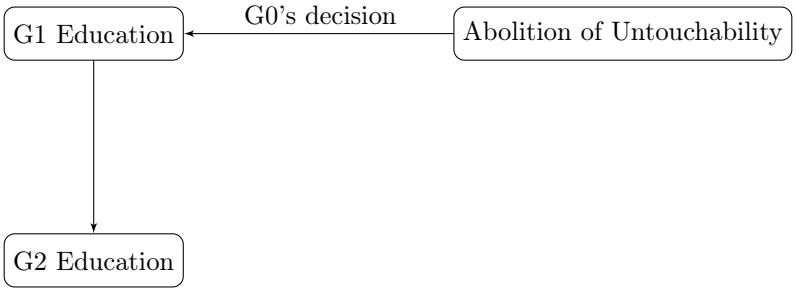

Figure 1
Identification Strategy

5 Models of Persistence

Table A2, Table A3, Table A4, and Table A1 report diagnostic tests associated with each of the models to arrive at the ideal model specification. Table A2, Table A3, and Table A4 determine those categories for which persistence must be measured using an IV specification due to endogeneity of the independent variable - *father's education*. Endogeneity is measured using two tests: the Wu-Hausman and the Durbin-Wu-Hausman tests. Where endogeneity is detected, additional identification tests such as Anderson-Rubin and Cragg-Donald Wald Tests are used to check for whether the instrumental variable - *grandfather's education* - is a good instrument for the father's education. Where there is no endogeneity, the choice is between Model I and Model II as shown in Equation (1) and Equation (2). This choice is guided by the Hausman test results as shown in Table A1.

Table 1 and Table 2 report β coefficients of persistence associated with the model that best explains intergenerational/ multigenerational education mobility for each of the caste and religious categories under each zone. We find that for Hindu and General category households, the transmission of educational persistence across generations is best captured by Model III regardless of regional influences. General category and Hindu religious groups also have the highest persistence, while the Scheduled Caste and Scheduled Tribes (SC/ST) and Christian religious groups have the lowest persistence. Among caste groups, persistence is extremely high in the Central region, while it is lowest on average, in the Northern region. Among religious groups, persistence is highest in the Central region, while it is lowest on average, in the South.

5.1 Four Emergent Scenarios

Table A5 and Table A6 report the average years of education for each of the three generations for caste and religious groups under each zone. When studied in conjunction with Table 1 and Table 2, these

tables reveal the emergence of four distinct scenarios concerning persistence, average years of education, and position in the social hierarchy[7].

1. **High persistence, favorable social position, and high education levels**

   General category comprises the *Brahmins*, *Forward Castes*, and other caste categories that do not belong to any of the caste categories that benefit from affirmative action or reservation policy. Thus, they represent that part of society that is advantaged (or at least not disadvantaged) by birth. In other words, their categorical social identity is such that they are right on top of the social hierarchy. This, combined with a higher average education that generations have benefited from[8] and a high educational persistence ensure that there is a reinforcement of advantage in the educational sphere.

2. **Low persistence, unfavorable social position, and low education levels**

   The Scheduled Castes (SC) and the Scheduled Tribes (ST) are two groups that form part of the bottom-rung of the caste hierarchy. Added to their inherent disadvantage is the low average educational attainment for all three generations across zones as observed in Table A5 and Table A6. However, their low educational persistence makes this disadvantage less vigorous, pointing towards a probable disintegration of disadvantage.

3. **High persistence, unfavorable social position, and low education levels**

   Among religious groups, Muslims' disadvantage in society cannot be attributed to their minority status alone; since there are minority religious groups that do not face stigma the way Muslims do. Their stigmatized identity, low educational levels, and high intergenerational persistence make them particularly vulnerable, which points to reinforcement or even an aggravation of their disadvantage.

4. **Low persistence, unfavorable social position, and high education levels**

   Among religious groups, Christians have the lowest educational persistence across zones and have reasonably high levels of average education except in the Northern zone. Christians are a minority everywhere except in the North-eastern states and Kerala. The lack of generalisability regarding the social status of Christians makes it rather difficult to categorically state the outcomes associated with low persistence and high education levels. It is, however, useful to note that regardless of social position, high education combined with low educational persistence could mean that the educational achievements for each generation among Christians are the outcome of individual decisions in favor of education, rather than an inheritance of educational advantage from their previous generations. This scenario is unique in identifying an identity group that has come up despite their historical disadvantage of being a minority but has consistently shown educational progress over generations.

The findings in this section reveal that persistence is indeed a cultural phenomenon that affects each identity group differently. The use of a predetermined model specification tends to assume away such differences and, therefore, compromises the study of the underlying mechanisms that influence educational mobility.

A positive and large β coefficient signifies high persistence and low educational mobility. Whether persistence is good or bad varies from one identity group to another. A group with a higher social standing is likely to benefit from high intergenerational persistence of outcomes, since that would imply the inheritance of advantages. Conversely, a group with a low social standing will only benefit from higher mobility or lower intergenerational persistence. It is also important to consider how the three generations within each religious group have fared, on average, as far as educational outcomes are concerned, so that we are better equipped to comment on their associated persistence coefficients.

Although we observe patterns in educational persistence that can be generalized for caste, religion, and regions, it must be borne in mind that each individual experiences privilege or disadvantage based on the intersectionality of all three, if not more characteristics or affiliations. This also means that multigenerational persistence mechanisms are regulated by identity affiliations - with each identity adding

[7] Of course, religious affiliation does not subscribe to a hierarchy, but it is interpreted in terms of minority and majority.

[8] Note that, among caste groups in Table A5 and Table A6 all three generations of the General category have on average, the highest educational attainment.

to a cumulative effect that culminates in degrees of advantage or disadvantage and the region to which they belong. We pursue this issue in the following subsection.

Table 1
Educational Persistence among Caste and Religious Groups in the North, North-East, and Central Zones

	North				North-East				Central			
	β	Model	$R^2/AdjR^2$	N	β	Model	$R^2/AdjR^2$	N	β	Model	$R^2/AdjR^2$	N
General	0.437***	III	0.204	1,205	0.708***	III	0.353	192	0.669***	III	0.254	504
	(0.041)				(0.095)				(0.062)			
OBC	0.428***	I	0.252	857	0.266***	I	0.036	207	0.712***	III	0.155	1,039
	(0.025)				(0.065)				(0.075)			
SC/ST	0.392***	I	0.173	815	0.380***	I	0.162	87	0.751***	III	0.087	521
	(0.030)				(0.048)				(0.133)			
Hindu	0.521***	III	0.230	2153	0.616***	III	0.223	300	0.687***	III	0.195	1761
	(0.035)				(0.085)				(0.040)			
Muslim	0.443***	I	0.196	390	0.510***	I	0.288	89	0.548***	I	0.317	294
	(0.046)				(0.086)				(0.047)			
Christian	0.580***	I	0.319	23	0.276***	II	0.163	82				
	(0.185)				(0.101)							
Other	0.400***	I	0.256	340	0.335***	I	0.399	24				
	(0.037)				(0.088)							

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Notes. The table reports the persistence associated with each caste and religious group for the North, North-East, and Central zones, respectively. Persistence is calculated using three model specifications, of which the best model specification and its associated β coefficient are reported for each of the groups. Model I is an intergenerational model, Model II is a multigenerational model which assumes that the grandfather's effect on the grandson's educational outcome is independent of the father's effect, and Model III measures long-term intergenerational mobility by instrumenting the father's educational outcome with the grandfather's education, which is based on the premise that the grandfather's educational outcome affects the grandson's outcome indirectly through the father.
The table does not report results for *Christian* and *Other* religious groups in the Central zone since the number of observations was below 20.

Table 2
Educational Persistence among Caste and Religious Groups in the Eastern, Western, and Southern Zones

	East				West				South			
	β	Model	$R^2/AdjR^2$	N	β	Model	$R^2/AdjR^2$	N	β	Model	$R^2/AdjR^2$	N
General	0.658***	III	0.413	576	0.495***	III	0.212	639	0.533***	III	0.164	293
	(0.049)				(0.055)				(0.076)			
OBC	0.792***	III	0.103	595	0.453***	I	0.265	642	0.547***	III	0.150	1,294
	(0.087)				(0.030)				(0.053)			
SC/ST	0.617***	III	0.242	564	0.395***	I	0.221	362	0.567***	III	0.085	666
	(0.074)				(0.039)				(0.087)			
Hindu	0.659***	III	0.311	1447	0.490***	III	0.265	1429	0.547***	III	0.133	1884
	(0.035)				(0.036)				(0.041)			
Muslim	0.459***	II	0.387	226	0.449***	I	0.221	109	0.422***	I	0.203	248
	(0.061)				(0.081)				(0.053)			
Christian	0.292**	I	0.256	20	0.302**	I	0.265	20	0.406***	I	0.234	117
	(0.117)				(0.119)				(0.069)			
Other	0.413***	I	0.190	42	0.399***	I	0.243	72	0.441**	I	0.273	16
	(0.135)				(0.084)				(0.192)			

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Notes. The table reports the persistence associated with each caste and religious group for the Eastern, Western, and Southern zones, respectively. Persistence is calculated using three model specifications, of which the best model specification and its associated β coefficient are reported for each of the groups. Model I is an intergenerational model, Model II is a multigenerational model which assumes that the grandfather's effect of the grandson's educational outcome is independent of the father's effect, and Model III measures long-term intergenerational mobility by instrumenting the father's educational outcome with the grandfather's education, which is based on the premise that the grandfather's educational outcome affects the grandson's outcome indirectly through the father.

5.2 Persistence, Identity, Region, and Abolition

In the previous subsection, we explored different models of persistence and observed general patterns with regard to identity for region-wise sub-samples. In this section, we explore persistence for all G0-G1-G2 triads with various model specifications with region fixed effects and interaction variables. Interactions are crucial to the study of intersectionality.

Table 3 reports the results from two specifications of regressions of G2's education on G1's education. The table compares the results of a simple OLS regression with region fixed effects, with a two-stage least square model with the abolition of untouchability as an instrument for G1's education as shown in Figure 1.

The abolition of untouchability creates a natural experiment setting where people in the G1 generation fall into one of the three categories - Control, Partial Treatment, and Treatment - depending on their age in 1955. The 'Treatment' group constitutes children aged 5 years or less at the time of abolition, while the 'Partial Treatment' group constitutes children who were already in school during abolition, and the 'Control' group comprises of G1 individuals who had completed their education by 1955. We observe in Table 3 that the educational persistence becomes stronger in magnitude with the introduction of the instrumental variable[9]. It is also observed that the treatment group benefited most from the ban on untouchability. Although the interaction between caste and untouchability is not significant, it is evident from the sign of the slope coefficient that SCs benefited greatly from the abolition, followed by OBCs. The partial treatment groups belonging to SC, ST, and OBC categories and the treatment groups from ST categories did not benefit from the abolition as far as educational attainment is concerned. In the absence of interactions, the table shows a positive and significant effect of abolition (*Untouchable*) on the educational outcomes of the father's generation (G1).

Table 4 attempts to study the joint effect of caste affiliation and father's education, religion and father's education, as well as region and father's education (G0 and G1) on the son's education (G1 and G2). Here, the term 'father' refers to both G0 and G1 and 'son' refers to G1 and G2. The two specifications in Table 4 pertain to contiguous father-son pairs - G1-G2 and G0-G1, respectively. In the latter, *Untouchable* is used as an additional variable to test the effect on G1's educational attainment.

Table 4 reveals that the G0-G1 pair has a higher educational persistence than the G1-G2 pair. The interaction variable *Region*Father's Education* is significant and positive for both pairs in the Central and Eastern regions signifying that education and region jointly affect the educational persistence of the next generation, whether it is G1 or G2. The significance of the interaction coefficients *Caste*Father's Education* and *Religion*Father's Education* also reveal the existence of a combined effect. This brings to light the importance of an intersection of factors in determining the direction and magnitude of persistence. It is, therefore, not enough to say that an educated parent will educate their child *ceteris paribus*; rather, we must pay attention to other factors that may add up to be advantageous or disadvantageous to a household such as its caste identity, religious identity, and the region where it is located.

[9] Similar to the findings of Moreno (2021).

Table 3
OLS and IV regressions of Son's education (G2) on Father's education (G1)

	(1) OLS	(2) IV
Reduced form		
Father's Education (G1)	0.440***	0.684***
	(0.007)	(0.025)
Observations	11174	11144
R-squared	0.275	0.198
First stage		
Untouchable		
Partial Treatment		1.317***
		(0.384)
Treatment		1.861***
		(0.365)
Caste		
OBC		-2.098***
		(0.541)
SC		-3.889***
		(0.679)
ST		-1.256
		(2.662)
*Untouchable*Caste*		
Partial Treatment*OBC		-0.200
		(0.579)
Partial Treatment*SC		-0.101
		(0.721)
Partial Treatment*ST		-1.598
		(2.764)
Treatment*OBC		0.176
		(0.555)
Treatment*SC		0.460
		(0.692)
Treatment*ST		-0.828
		(2.700)
R-squared		0.096
F-stat		73.41

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Notes. The table reports the results of OLS and IV regressions of the Son's education (G2) on the Father's education (G1). The IV model uses the abolition of untouchability as an instrument for the Father's education. The first stage and reduced form results are shown. There are three categories of interest as far as the abolition of untouchability is concerned: Treatment, Partial Treatment, and Control. The treatment group refers to those people in the father's generation who were 5 years, younger than 5, or not even born at the time of abolition. The Control group constitutes those who were above the age of 26 during that time. The Partial Treatment group constitutes those who had already started schooling when abolition was legalized.

Table 4
The Effect of Caste, Religion, and Region

	(1)	(2)
Father's Education	0.354***	0.503***
	(0.016)	(0.022)
*Caste*Father's Education*		
OBC	0.024	-0.002
	(0.017)	(0.025)
SC	0.023	0.105***
	(0.019)	(0.030)
ST	0.067	-0.014
	(0.067)	(0.076)
*Religion*Father's Education*		
Muslim	0.092***	0.081**
	(0.023)	(0.033)
Christian	0.013	-0.144**
	(0.051)	(0.061)
Other	-0.009	0.025
	(0.032)	(0.052)
*Region*Father's Education*		
North-East	0.025	0.031
	(0.034)	(0.045)
Central	0.058***	0.102***
	(0.021)	(0.033)
East	0.111***	0.114***
	(0.021)	(0.030)
West	0.016	0.072**
	(0.021)	(0.031)
South	-0.017	0.084***
	(0.022)	(0.032)
Untouchable		
Partial Treatment		1.261***
		(0.219)
Treatment		1.950***
		(0.209)
R-squared	0.292	0.288
Observations	11,144	11,032
F-stat	201.07	234.1

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Notes. The table shows OLS estimates for two consecutive generations. The dependent variables are Grandson's education (G2) and Father's education (G1), respectively for (1) and (2). Here, 'Father's Education' refers to the education of the father of G2 (G1), i.e; G1 (G0) depending on the specification. This table explores the effect of interactions of the education of the previous generation with caste, religion, as well as region. Specification (2) also includes an additional variable 'Untouchability' to test for the effect of abolition on G1's education.

6 The Great Gatsby Relation

The Great Gatsby curve was first used by Corak (2013) to show that countries with high income inequality had high generational earnings elasticity or persistence, while those with low income inequality had low persistence. *The Great Gatsby* relation establishes a direct relationship between inequality and generational persistence. By being the first to test this relation for areas within a country, this study aims at exploring if *The Great Gatsby* relation exists for regions within India.

Figure 2 - Figure 9 analyze the relationship between education inequality in the father's generation and the persistence of each of the caste/ religious groups in each region to determine whether the Great Gatsby relation holds for identity-region cells in India. We also test if this relation holds regardless of the model specification employed to measure persistence.

It may be observed from all the figures that regardless of the identity group chosen, *The Great Gatsby* relation holds or the curve is positively sloped. In other words, caste-region cells or religion-region cells with high educational inequality (measured by the Gini coefficient) have high generational educational persistence (β coefficients) and therefore, low generational educational mobility. This positive relation holds regardless of how the β coefficients are measured. The only exception is Figure 3 where the relation is almost non-existent but close to a positive relationship. It is also observed that the caste-region interaction produces a much steeper *Great Gatsby* relation than the religion-region interaction.

Another noteworthy pattern is that of how the General category corresponds to high persistence and relatively low educational inequality. This means that an individual who belongs to the General category is likely to have high educational attainment, that is most likely because his father is also highly educated; and since educational inequality is low among General category individuals, it can be inferred that most individuals with this caste affiliation are highly educated. There is a cluster of SCs and STs on that part of the graphs that correspond to high inequality and low generational persistence, which means that within this group, there may be a few individuals with high education, and high mobility ensures that there is ample scope for breaking out of a predominantly low education trap.

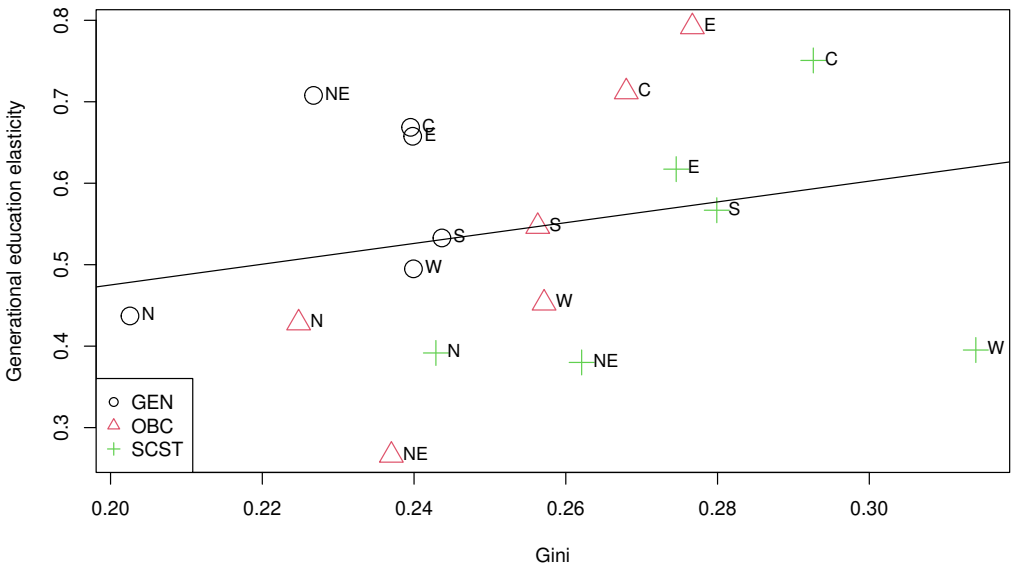

Figure 2

Great Gatsby Curve for Caste Groups based on the Best Model Specification

Notes. The figure plots the Gini coefficient for education for each caste group in each zone on the x-axis and the corresponding β coefficient using the best model specification on the y-axis.

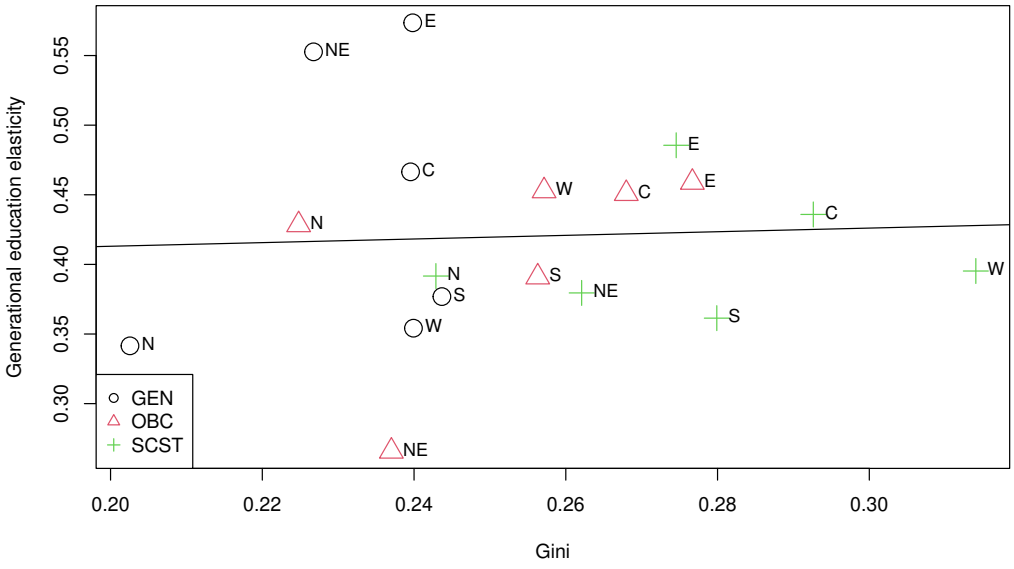

Figure 3

Great Gatsby Curve for Caste Groups based on Model I

Notes. The figure plots the Gini coefficient for education for each caste group in each zone on the x-axis and the corresponding β coefficient using Model I i.e, Equation (1) on the y-axis.

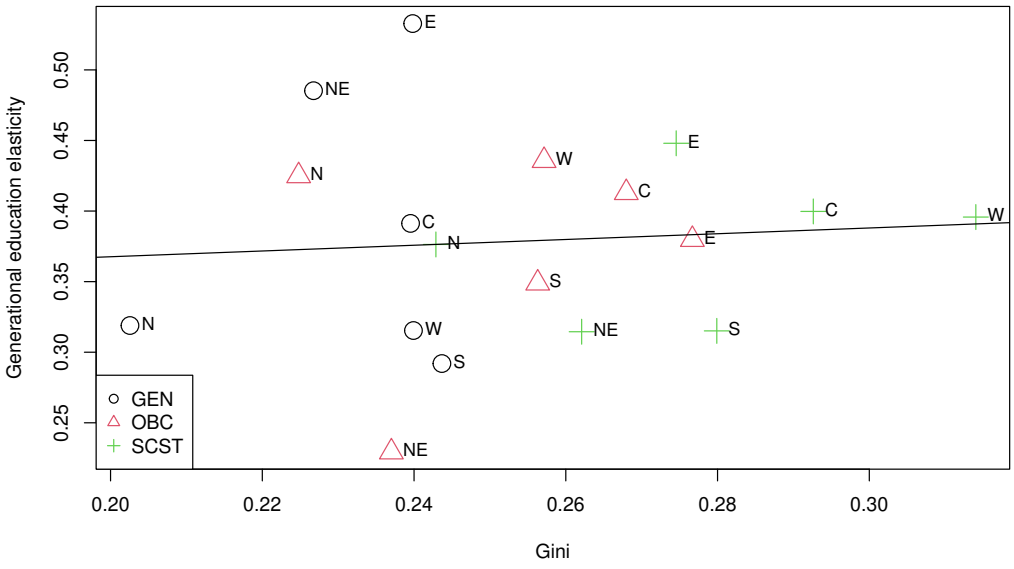

Figure 4

Great Gatsby Curve for Caste Groups based on Model II

Notes. The figure plots the Gini coefficient for education for each caste group in each zone on the x-axis and the corresponding β coefficient using Model II i.e, Equation (2) on the y-axis.

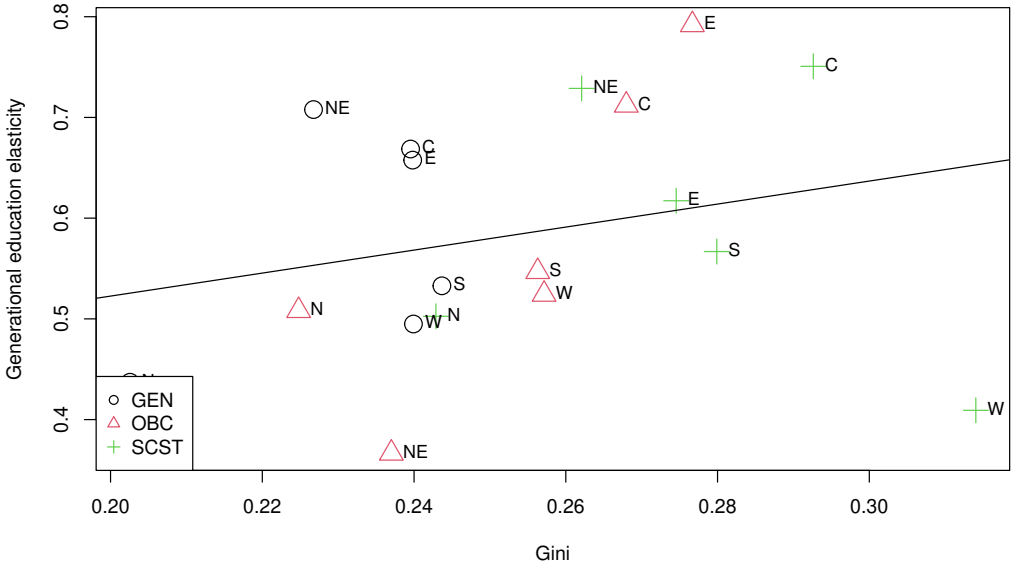

Figure 5

Great Gatsby Curve for Caste Groups based on Model III

Notes. The figure plots the Gini coefficient for education for each caste group in each zone on the x-axis and the corresponding β coefficient using Model III i.e, Equation (3) on the y-axis.

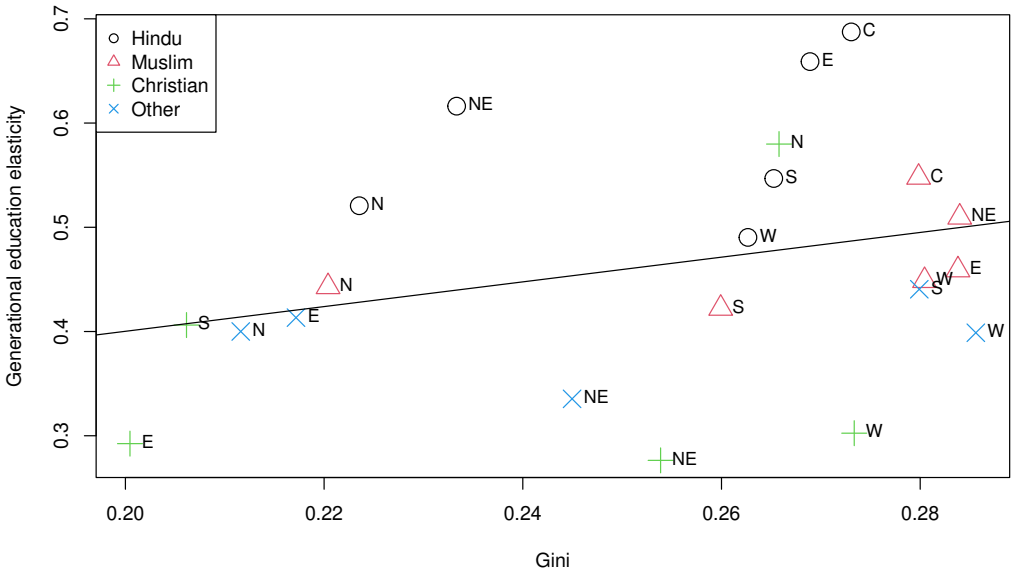

Figure 6

Great Gatsby Curve for Religious Groups based on the Best Model Specification

Notes. The figure plots the Gini coefficient for education for each religious group in each zone on the x-axis and the corresponding β coefficient using the best model specification on the y-axis.

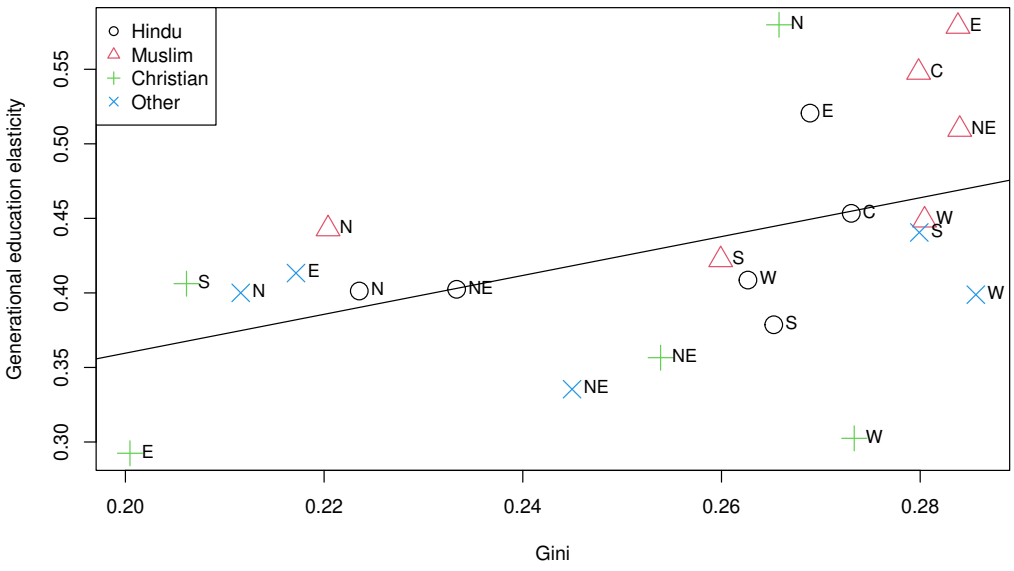

Figure 7

Great Gatsby Curve for Religious Groups based on Model I

Notes. The figure plots the Gini coefficient for education for each religious group in each zone on the x-axis and the corresponding β coefficient using Model I i.e, Equation (1) on the y-axis.

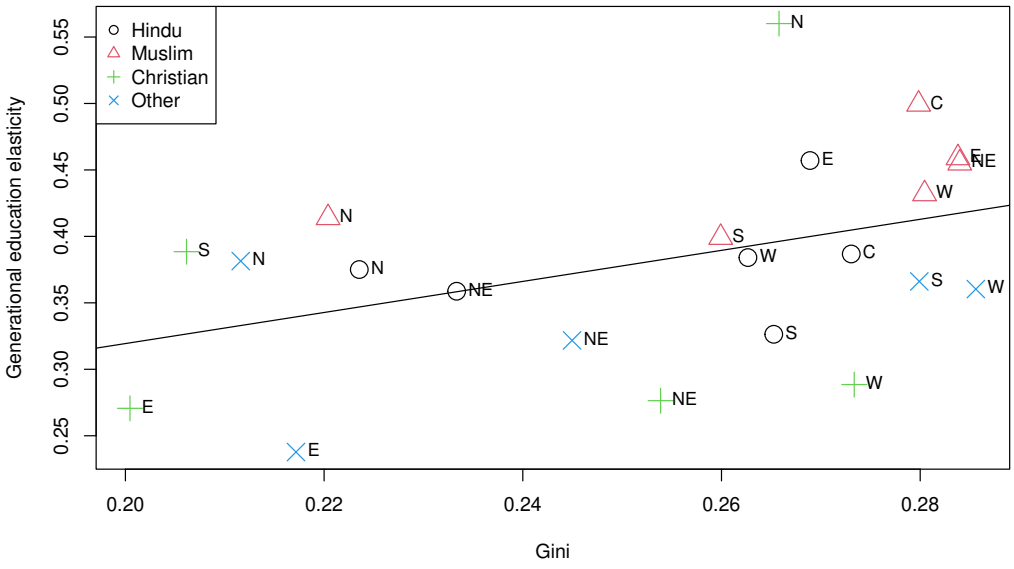

Figure 8

Great Gatsby Curve for Religious Groups based on Model II

Notes. The figure plots the Gini coefficient for education for each religious group in each zone on the x-axis and the corresponding β coefficient using Model I i.e, Equation (2) on the y-axis.

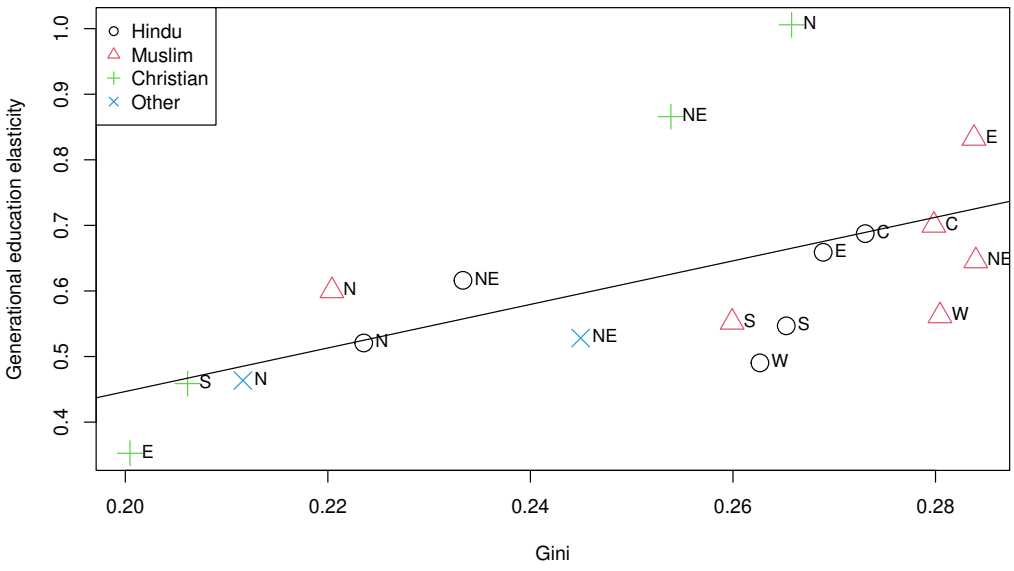

Figure 9

Great Gatsby Curve for Religious Groups based on Model III

Notes. The figure plots the Gini coefficient for education for each religious group in each zone on the x-axis and the corresponding β coefficient using Model I i.e, Equation (3) on the y-axis.

7 Discussion & Conclusion

While this study analyzes caste and religious groups separately, they represent aspects of identity that do not exist in isolation from one another. In fact the two identity groups overlap for all individuals, thereby creating an intersectionality of identities that reinforce advantages or disadvantages depending on their position in the caste hierarchy and the status of their religious affiliation (majority or minority). The study does not take for granted the fact that socially advantaged groups have better educational outcomes than those groups with an inherent disadvantage imposed by their identity. It seeks to further develop this notion in terms of how far these disadvantages or advantages manifest in the future generations, what is the average educational outcomes for each of the generations under study and what it means for their future progress (mobility) when the group has relatively high or low educational inequality. Another aspect that this study accounts for throughout, is the role of one's location in addition to their identity affiliations. It is apparent that the region in which an identity group is located also matters when it comes to not just their educational outcomes but also that of its persistence across generations.

The purpose of this study is manifold. On one hand, it seeks to study how generational persistence operates for each identity group. We find evidence for differential persistence mechanisms for different identity groups in specific zones. For some groups there exist no grandfather effect and therefore persistence is a phenomenon that lasts two generations; for others, the grandfather's educational outcome affects the grandson's outcome independently while for some other groups, this effect operates indirectly through the father.

The study is also the first to use an identification strategy for studying generation persistence in India. 'The Untouchability (Offences) Act, 1955', which marked the abolition of the practice of untouchability, is used to instrument for the father's education (G1). This novel approach also proves to be a good way to account for the external variation in G1's education. The paper also explores various interactions between identity, region, and education of the father to establish the importance of various factors in adding to the advantage or disadvantage of a individual's educational outcome through their combined effect on persistence.

Another aspect of generational persistence this study covers is the relationship between persistence and inequality of education through *The Great Gatsby* curve, and how the relationship holds regardless of the identity group considered.

The study of inequality is incomplete if it ignores the influence of one's circumstances on their outcomes, and identity is one such circumstance. Marrero and Rodríguez (2013) conceptualizes inequality as the sum of inequality of opportunity and inequality of effort. Policies must identify and mitigate the former. What makes this especially challenging is multigenerational persistence. While circumstances such as living conditions can be ameliorated, identity which follows a social hierarchy[10], is difficult to be 'compensated for'. As our findings indicate, identity, the region to which the person belongs, the inequality in their region, their father's and grandfather's outcome, all add to their outcomes. Another important consideration that this paper emphasizes is that the transmission mechanism varies from one identity group to another.

[10]Either in strict terms such as caste hierarchy, or in terms of majority and minority such as religious identity.

References

Alesina, A., Hohmann, S., Michalopoulos, S., & Papaioannou, E. (2021). Intergenerational mobility in Africa. *Econometrica*, *89*(1), 1–35.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, *106*(4), 979–1014.

Arkes, J. (2010). Using unemployment rates as instruments to estimate returns to schooling. *Southern Economic Journal*, *76*(3), 711–722.

Azam, M. (2019). Household income mobility in India, 1993-2011. *Available at SSRN 2815008*.

Azam, M., & Bhatt, V. (2015). Like father, like son? Intergenerational educational mobility in India. *Demography*, *52*(6), 1929–1959.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy*, *82*(6), 1063–1093.

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, *87*(6), 1153–1189.

Behrman, J., & Taubman, P. (1985). Intergenerational earnings mobility in the United States: some estimates and a test of Becker's intergenerational endowments model. *The Review of Economics and Statistics*, 144–151.

Bernheim, B. D., Shleifer, A., & Summers, L. H. (1986). The strategic bequest motive. *Journal of Labor Economics*, *4*(3, Part 2), S151–S182.

Black, S. E., & Devereux, P. J. (2010). Recent developments in intergenerational mobility. *NBER Working Paper 15889*.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics*, *107*(1), 123–150.

Boserup, S. H., Kopczuk, W., & Kreiner, C. T. (2013). Intergenerational wealth mobility: evidence from Danish wealth records of three generations.

Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, *11*, 123–160.

Checchi, D., Fiorio, C. V., & Leonardi, M. (2013). Intergenerational persistence of educational attainment in Italy. *Economics Letters*, *118*(1), 229–232.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, *129*(4), 1553–1623.

Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *American Economic Review*, *104*(5), 141–47.

Clark, G. (2012). What is the true rate of social mobility in Sweden? A surname analysis, 1700-2012. *Manuscript, Univ. California, Davis*.

Clark, G., & Cummins, N. (2014). Surnames and social mobility in England, 1170–2012. *Human Nature*, *25*(4), 517–537.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, *27*(3), 79–102.

Corak, M., Lindquist, M. J., & Mazumder, B. (2014). A Comparison of Upward and Downward Intergenerational Mobility in Canada, Sweden and the United States. *Labour Economics*, *30*, 185–200.

Das, M. (2007). Persistent inequality: an explanation based on limited parental altruism. *Journal of Development Economics*, *84*(1), 251–270.

Desai, S., & Vanneman, R. (2010). National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005. Inter-university Consortium for Political and Social Research [distributor]. *Ann Arbor, MI*, 06–29.

Desai, S., & Vanneman, R. (2015). *India Human Development Survey-II (IHDS-II), 2011-12*. Inter-university Consortium for Political and Social Research Ann Arbor, MI.

Dhar, D., Jain, T., & Jayachandran, S. (2019). Intergenerational transmission of gender attitudes: Evidence from India. *The Journal of Development Studies*, *55*(12), 2572–2592.

Editors. (2020). "untouchable". In *Encyclopaedia Britannica.*

Fernández, R., & Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, *1*(1), 146–77.

Ferreira, F. H., Messina, J., Rigolini, J., López-Calva, L.-F., Lugo, M. A., Vakis, R., ... others (2012). *Economic mobility and the rise of the Latin American middle class*. World Bank Publications.

Giuliano, P. (2007). Living arrangements in Western Europe: Does cultural origin matter? *Journal of the European Economic Association*, *5*(5), 927–952.

Gong, H., Leigh, A., & Meng, X. (2012). Intergenerational income mobility in urban China. *Review of Income and Wealth*, *58*(3), 481–503.

Güell, M., Rodríguez Mora, J. V., & Solon, G. (2018). New directions in measuring intergenerational mobility: introduction. *The Economic Journal*, *128*(612), F335–F339.

Heath, A., & Zhao, Y. (2021). Rethinking occupational mobility in developing countries. *Social Mobility in Developing Countries*.

Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., & Verashchagina, A. (2008). The inheritance of educational inequality: international comparisons and fifty-year trends. *The BE Journal of Economic Analysis & Policy*, *7*(2).

Iversen, V., Krishna, A., Sen, K., et al. (2016). *Rags to riches? Intergenerational occupational mobility in India* (Tech. Rep.). GDI, The University of Manchester.

Kotlikoff, L. J., & Spivak, A. (1981). The family as an incomplete annuities market. *Journal of political economy*, *89*(2), 372–391.

Kundu, A., & Sen, K. (2022). Multigenerational mobility among males in India. *Review of Income and Wealth*.

Lambert, S., Ravallion, M., & Van de Walle, D. (2014). Intergenerational mobility and interpersonal inequality in an African economy. *Journal of Development Economics*, *110*, 327–344.

Li, Y., & Zhao, Y. (2017). Double disadvantages: A study of ethnic and hukou effects on class mobility in China (1996–2014). *Social Inclusion*, *5*(1), 5–19.

Lindahl, M., Palme, M., Massih, S. S., & Sjögren, A. (2015). Long-term intergenerational persistence of human capital: an empirical analysis of four generations. *Journal of Human Resources*, *50*(1), 1–33.

Lindahl, M., Palme, M., Sandgren-Massih, S., & Sjögren, A. (2014). A test of the Becker-Tomes model of human capital transmission using microdata on four generations. *Journal of Human Capital*, *8*(1), 80–96.

Marrero, G. A., & Rodríguez, J. G. (2013). Inequality of opportunity and growth. *Journal of Development Economics*, *104*, 107–122.

Moreno, H. (2021). The influence of parental and grandparental education in the transmission of human capital. *ECINEQ WP*, *588*.

Narayan, A., Van der Weide, R., Cojocaru, A., Lakner, C., Redaelli, S., Mahler, D. G., ... Thewissen, S. (2018). *Fair progress?: Economic mobility across generations around the world*. World Bank Publications.

Neidhöfer, G., Serrano, J., & Gasparini, L. (2018). Educational inequality and intergenerational mobility in Latin America: A new database. *Journal of Development Economics*, *134*, 329–349.

Pollak, R. A. (1988). Tied transfers and paternalistic preferences. *The American Economic Review*, *78*(2), 240–244.

Solon, G. (2013). Theoretical models of inequality transmission across multiple generations. *NBER Working Paper 18790*.

Solon, G. (2014). Theoretical models of inequality transmission across multiple generations. *Research in Social Stratification and Mobility*, *35*, 13–18.

Solon, G. (2018). What do we know so far about multigenerational mobility? *The Economic Journal*, *128*(612), F340–F352.

Stuhler, J. (2013). Mobility across multiple generations: The iterated regression fallacy. *Available at SSRN 2192768*.

Wu, X., & Treiman, D. J. (2007). Inequality and equality under Chinese Socialism: The hukou system and intergenerational occupational mobility. *American Journal of Sociology*, *113*(2), 415–445.

A Appendix

A.1 Appendix A.1: Definition of Key Variables

1. *Untouchable*: Refers to a group of people belonging to a low-caste or outside the caste system. They are also called *Dalits*.(Editors, 2020)
2. *Caste*: There are three caste categories used in this paper: *General Category*, *Other Backward Classes (OBC)*, and *Scheduled Caste & Scheduled Tribe (SCST)*.
   General Category is a combination of *Brahmin* and *Forward Caste* categories. *Scheduled Caste (SC)* and *Scheduled Tribe (ST)* categories are combined for analyses in this paper as they represent the most marginalized groups in India.
3. *Religion*: The broad classifications under this variable are *Hindu*, *Muslim*, *Christian* and *Other*. The *Other* category is a combination of *Sikh*, *Buddhist*, *Jain*, *Tribal* and *Other*. There was an additional category - *None*. All households that reported *None* was removed.

A.2 Appendix A.2: Zonal Councils

S.No	Zone	States
1	North	Chandigarh, Delhi, Haryana, Himachal Pradesh, Jammu & Kashmir, Punjab, Rajasthan
2	North-East	Assam, Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Tripura, Sikkim
3	Central	Chhattisgarh, Madhya Pradesh, Uttarakhand, Uttar Pradesh
4	Eastern	Bihar, Jharkhand, Odisha, West Bengal
5	Western	Dadra & Nagar Haveli, Daman & Diu, Goa, Gujarat, Maharashtra
6	Southern	Andhra Pradesh, Karnataka, Kerala, Puducherry, Tamil Nadu, Andaman & Nicobar, Lakshadweep

Notes. The table includes only those states and union territories that are part of the IHDS data. Two union territories, *Andaman & Nicobar*, and *Lakshadweep* were excluded from the analysis.

A.3 Appendix A.3: Tables

Table A1
Hausman Test for Caste and Religious Groups in each Zone

	North	North-East	Central	East	West	South
General						
OBC	0.15	1.05			0.98	
	(0.699)	(0.305)			(0.322)	
	(I)	(I)			(I)	
SC/ST	1.9				0	
	(0.168)				(0.982)	
	(I)				(I)	
Hindu						
Muslim	1.34	0.87	3.44	13.4***	0.27	0.63
	(0.246)	(0.351)	(0.064)	(0.000)	(0.601)	(0.429)
	(I)	(I)	(I)	(II)	(I)	(I)
Christian	0.13	6.82***		0.08	0.07	0.19
	(0.716)	(0.009)		(0.778)	(0.797)	(0.662)
	(I)	(II)		(I)	(I)	(I)
Other	0.86	0.23		3.37	0.48	0.12
	(0.355)	(0.630)		(0.067)	(0.487)	(0.724)
	(I)	(I)		(I)	(I)	(I)

Notes. *** p<0.01, ** p<0.05, * p<0.1
The table reports the results for the Hausman specification test and the corresponding p-values, based on which the ideal model for explaining educational persistence is chosen. The choice is made between two models- Model I that assumes intergenerational mobility where the father's educational outcome affects the son's outcome, and Model II that presupposes multigenerational mobility wherein the father's and grandfather's education affect the son/ grandson's outcome independent of one another.

Table A2
Diagnostic Tests for North and North-Eastern Zones

	North					North-East				
	Endogeneity Tests		Endogenous	Identification Tests		Endogeneity Tests		Endogenous	Identification Tests	
	Wu-Hausman	Durbin-Wu-Hausman		Anderson	Cragg-Donald	Wu-Hausman	Durbin-Wu-Hausman		Anderson	Cragg-Donald
General	6.758***	6.737***	Yes	56.046***	78.327***	4.071***	4.049***	Yes	56.046***	78.327***
	(0.009)	(0.009)		(0.000)	(0.000)	(0.045)	(0.044)		(0.000)	(0.000)
OBC	1.152	1.154	No			0.722	0.742	No		
	(0.284)	(0.283)				(0.398)	(0.389)			
SC/ST	1.884	1.887	No			9.934***	9.612***	Yes	32.692***	38.449***
	(0.170)	(0.170)				(0.002)	(0.002)		(0.000)	(0.000)
Hindu	14.684***	14.605***	Yes	410.875***	507.307***	8.211***	8.071***	Yes	61.129***	76.261***
	(0.000)	(0.000)		(0.000)	(0.000)	(0.004)	(0.005)		(0.000)	(0.000)
Muslim	2.856	2.857	No			1.060	1.085	No		
	(0.092)	(0.091)				(0.306)	(0.298)			
Christian	0.237	0.269	No			4.190**	4.130**	No		
	(0.632)	(0.604)				(0.044)	(0.042)			
Other	0.857	0.863	No					No		
	(0.355)	(0.353)								

Notes. *** p<0.01, ** p<0.05, * p<0.1
The table reports the results of endogeneity tests for each caste and religious groups and consequently, identification tests for instrumenting the father's education level with that of the grandfather for the Northern and North-Eastern regions. Endogeneity is tested for using the *Wu-Hausman F test* and the *Durbin-Wu-Hausman chi^2 test*. Where endogeneity has been detected, identification tests such as *Anderson-Rubin Test* and *Cragg-Donald Wald Test* are used to test the validity of the instrument (grandfather's education).

Table A3
Diagnostic Tests for Central and Eastern Zones

	Central					East				
	Endogeneity Tests		Endogenous	Identification Tests		Endogeneity Tests		Endogenous	Identification Tests	
	Wu-Hausman	Durbin-Wu-Hausman		Anderson	Cragg-Donald	Wu-Hausman	Durbin-Wu-Hausman		Anderson	Cragg-Donald
General	16.474***	16.045***	Yes	136.674***	186.784***	4.450**	4.439**	Yes	195.453***	294.813***
	(0.000)	(0.000)		(0.000)	(0.000)	(0.035)	(0.035)		(0.000)	(0.000)
OBC	15.545***	15.359***	Yes	131.391***	150.122***	21.469***	20.823***	Yes	112.815***	138.742***
	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
SC/ST	6.975***	6.923***	Yes	53.654***	59.584***	4.185**	4.176**	Yes	125.084***	160.161***
	(0.009)	(0.009)		(0.000)	(0.000)	(0.041)	(0.041)		(0.000)	(0.000)
Hindu	49.807***	48.517***	Yes	390.125***	500.577***	23.982***	23.639***	Yes	458.204***	669.608***
	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
Muslim	3.352	3.348	No			12.005***	11.545***	Yes	72.399***	105.580***
	(0.068)	(0.067)				(0.001)	(0.001)		(0.000)	(0.000)
Christian										
Other						3.067	3.062	No		
						(0.088)	(0.080)			

Notes. *** p<0.01, ** p<0.05, * p<0.1
The table reports the results of endogeneity tests for each caste and religious groups and consequently, identification tests for instrumenting the father's education level with that of the grandfather for the Central and Eastern regions. Endogeneity is tested for using the *Wu-Hausman F test* and the *Durbin-Wu-Hausman chi^2 test*. Where endogeneity has been detected, identification tests such as *Anderson-Rubin Test* and *Cragg-Donald Wald Test* are used to test the validity of the instrument (grandfather's education).
The table does not report results for *Christian* and *Other* religious groups in the Central zone since the number of observations was below 30.

Table A4
Diagnostic Tests for Western and Southern Zones

	West					South				
	Endogeneity Tests		Endogenous	Identification Tests		Endogeneity Tests		Endogenous	Identification Tests	
	Wu-Hausman	Durbin-Wu-Hausman		Anderson	Cragg-Donald	Wu-Hausman	Durbin-Wu-Hausman		Anderson	Cragg-Donald
General	6.907***	6.816***	Yes	101.100***	153.310***	9.231***	9.142***	Yes	129.315***	161.616***
	(0.009)	(0.009)		(0.000)	(0.000)	(0.002)	(0.003)		(0.000)	(0.000)
OBC	11.714***	11.635***	Yes	248.936***	307.756***	1.549	1.553	No		
	(0.001)	(0.001)		(0.000)		(0.214)	(0.213)			
SC/ST	7.035***	6.992***	Yes	132.926***	165.573***	0.020	0.020	No		
	(0.008)	(0.008)		(0.000)	(0.000)	(0.887)	(0.887)			
Hindu	22.959***	22.718***	Yes	427.924***	553.097***	6.651***	6.634***	Yes	340.457***	446.315***
	(0.000)	(0.000)		(0.000)	(0.000)	(0.010)	(0.010)		(0.000)	(0.000)
Muslim	1.387	1.397	No			0.285	0.292	No		
	(0.240)	(0.237)				(0.595)	(0.589)			
Christian	0.196	0.201	No							
	(0.659)	(0.654)								
Other			No			0.386	0.401	No		
						(0.537)	(0.527)			

Notes. *** p<0.01, ** p<0.05, * p<0.1
The table reports the results of endogeneity tests tests for each caste and religious groups and consequently, identification tests for instrumenting the father's education level with that of the grandfather for the Western and Southern regions. Endogeneity is tested for using the *Wu-Hausman F test* and the *Durbin-Wu-Hausman chi^2 test*. Where endogeneity has been detected, identification tests such as *Anderson-Rubin Test* and *Cragg-Donald Wald Test* are used to test the validity of the instrument (grandfather's education).

Table A5
Average Years of Education among Caste and Religious Groups in the Northern, North-Eastern, and Central Zones

	North			North-East			Central		
	Son	Father	Grandfather	Son	Father	Grandfather	Son	Father	Grandfather
General	11.15	7.49	2.99	10.03	6.84	3.39	10.45	7.86	3.54
OBC	9.34	5.76	1.44	10.93	8.98	3.91	8.32	4.79	1.40
SC/ST	8.61	4.32	0.87	9.47	6.02	2.76	7.46	3.36	0.71
Hindu	10.10	6.39	1.94	10.64	7.71	3.73	8.89	5.35	1.76
Muslim	9.07	4.87	1.85	7.66	4.34	1.71	6.84	4.00	1.67
Christian	7.57	4.13	1.57	9.77	6.78	3.17	12.67	11.00	7.00
Other	9.68	5.76	2.05	9.96	5.17	1.50	13.25	8.25	3.25

Notes. The table shows the average years of education for each generation (son, father, and grandfather) for each caste and religious group in the Northern, North-Eastern, and Central zones, respectively.

Table A6
Average Years of Education among Caste and Religious Groups in the Eastern, Western, and Southern Zones

	Eastern			Western			Southern		
	Son	Father	Grandfather	Son	Father	Grandfather	Son	Father	Grandfather
General	10.32	7.74	4.62	11.13	7.23	3.10	11.06	7.23	3.52
OBC	8.50	4.96	2.13	9.57	5.56	2.10	10.39	5.32	2.18
SC/ST	7.64	3.82	1.45	9.38	4.52	1.61	9.40	3.43	1.34
Hindu	9.06	5.71	2.79	10.26	6.05	2.44	10.16	4.81	1.93
Muslim	7.41	4.58	2.52	8.18	4.75	1.28	9.68	5.10	2.32
Christian	9.80	6.20	3.05	11.00	6.80	3.79	11.89	7.98	4.68
Other	7.83	3.74	1.90	10.61	6.51	2.70	8.88	7.31	2.21

Notes. The table shows the average years of education for each generation (son, father, and grandfather) for each caste and religious group in the Eastern, Western, and Southern zones, respectively.