



UNITED NATIONS
UNIVERSITY
UNU-WIDER

WIDER Working Paper 2023/81

Identity and multigenerational persistence

Chinmayi Srikanth and Shubhasis Dey*

July 2023

Abstract: The study aims at bridging gaps in both theoretical and empirical literature pertaining to multigenerational persistence. From a theoretical standpoint, it argues that parental altruism is influenced by social heterogeneity rather than income-based heterogeneity. One's position in the social hierarchy, therefore, determines the motive for altruism and is in turn observed in the persistence mechanism. The study finds that for India, persistence mechanisms for education are indeed different for various identity groups (caste and religion) depending on the region to which they belong. We also study what this means for reinforcement/disintegration of (dis)advantage for these groups. By finding evidence for the *The Great Gatsby* relation, the study contends that policy-makers must target inequality rather than generational persistence. This study is also the first to develop an identification strategy that is unique to India—the abolition of untouchability in 1955, to account for endogeneity in the father's education.

Key words: persistence, multigenerational mobility, identity, education, untouchability, India

JEL classification: I24, J15, J62

Acknowledgements: We are thankful for useful comments from François Bourguignon, participants at Delhi School of Economics Winter School 2022, 17th Annual Conference on Economic Growth and Development ISI Delhi, and the Asia Meeting of the Econometric Society 2023. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

An earlier version of this paper titled *Multigenerational Persistence and the Great Gatsby Relation for India* was registered as an SSRN Working Paper and as IIMK Working Paper IIMK/WPS/544/ECO/2022/01.

Note: As the research is part of Chinmayi Srikanth's PhD thesis, the authors will hold copyright to facilitate publication of the thesis.

* Indian Institute of Management Kozhikode (IIMK), Kunnammangalam, Kerala, India; corresponding author srikanthc12fpm@iimk.ac.in

This study is published within the UNU-WIDER project Academic excellence.

Copyright © The Authors 2023

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-389-5

<https://doi.org/10.35188/UNU-WIDER/2023/389-5>

Typescript prepared by Siméon Rapin.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland and Sweden, as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

1 Introduction

Inequality and its potential causes have puzzled economists for the longest time. Inequality is a serious concern because it is persistent across generations. Generational persistence or social mobility has been widely researched both theoretically as well as empirically. However, the theoretical and empirical literature exists independently of one another. This study aims at bridging this gap by making identity affiliations the entry point, especially in a hierarchical social structure.

Extant theoretical literature on intergenerational mobility suggests that inequality arises from heterogeneity in income following which parental altruism towards their children's investment in human capital also varies. Through this paper, we contend that altruism is primarily driven by social heterogeneity which has a direct bearing on the differences in motives of different parents. By making social identity differences salient, we argue that parents' need to invest comes from the need to escape the stranglehold of their caste and/ or religious identity or to maintain them depending on the inherent advantage or disadvantage of their group in society.

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 the economy and consequently jobs, the concept of caste and its inherent social hierarchy continues to perpetuate in both occupational and educational spheres. Caste and religious identities constitute categorical social identities which 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. If parental altruism is indeed as heterogeneous as their identity affiliations in a hierarchical society, then the channels or mechanisms through which one generation's advantage or disadvantage is inherited by the next generation should also be different. The only study that explores multigenerational persistence in India finds evidence for a 'grandfather effect', where the effect of the grandfather's outcome on the grandson's outcome is independent of the father's educational and occupational outcome (Kundu and Sen 2023). The study, however, makes an a priori assumption about the channel through which persistence is carried forward to future generations. Whether the grandfather affects the grandson directly or indirectly is a proposition that itself needs to be tested in order to assess the persistence of categorical affiliations over generations. The empirical contribution of this paper is, therefore, linked to its theoretical contribution.

Another aspect of multigenerational persistence that is often overlooked is the influence of external shocks that are likely to curb or are at least aimed at curbing the hierarchical structures embedded in society. By studying the effect of the ban on untouchability, a law that has direct consequences for the rigidities of the caste system on the education of the father's generation, we not only account for external shocks but also establish a novel instrument to account for the endogeneity that exists in the father's education when studying the effect on the son's education.

Finally, this study also offers insights into how interventions can effectively reduce the persistence of disadvantage by targeting inequality in education rather than targeting persistence itself. Although we do establish that the law on the abolition of untouchability has had a major influence on the father's generation's educational attainment, the process of changing social structures is gradual and slow. 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). We test for whether *The Great Gatsby* relation holds within regions in India and find that there is a positive correlation between inequality in educational attainment and educational persistence or generational education elasticity of identity groups segregated on the basis of caste, religion, and the region to which they belong. This finding establishes that inherent inequality can be attenuated by targeting inequality rather than persistence.

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 by bearing in mind the theoretical framework of parental altruism as motivated by the need to maintain their social advantage or breakaway from their inherent social disadvantage through their aspirations for their children as measured by the educational attainment.

The study uses data from the India Human Development Survey (IHDS) 2011–12. IHDS is a rich source of data on Indian households on socioeconomic 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).

This paper is organized as follows: Section 2 offers a brief review of literature on intergenerational and multigenerational mobility, Section 3 explains the theoretical contribution of the paper which provides a framework to analyse the empirical literature, and Section 4 discusses the data source and construction of G0–G1–G2 triads. In Section 5, the models for educational transmission are discussed; Section 6 discusses persistence for each identity group based on the methods proposed in the previous section, while Section 7 tests *The Great Gatsby* for regions within India. Section 8 entails a discussion based on the findings in this paper.

2 Review of literature

Becker and Tomes (1979) are 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, affects 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 and 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 et al. 1986; Kotlikoff and 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). Despite their inter-relatedness, the theoretical and empirical literature on multigenerational mobility has existed independently of one another.

On the empirical side, although much of the methodological advances in social mobility literature have used earnings and wealth as the key variables,¹ a large part of the extant literature studies occupational (Wu and Treiman 2007; Corak et al. 2014; Li and Zhao 2017; Heath and Zhao 2021) and educational mobility (Hertz et al. 2008; Ferreira et al. 2012; Narayan et al. 2018; Alesina et al. 2021) because of the lack of reliable data on earnings (Iversen et al. 2016).

Mobility literature has largely been confined to within-country analyses, i.e the USA (Chetty, Hendren, Kline, and Saez 2014; Chetty, Hendren, Kline, Saez, and Turner 2014), Italy (Checchi et al. 2013), India (Azam and Bhatt 2015; Azam 2019; Kundu and Sen 2023), China (Gong et al. 2012), and Senegal (Lambert et al. 2014), to name a few. Neidhöfer et al. (2018) is a cross-country study of Latin America.

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

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 et al.’s (2019) unconventional study uses intergenerational transmission to study gender attitudes in India. Hnatkovska et al. (2012) find evidence for a convergence of wages and consumption between *Scheduled Castes* (SC) / *Scheduled Tribes* (ST) and non-SC/ST identity groups using National Sample Survey (NSS) data for the period spanning 1983–2005.

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 et al. 2015). Multigenerational mobility is, therefore, more desirable, where data are available. The underlying mechanism of generational persistence is also influenced by the model used (Clark 2012; Solon 2014; Stuhler 2012; Clark and Cummins 2014). 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). Studying this phenomenon is particularly challenging given the paucity or even absence of data at a multigenerational level (Güell et al. 2018), especially in the Indian context. Kundu and Sen (2023) make use of data for three generations. It is the only study on multigenerational persistence for India and it establishes the existence of a ‘grandfather effect’ for India. Multigenerational mobility literature has established specifications that either assume 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 common in the literature (Boserup et al. 2013; Lindahl et al. 2014), this instrumentation strategy is not without its deficiencies. It is not theoretically impossible to argue that the unobserved factors, which explained the outcomes of the grandfather and the father may continue to be present for the grandson. As a result, the strategy of using the grandfather’s outcome as an exogenous variable to correct for the endogeneity of the father’s outcome, as is commonly used in the literature, may be problematic. Moreno (2021) tries to overcome this flaw in the instrumentation strategy by using the incidence of religious conflict (Moreno 2021) as a proxy to instrument out the father’s education in explaining the grandson’s education. In this case, the religious conflict acts as a truly exogenous shock that affects the father’s educational attainment, which is unique to that generation. Our study exploits an exogenous shock to the existing social hierarchy in Indian society—the abolition of untouchability in 1955—to account for the potential problem of instrumentation arising due to the existence of unobserved variables across multiple generations.

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).² We also test whether the relationship holds for Zonal Councils within India.

² See Sakri et al. (2023) who study the relationship in a developing country context, i.e. Indonesia.

3 Theoretical contribution

Becker and Tomes (1979) distinguish between intergenerational mobility in open and closed societies. The degree of inheritability and propensity to invest is high in closed societies where parents have some ‘control’ over their children. Each household in a closed society aims at some form of standardization wherein most members have similar characteristics because the family has a reputation to maintain. Suppose all closed societies had cultural diversity to the extent that the inequalities were horizontal. In that case, all groups take pride in their group or community membership and have a ‘reputation’ to maintain. However, in a closed society like that of India, society is divided into groups marked by a clear hierarchy from ‘superior’ or ‘high caste’ to the ‘untouchable’ castes. For those higher up the caste hierarchy, there is indeed a social reputation to maintain. But for those at the bottom of the hierarchy, this reputation maintenance primarily serves the purpose of distinguishing themselves from those groups below them in the social ladder. In the larger scheme of things, stigmatized groups or those lower down the social hierarchy do not have a reputation to maintain as far as their social identities are concerned. In fact, they are more likely to dissociate themselves from their stigmatized identities (Davis 2015; Srikanth and Dey 2023).

Society in India is divided by a hierarchical order of caste called the *Varna* system comprising the *Brahmins*, *Kshatriyas*, *Vaisyas*, and *Shudras*. The ‘untouchable’ castes form the bottom rung of the hierarchy and are called the *Avarnas* or those who are not part of the occupation-based *Varna* system. When economies grew in complexity, a more nuanced and complex system of *Jati* came into being while still maintaining the hierarchy established by the *Varna* order. While caste hierarchy is universal in the sense that it does not change based on social setting, religious identity, and even gender identity to some extent can vary based on the sociocultural context in which we examine these. Religious groups do not fit into a hierarchy, but a lot can be said about the stigma they face based on the minority/majority status of a religious group which depends on the region or cultural context studied. Muslims in India are the second highest (after Hindus) in terms of population proportion but the most socioeconomically marginalized (Sachar et al. 2006). Other religious minorities such as Sikhs, Jains, and Buddhists are part of *Dharmic* religions much like Hinduism itself, and therefore, are seen as similar to Hindus. They also benefit from the same affirmative action privileges as the Hindus. Gender roles, although pervasive, can be surprisingly distinct in different regions. Thus, contrary to Becker and Tomes’s (1979) stance, a closed society cannot have a one-size-fits-all proposition about reputation maintenance as the main driving force of parental altruism. Altruism as a form of ‘giving’ to one’s children may be culturally driven, as opposed to solely maintaining a desired standard of living; but Becker and Tomes’s (1979) discussion of the role of one’s identity in such altruistic pursuits features in the form of an *ad hoc* relaxation of the assumption of symmetrically distributed income in all generations or simply that all families in society are identical. Macroeconomists have sought to explain income inequality by relaxing this very assumption of identical income distribution, in the presence of which there would be no inequality. Moav (2002) explains that inequality is an outcome of a convex bequest function but does little to explain why it is so. Das (2007), on the other hand, not only demonstrates that low-income parents spend less on their children’s human capital investment but also contends that they are less altruistic than their high-income counterparts.

In Das’s (2007) conceptual framework, a poor person belonging to a ‘higher caste’ group and an equally poor person belonging to a ‘lower caste’ group would ideally spend equally less on their child’s education, since their ability and willingness to pay is low and only influenced by their income. Heterogeneity in income and wealth has been explored in depth to explain inequality, but the extant literature remains silent or even assumes away the inherent heterogeneity in society. This is to say that a poor person belonging to a ‘lower caste’ will act differently compared to a poor person belonging to a ‘higher caste’.

As a start, it is necessary to think about the reasons why parents would want to invest in their children's human capital. Altruism surely drives such decisions but they should not be narrowly defined as driven by expected pecuniary gains from one's children in the future. Rather, this altruism comes from the need to help the next generation escape the stranglehold of the social structures that are likely to tie them down. And if this is indeed the case, altruism would be higher among those belonging to the bottom of the social ladder since it enables them to escape their otherwise inevitable social plight. Greater social mobility ensues for these groups which, in principle, runs counter to Becker and Tomes's (1979) hypothesis that families' altruism is an attempt to maintain the family reputation which leads the next generation to become similar to the current generation.

This study is motivated by the need to establish that persistence (mobility) can vary on several grounds, especially in a 'closed' society like that of India. Such closed societies are made up of heterogeneous groups that have different aspirations for the next generations. Serneels and Dercon (2021) show that parental aspirations are an important channel for intergenerational education mobility, especially in the case of 'backward castes'.

In this paper, we argue that parental aspirations materialize in their high or low investment in human capital which is captured by the educational attainment of their children. The slope coefficients for intergenerational persistence encapsulate the stickiness (advantage or disadvantage) of one's identity, the subsequent access (or lack of it) to educational opportunities, and to a large extent the need to maintain or escape the advantages or disadvantages that come with one's identity affiliation. As far as this conceptualization is concerned, 'parental altruism' is not narrowly defined in pecuniary terms, but is a direct consequence of the need to maintain or escape one's inherent advantage or disadvantage. Certainly, this altruism is limited by their economic means as shown by Das (2007), but we posit that categorical identity affiliations have a larger role to play in their willingness to spend on human capital investments in the children. Thus, we contest Das's (2007) argument—poor parents are less altruistic than rich parents because their ability and willingness to spend on their children's human capital are low—by accounting for differences in identity and regional influences in our empirical model. Altruism is, therefore, a function of the desperation to retain or escape one's social advantage or disadvantage, more than the ability to afford human capital investment. This study seeks to shift the discussion of parental altruism from the ability to pay to the willingness to pay by further reinterpreting this willingness to be socioeconomically driven rather than just economically driven.

4 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% were re-interviewed from 2004–05 (Desai and 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,³ makes it easier to compare across households since the within-household environment is 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 (2023). 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,⁴ since it is much more consistent in terms of boundaries than states, for the period spanning three generations.

5 Methodology

5.1 Transmission models

The use of distinct model specifications to ascertain the nature of persistence for each identity group forms the crux of this study. The underlying intuition is that parental altruism as captured by the persistence or β is motivated by the aspirations they have for their children which is directly linked to their inherited social identity. When motivations are distinct, the underlying mechanisms of persistence will also be different. We study these mechanisms by testing for three model specifications and identifying the mechanism for all identity groups in each region. Here, each identity group belonging to a particular region is studied as a separate sub-sample of the data. The specifications include one that assumes a direct and independent effect of the grandfather’s educational attainment on the grandson’s education, another that assumes an indirect effect of the grandfather’s education on the grandson’s education, and a third model specification that 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:

$$y_{it} = \beta_0 + \beta_1 y_{i,t-1} + u_{it} \quad (1)$$

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:

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

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).

³ 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.

⁴ Refer to Appendix Section B for the construction of regions based on Zonal Councils.

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:

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

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 validity of the proposed instrumental variable.

5.2 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 practised. Under the *Varna* system, the *Brahmins* were on the top of the social hierarchy, followed by the *Kshatriyas* and *Vaisyas*. The *Shudras* and *Avarnas* were at the bottom of the social order and were subjected to the practice of untouchability. According to the Indian legal system, 'General' category primarily consists of *Brahmins*, *Kshatriyas*, and *Vaisyas*, whereas the *Shudras* and *Avarnas* are distributed among three categories, namely, the *Other Backward Classes* (OBC), *Scheduled Castes* (SC), and *Scheduled Tribes* (ST). The prevalent practice of untouchability before its abolition, therefore, naturally prevented access to public goods such as education for the identity groups belonging to the *OBC*, *SC*, and *ST* categories.

In 1955, The Untouchability (Offences) 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 set in motion 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 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. We argue that the effect of the abolition of the practice would have been felt across all social identity groups. However, there will be inherent differences in the motives guiding various identity groups in such a choice. For the dominant social groups (General category), the motive comes from the increased competition that the law would enable in both occupational and educational spheres, whereas for those groups at the bottom of the social hierarchy (OBC, SC, ST), the motive comes from the need for increased social mobility. Thus, for those groups at the bottom of the hierarchy, this law is

an enabler of parental altruism whereas for the dominant groups, this ‘parental altruism’ will primarily reflect in the change in market opportunities in favour of the ‘lower castes’.

To formalize this argument, we 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. Here, the abolition of untouchability becomes an instrument for the educational attainment of G1 that further affects the next generation G2’s education, as shown in Figure 1.

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

$$Ed_{i2} = \alpha_2 + \alpha_3 \hat{Ed}_{i1} + \epsilon_{2it} \quad (5)$$

Figure 1 Identification strategy



Source: authors' elaboration.

6 Models of persistence

Table C1, Table C2, Table C3, and Table C4 report diagnostic tests associated with each of the models to arrive at the ideal model specification. Table C2, Table C3, and Table C4 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 C1.

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.

6.1 Four emergent scenarios

Table C5 and Table C6 report the average years of education for each of the three generations for caste and religious groups under each zone, and Table C7 reports the average equivalent incomes based on identity groups and zones for the whole IHDS sample. 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, average income, and position in the social hierarchy.⁵

1. High persistence, favourable social position, high education, and high income

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 from⁶ and a high educational persistence, ensures that there is a *reinforcement of advantage* in the educational sphere. The high average income of the household can be attributed to the father's generation ensuring greater investment in the child's education and therefore, resulting in high educational attainment of the son. This would be a case of 'reputation maintenance' where the family tries to standardize its characteristics—of which caste advantage is a major factor—across generations.

2. Low persistence, unfavourable social position, low education, and low income

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 C5 and Table C6. However, their low educational persistence makes this disadvantage less vigorous, pointing towards a probable *disintegration of disadvantage*. Even though their means are limited, the lower generational elasticity of education indicates that parents are altruistic enough to ensure that the following generation does not face the same magnitude of disadvantage that they faced. In this case, the motive of breaking away from the caste barriers is encapsulated in the low generational elasticity of education despite low education and income.⁷

3. High persistence, unfavourable social position, low education, and low income

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 an aggravation of or *reinforcement of disadvantage*.

⁵ Of course, religious affiliation does not subscribe to a hierarchy, but it is interpreted in terms of minority and majority where social stigma also plays a role.

⁶ Note that, among caste groups in Table C5 and Table C6, all three generations of the General category have, on average, the highest educational attainment.

⁷ Refer to Hnatkovska et al. (2012) for further empirical support on this scenario.

The channel through which this is likely to operate is that their categorical identity places them in a disadvantageous position that translates to low education and low income. But for the following generation, the high persistence can be largely attributed to their ‘devalued’ identity and partly due to lower investment in their education by their parents due to very low average income. The *reinforcement of disadvantage* that we observe can be attributed to the inadequate application of affirmative action policies for Indian Muslims.⁸ Muslim parents would find it irrational to spend disproportionately high amounts of resources on their children’s education because of the persistent stigma that Muslim persons face in the occupational domain despite being adequately qualified. They would rather prefer to educate their children in the skills and tricks of the trade that they themselves have been trained in by their parents. Hence we argue that Muslim parents have been able to rationalize how the high educational attainment of their children will not help them escape their stigmatized identity.

4. Low persistence, unfavourable social position, high education, and high income

Among religious groups, Christians have the lowest educational persistence across zones and have reasonably high levels of average education except in the Northern zone. This points to a clear *disintegration of disadvantage*. This scenario uniquely identifies an identity group that has come up despite their historical disadvantage of being a minority but has consistently shown educational progress over generations. It is also worth noting that their income levels are highest among all religious groups. Low generational elasticity of education would mean on one hand that the disadvantages are not very persistent across generations which makes it easier for them to escape their inherent social disadvantages, and on the other hand, higher education across generations can be directly attributed to ‘parental altruism’ in favour of human capital investment due to their high income. Here, the ability and willingness to pay are high and the low persistence makes it easier for them to escape their unfavourable position.

The findings in this section reveal that persistence is a cultural phenomenon affecting 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 essential to consider how the three generations within each religious group have fared, on average, regarding educational outcomes, 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 multi-generational persistence mechanisms are regulated by identity affiliations—with each identity adding 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.

⁸ Alam (2010) explains how the lack of affirmative action policies in educational and occupational domains may play a role in the social exclusion of Muslims in India.

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

Note: standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 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*** (0.049)	III	0.413	576	0.495*** (0.055)	III	0.212	639	0.533*** (0.076)	III	0.164	293
OBC	0.792*** (0.087)	III	0.103	595	0.453*** (0.030)	I	0.265	642	0.547*** (0.053)	III	0.150	1,294
SC/ST	0.617*** (0.074)	III	0.242	564	0.395*** (0.039)	I	0.221	362	0.567*** (0.087)	III	0.085	666
Hindu	0.659*** (0.035)	III	0.311	1447	0.490*** (0.036)	III	0.265	1429	0.547*** (0.041)	III	0.133	1884
Muslim	0.459*** (0.061)	II	0.387	226	0.449*** (0.081)	I	0.221	109	0.422*** (0.053)	I	0.203	248
Christian	0.292** (0.117)	I	0.256	20	0.302** (0.119)	I	0.265	20	0.406*** (0.069)	I	0.234	117
Other	0.413*** (0.135)	I	0.190	42	0.399*** (0.084)	I	0.243	72	0.441** (0.192)	I	0.273	16

Note: standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 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.

6.2 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 two two-stage least square models with region fixed effects, one with grandfather's education as the instrument and the other 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 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 abolition of untouchability as the instrumental variable⁹ which increases the magnitude of educational persistence far more when compared to the use of *grandfather's education* as an instrument. 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 IV regressions of son's education (G2) on father's education (G1)

	(1) IV Grandfather	(2) IV Untouchability
Reduced form		
Father's Education (G1)	0.591*** (0.015)	0.684*** (0.025)
Observations	11,061	11,144
R-squared	0.247	0.198
First stage		
Grandfather's Education	0.660*** (0.011)	
<i>Untouchable</i>		
Partial Treatment		1.317*** (0.384)
Treatment		1.861*** (0.365)
<i>Caste</i>		
OBC		-2.098*** (0.541)
SC		-3.889*** (0.679)
ST		-1.256 (2.662)
<i>Untouchable*Caste</i>		
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	0.243	0.096
Wald chi-square	1905.82 (0.000)	1048.21 (0.000)
<i>Endogeneity Tests</i>		
Durbin chi-square	141.201 (0.000)	110.992 (0.000)
Wu-Hausman	142.923 (0.000)	112.029 (0.000)

Note: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The OLS regression is not reported in the table as the Wu-Hausman & Durbin chi-square endogeneity tests find the father's education to be endogenous. The table reports the results of two IV regressions of son's education (G2) on the father's education (G1). The first IV model uses grandfather's education as the instrument for the father's education, while the second 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*** (0.016)	0.503*** (0.022)
<i>Caste*Father's Education</i>		
OBC	0.024 (0.017)	-0.002 (0.025)
SC	0.023 (0.019)	0.105*** (0.030)
ST	0.067 (0.067)	-0.014 (0.076)
<i>Religion*Father's Education</i>		
Muslim	0.092*** (0.023)	0.081** (0.033)
Christian	0.013 (0.051)	-0.144** (0.061)
Other	-0.009 (0.032)	0.025 (0.052)
<i>Region*Father's Education</i>		
North-East	0.025 (0.034)	0.031 (0.045)
Central	0.058*** (0.021)	0.102*** (0.033)
East	0.111*** (0.021)	0.114*** (0.030)
West	0.016 (0.021)	0.072** (0.031)
South	-0.017 (0.022)	0.084*** (0.032)
<i>Untouchable</i>		
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

Note: robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. 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), which is 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.

7 The Great Gatsby relation

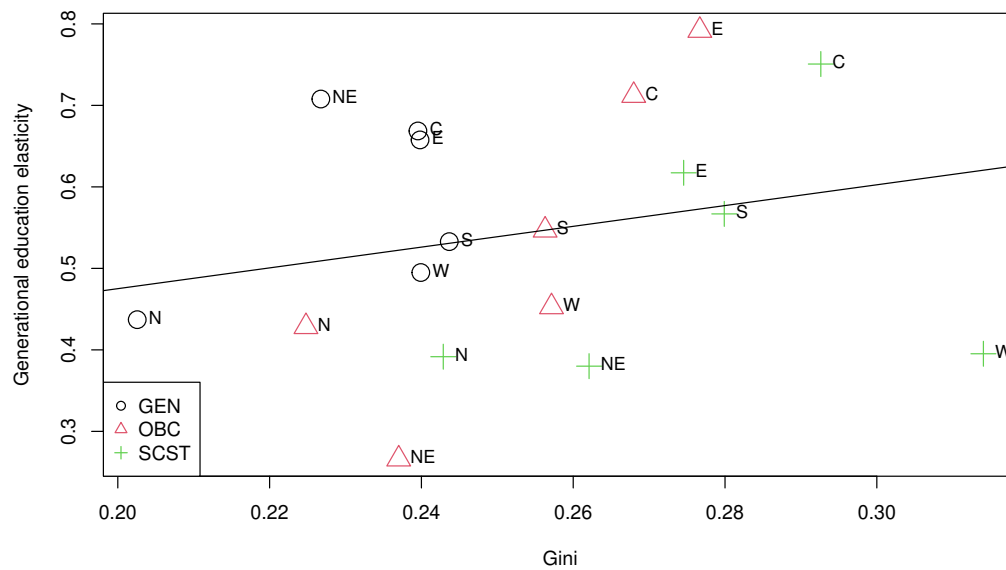
The Great Gatsby curve was first used by Corak (2013) to show that countries with high income inequality had high generational earnings elasticity or persistence, while those with low income inequality had low persistence. *The Great Gatsby* relation establishes a direct relationship between inequality 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 analyse 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.

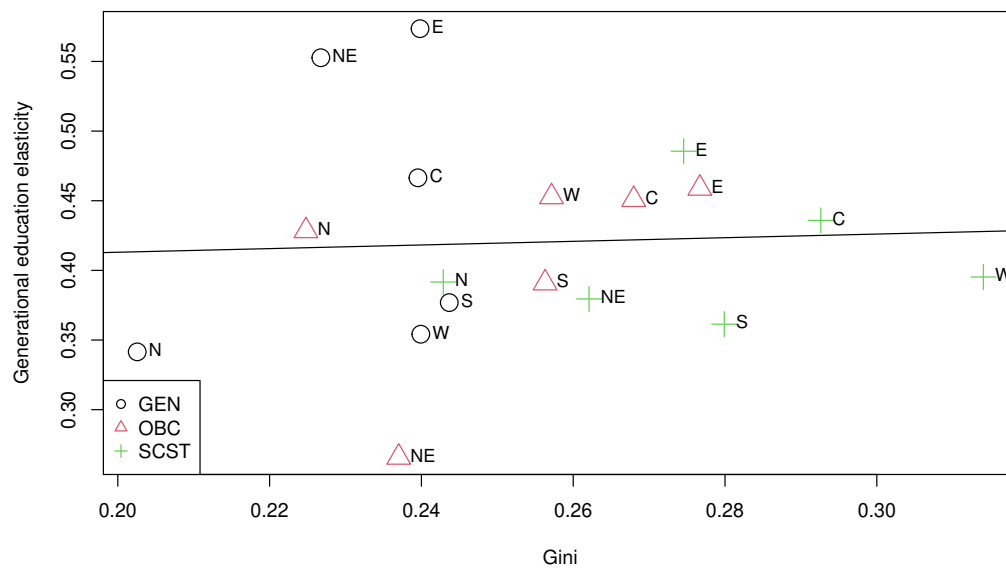
Another noteworthy pattern is that of how the General category corresponds to high persistence and relatively low educational inequality. This means that an individual who belongs to the General category is likely to have high educational attainment, that is most likely because his father is also highly 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 corresponds 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



Note: 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



Note: 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.

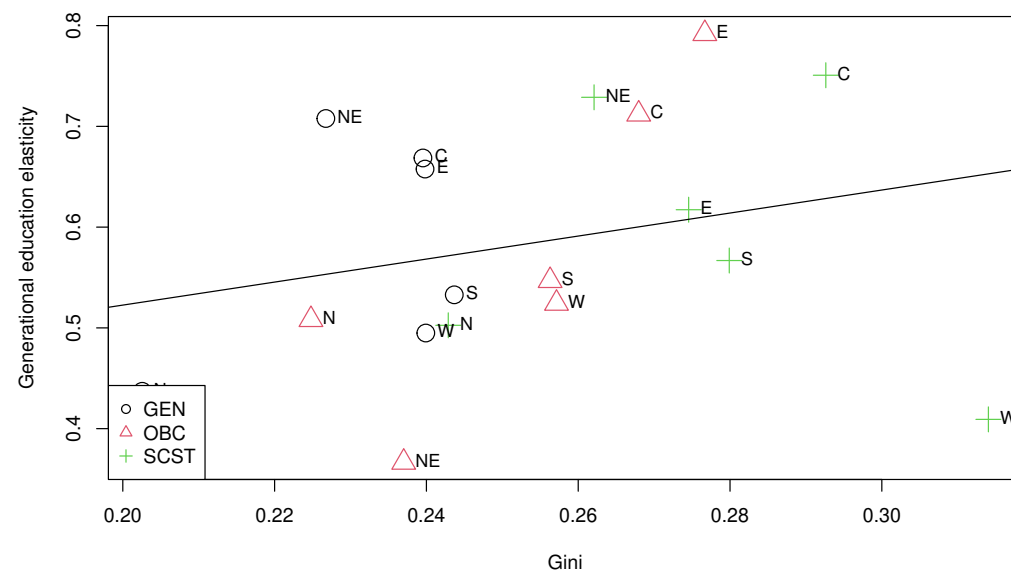
Generational education elasticity

Gini

○ GEN
△ OBC
+ SCST

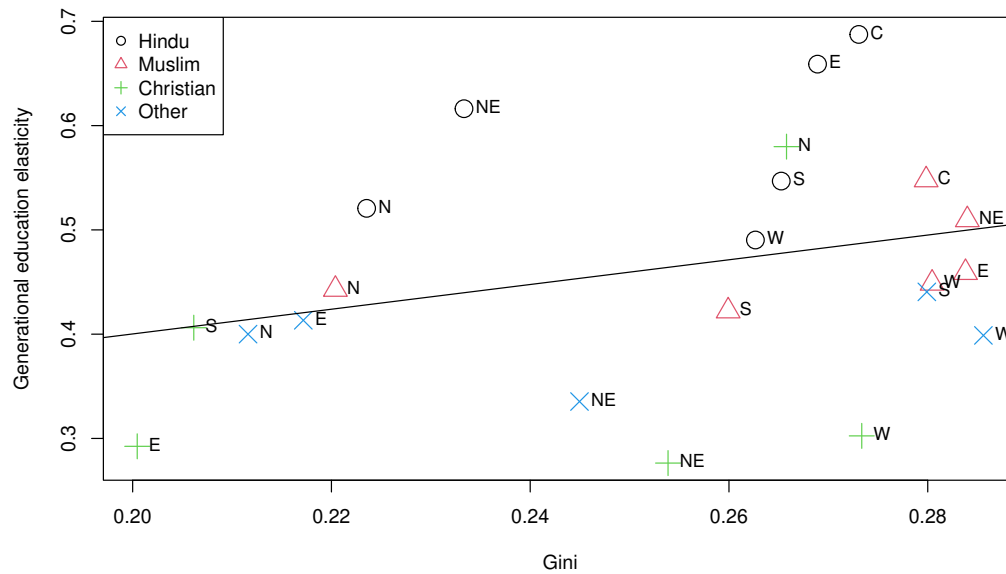
Country	Label	Gini	Elasticity
GEN	ON	0.205	0.32
GEN	OC	0.240	0.39
GEN	OE	0.240	0.52
GEN	OW	0.245	0.31
GEN	OS	0.245	0.29
GEN	ONE	0.225	0.49
OBC	N	0.225	0.43
OBC	W	0.255	0.44
OBC	C	0.270	0.41
OBC	E	0.275	0.38
OBC	S	0.255	0.35
OBC	NE	0.235	0.23
SCST	E	0.275	0.45
SCST	C	0.290	0.40
SCST	W	0.310	0.39
SCST	NE	0.265	0.31
SCST	S	0.280	0.31
SCST	N	0.245	0.37

Figure 5 Great Gatsby curve for caste groups based on Model III



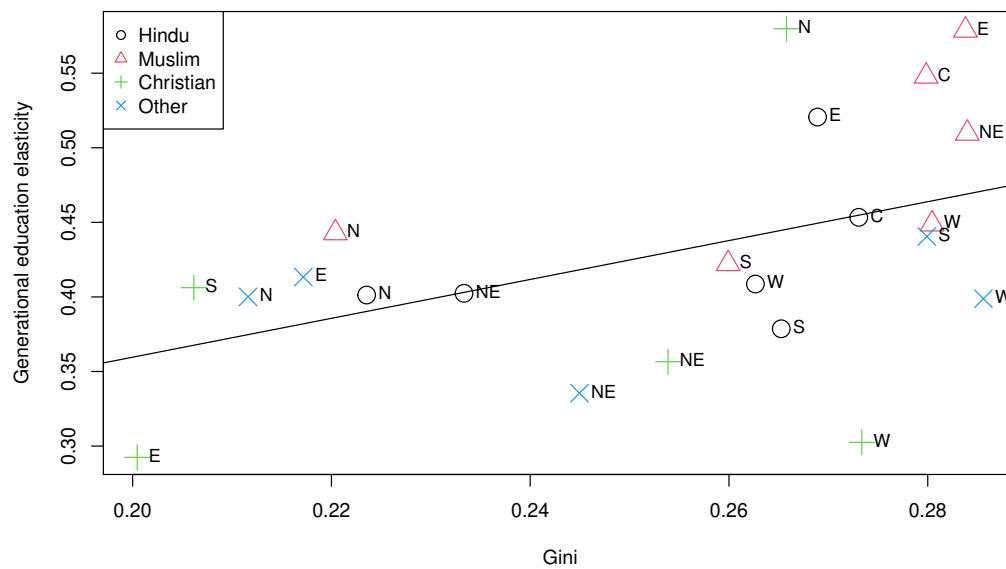
18

Figure 6 Great Gatsby curve for religious groups based on the best model specification



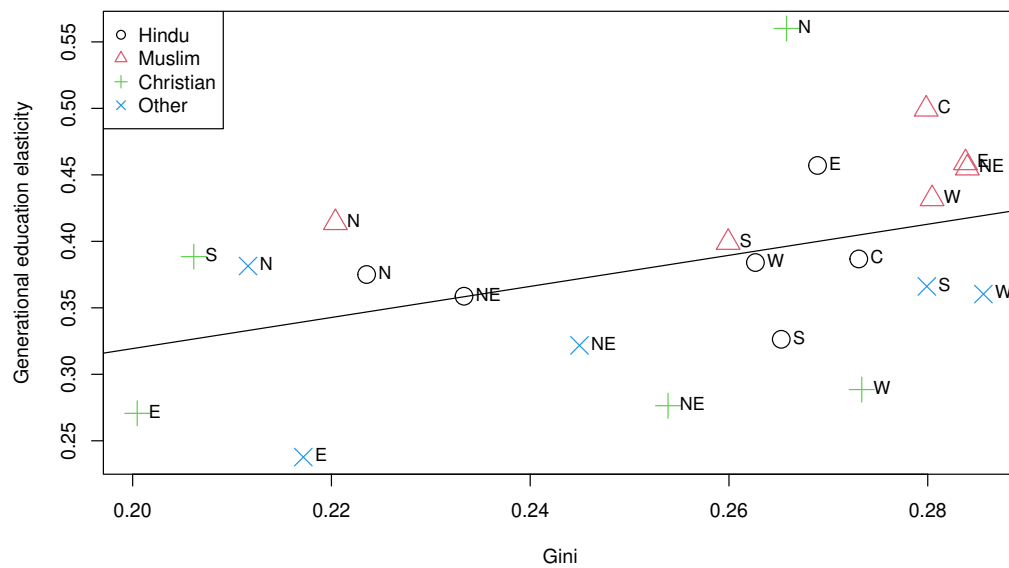
Note: the figure plots the Gini coefficient for education for each religious group in each zone on the x-axis and the corresponding β coefficient using the best model specification on the y-axis.

Figure 7 Great Gatsby curve for religious groups based on Model I



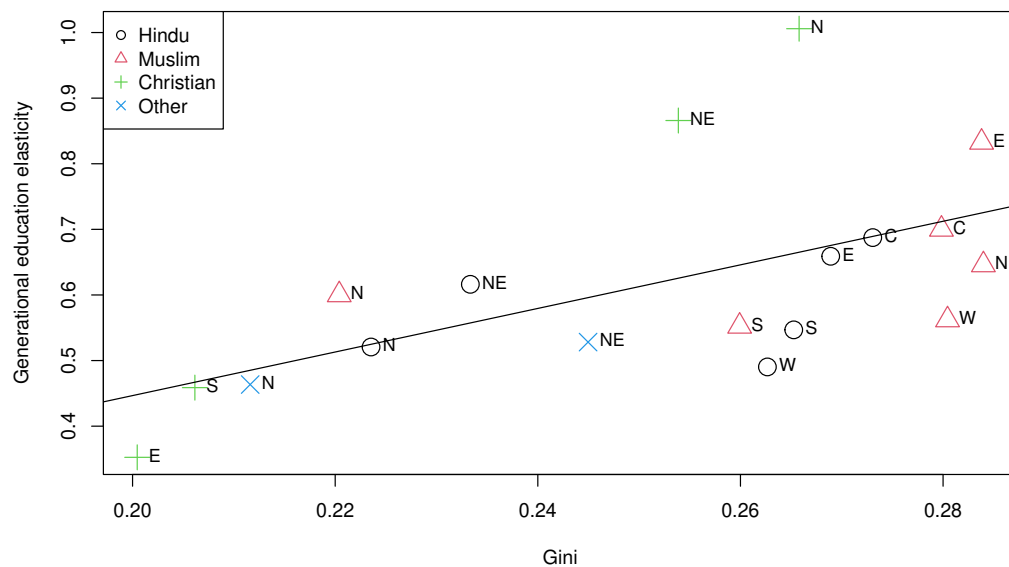
Note: the figure plots the Gini coefficient for education for each religious group in each zone on the x-axis and the corresponding β coefficient using Model I, i.e. Equation (1) on the y-axis.

Figure 8 Great Gatsby curve for religious groups based on Model II



Note: 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 II, i.e Equation (2) on the y-axis.

Figure 9 Great Gatsby curve for religious groups based on Model III



Note: 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 III, i.e Equation (3) on the y-axis.

8 Discussion and conclusion

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) conceptualize 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, an identity that follows a social hierarchy¹⁰ is difficult to be 'compensated for'. As our findings indicate, identity, the region to which the person belongs, the educational inequality in their region, and their father's and grandfather's outcome, all add to their outcomes.

The purpose of this study is manifold. On the one hand, it aims at establishing that parental altruism is socioeconomically driven and cannot be narrowly understood as limited by their income. Altruism in a society marked by a hierarchy can mean more things than just the willingness to spend on human capital investments as dictated by the parent's ability to pay. We argue that the motives of parents reflect their aspirations for their children which is a direct consequence of their position in society. The paper as a result establishes that the persistence mechanisms are indeed different for various identity groups and are also influenced by the region to which they belong. On the other hand, the study accounts for the effect of an external shock—the legal abolition of untouchability—on the inherent social structures that keep the caste system intact and explore its implications on the father's generation's educational attainment. Furthermore, the same law is used as an instrumental variable to account for the endogeneity in the father's education. The study also aims at exploring the generational elasticity of education in tandem with the average educational attainment across generations, the average income of households, and their position in the social hierarchy so as to comment on whether their inherent advantage or disadvantage is likely to be reinforced or disintegrated. We find that for those belonging to the General category, there is a *reinforcement of advantage*, while for those belonging to the Scheduled Caste (SC) or Scheduled Tribes (ST), there is a *disintegration of disadvantage* but one that is limited by their lower income and consequently lower spending on children's human capital. In contrast, among religious groups, Christians also observe a *disintegration of disadvantage* and are further backed by higher average income which materializes into greater educational investment in the next generation. Among Muslims, there is a tendency of *reinforcement of disadvantage*. For Muslims, discrimination is both structural and social. The lack of affirmative action in the educational and occupational domains and persistent discrimination in the labour market may explain why investments in children's educational attainment do not help the next generation escape their stigmatized identity.

Muslims are socioeconomically the most vulnerable among identity groups, and therefore, need interventions specifically designed to lift them out of their inherent disadvantage. However, directly targeting persistence can be very difficult for policy-makers. Since our study establishes that region–identity (caste or religion) cells with high educational inequality correspond to high generational persistence or low social mobility (*The Great Gatsby* relation), we contend that policy-makers should target educational inequality, which is much easier than targeting persistence itself.

References

- Alam, M. S. (2010). 'Social Exclusion of Muslims in India and Deficient Debates about Affirmative Action: Suggestions for a New Approach'. *South Asia Research*, 30(1): 43–65. <https://doi.org/10.1177/026272800903000103>
- Alesina, A., Hohmann, S., Michalopoulos, S., and Papaioannou, E. (2021). 'Intergenerational Mobility in Africa'.

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

- Econometrica*, 89(1): 1–35. <https://doi.org/10.3982/ECTA17018>
- Azam, M. (2019). ‘Household Income Mobility in India, 1993–2011’. Available at SSRN. <https://doi.org/10.2139/ssrn.2815008>
- Azam, M., and Bhatt, V. (2015). ‘Like Father, like Son? Intergenerational Educational Mobility in India’. *Demography*, 52(October): 1929–59. <https://doi.org/10.1007/s13524-015-0428-8>
- Becker, G. S. (1974). ‘A Theory of Social Interactions’. *Journal of Political Economy*, 82(6): 1063–93.
- Becker, G. S., and Tomes, N. (1979). ‘An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility’. *Journal of Political Economy*, 87(6): 1153–89.
- Behrman, J., and 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*, 67(1): 144–51. <https://doi.org/10.2307/1928446>
- Bernheim, B. D., Shleifer, A., and Summers, L. H. (1986). ‘The Strategic Bequest Motive’. *Journal of Labor Economics*, 4(3, Part 2): S151–S182. <https://doi.org/10.1086/298126>
- Black, S. E., and Devereux, P. J. (2010). ‘Recent Developments in Intergenerational Mobility’. NBER Working Paper 15889. Cambridge, MA: National Bureau of Economic Research. <https://doi.org/10.3386/w15889>
- Boserup, S. H., Kopczuk, W., and Kreiner, C. T. (2013). *Intergenerational Wealth Mobility: Evidence from Danish Wealth Records of Three Generations*. Mimeo, University of Copenhagen.
- Checchi, D., Fiorio, C. V., and Leonardi, M. (2013). ‘Intergenerational Persistence of Educational Attainment in Italy’. *Economics Letters*, 118(1): 229–32. <https://doi.org/10.1016/j.econlet.2012.10.033>
- Chetty, R., Hendren, N., Kline, P., and 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): 1553–623. <https://doi.org/10.1093/qje/qju022>
- Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014). ‘Is the United States still a Land of Opportunity? Recent Trends in Intergenerational Mobility’. *American Economic Review*, 104(5): 141–47. <https://doi.org/10.1257/aer.104.5.141>
- Clark, G. (2012). *What Is the True Rate of Social Mobility in Sweden? A Surname Analysis, 1700–2012* [Unpublished manuscript, Univ. California, Davis].
- Clark, G., and Cummins, N. (2014). ‘Surnames and Social Mobility in England, 1170–2012’. *Human Nature*, 25(November): 517–37. <https://doi.org/10.1007/s12110-014-9219-y>
- Corak, M. (2013). ‘Income Inequality, Equality of Opportunity, and Intergenerational Mobility’. *Journal of Economic Perspectives*, 27(3): 79–102. <https://doi.org/10.1257/jep.27.3.79>
- Corak, M., Lindquist, M. J., and Mazumder, B. (2014). ‘A Comparison of Upward and Downward Intergenerational Mobility in Canada, Sweden and the United States’. *Labour Economics*, 30(October): 185–200. <https://doi.org/10.1016/j.labeco.2014.03.013>
- Das, M. (2007). ‘Persistent Inequality: An Explanation Based on Limited Parental Altruism’. *Journal of Development Economics*, 84(1): 251–70. <https://doi.org/10.1016/j.jdeveco.2005.10.006>
- Davis, J. B. (2015). ‘Stratification Economics and Identity Economics’. *Cambridge Journal of Economics*, 39(5): 1215–29. <https://doi.org/10.1093/cje/beu071>
- Desai, S., and Vanneman, R. (2010). *National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005*. Ann Arbor: Inter-University Consortium for Political and Social Research [distributor].
- Desai, S., and Vanneman, R. (2015). *India Human Development Survey-II (IHDS-II), 2011–12*. Ann Arbor: Inter-University Consortium for Political and Social Research.
- Dhar, D., Jain, T., and Jayachandran, S. (2019). ‘Intergenerational Transmission of Gender Attitudes: Evidence from India’. *The Journal of Development Studies*, 55(12): 2572–92. <https://doi.org/10.1080/00220388.2018.1520214>
- Encyclopædia Britannica (n.d.). *untouchable*. Encyclopædia Britannica. Available at: <https://www.britannica.com/topic/untouchable> (accessed in June 2023)
- Ferreira, F. H., Messina, J., Rigolini, J., López-Calva, L.-F., Lugo, M. A., and Vakis, R. (2012). *Economic Mobility and the Rise of the Latin American Middle Class*. Washington, DC: The World Bank. <https://doi.org/10.1596/978-0-8213-9634-6>
- Gong, H., Leigh, A., and Meng, X. (2012). ‘Intergenerational Income Mobility in Urban China’. *Review of Income and Wealth*, 58(3): 481–503. <https://doi.org/10.1111/j.1475-4991.2012.00495.x>
- Güell, M., Rodríguez Mora, J. V., and Solon, G. (2018). ‘New Directions in Measuring Intergenerational Mobility: Introduction’. *The Economic Journal*, 128(612): F335–F339. <https://doi.org/10.1111/econj.12607>
- Heath, A., and Zhao, Y. (2021). ‘Rethinking Occupational Mobility in Developing Countries’. In V. Iversen, A. Krishna, and K. Sen (eds), *Social Mobility in Developing Countries: Concepts, Methods, and Determinants*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780192896858.003.0008>

- Hertz, T., Jayasundera, T., Piraino, P., Selcuk, S., Smith, N., and Verashchagina, A. (2008). 'The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends'. *The BE Journal of Economic Analysis & Policy*, 7(2): n/a. <https://doi.org/10.2202/1935-1682.1775>
- Hnatkovska, V., Lahiri, A., and Paul, S. (2012). 'Castes and Labor Mobility'. *American Economic Journal: Applied Economics*, 4(2): 274–307. <https://doi.org/10.1257/app.4.2.274>
- Iversen, V., Krishna, A., and Sen, K. (2016). 'Rags to Riches? Intergenerational Occupational Mobility in India'. GDI Working Paper 2016-004. Manchester: Global Development Institute, The University of Manchester.
- Kotlikoff, L. J., and Spivak, A. (1981). 'The Family as an Incomplete Annuities Market'. *Journal of Political Economy*, 89(2): 372–91.
- Kundu, A., and Sen, K. (2023). 'Multigenerational Mobility among Males in India'. *Review of Income and Wealth*, 69(2): 395–418. <https://doi.org/10.1111/roiw.12568>
- Lambert, S., Ravallion, M., and Van de Walle, D. (2014). 'Intergenerational Mobility and Interpersonal Inequality in an African Economy'. *Journal of Development Economics*, 110(September): 327–44. <https://doi.org/10.1016/j.jdeveco.2014.05.007>
- Li, Y., and 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. <https://doi.org/10.17645/si.v5i1.857>
- Lindahl, M., Palme, M., Massih, S. S., and Sjögren, A. (2015). 'Long-term intergenerational persistence of human capital: An empirical analysis of four generations'. *Journal of Human Resources*, 50(1): 1–33. <https://doi.org/10.3368/jhr.50.1.1>
- Lindahl, M., Palme, M., Sandgren-Massih, S., and 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. <https://doi.org/10.1086/674104>
- Marrero, G. A., and Rodríguez, J. G. (2013). 'Inequality of Opportunity and Growth'. *Journal of Development Economics*, 104(September): 107–22. <https://doi.org/10.1016/j.jdeveco.2013.05.004>
- Moav, O. (2002). 'Income Distribution and Macroeconomics: The Persistence of Inequality in a Convex Technology Framework'. *Economics Letters*, 75(2): 187–92. [https://doi.org/10.1016/S0165-1765\(01\)00625-5](https://doi.org/10.1016/S0165-1765(01)00625-5)
- Moreno, H. (2021). 'The Influence of Parental and Grandparental Education in the Transmission of Human Capital'. ECINEQ Working Paper 588. Rome: ECINEQ – Society for the Study of Economic Inequality.
- Narayan, A., Van der Weide, R., Cojocar, A., Lakner, C., Redaelli, S., Mahler, D. G., Ramasubbaiah, R. G. N., and Thewissen, S. (2018). *Fair Progress? Economic Mobility across Generations around the World*. Washington, DC: World Bank.
- Neidhöfer, G., Serrano, J., and Gasparini, L. (2018). 'Educational Inequality and Intergenerational Mobility in Latin America: A New Database'. *Journal of Development Economics*, 134(September): 329–49. <https://doi.org/10.1016/j.jdeveco.2018.05.016>
- Pollak, R. A. (1988). 'Tied Transfers and Paternalistic Preferences'. *The American Economic Review*, 78(2): 240–44.
- Sachar, R., Hamid, S., Oommen, T., Basith, M., Basant, R., Majeed, A., and Shariff, A. (2006). *Social, Economic and Educational Status of the Muslim Community of India*. New Delhi: Prime Minister's High Level Committee, Cabinet Secretariat, Government of India.
- Sakri, D., Sumner, A., and Yusuf, A. A. (2023). *Great Gatsby and the Global South: Intergenerational Mobility, Income Inequality, and Development* [Elements in Development Economics Series]. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781009382700>
- Serneels, P., and Dercon, S. (2021). 'Aspirations, Poverty, and Education. Evidence from India'. *The Journal of Development Studies*, 57(1): 163–83. <https://doi.org/10.1080/00220388.2020.1806242>
- Solon, G. (2014). 'Theoretical Models of Inequality Transmission across Multiple Generations'. *Research in Social Stratification and Mobility*, 35(March): 13–18. <https://doi.org/10.1016/j.rssm.2013.09.005>
- Solon, G. (2018). 'What Do We Know so Far about Multigenerational Mobility?' *The Economic Journal*, 128(612): F340–F352. <https://doi.org/10.1111/ecoj.12495>
- Srikanth, C., and Dey, S. (2023). 'Conspicuous Consumption for Social Parity'. WIDER Working Paper 2023/78. Helsinki. <https://doi.org/10.35188/UNU-WIDER/2023/386-4>
- Stuhler, J. (2012). 'Mobility across Multiple Generations: The Iterated Regression Fallacy'. IZA Discussion Paper 7072. Bonn: Institute of Labor Economics. <https://doi.org/10.2139/ssrn.2192768>
- Wu, X., and Treiman, D. J. (2007). 'Inequality and Equality under Chinese Socialism: The Hukou System and Intergenerational Occupational Mobility'. *American Journal of Sociology*, 113(2): 415–45. <https://doi.org/10.1086/518905>

Appendix

A 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* (Encyclopædia Britannica n.d.).
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* were removed.

B 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

Note: 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.

C Tables

Table C1 Hausman test for caste and religious groups in each zone

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

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 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 C2 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*** (0.009)	6.737*** (0.009)	Yes	56.046*** (0.000)	78.327*** (0.000)	4.071*** (0.045)	4.049*** (0.044)	Yes	56.046*** (0.000)	78.327*** (0.000)
OBC	1.152 (0.284)	1.154 (0.283)	No			0.722 (0.398)	0.742 (0.389)	No		
SC/ST	1.884 (0.170)	1.887 (0.170)	No			9.934*** (0.002)	9.612*** (0.002)	Yes	32.692*** (0.000)	38.449*** (0.000)
Hindu	14.684*** (0.000)	14.605*** (0.000)	Yes	410.875*** (0.000)	507.307*** (0.000)	8.211*** (0.004)	8.071*** (0.005)	Yes	61.129*** (0.000)	76.261*** (0.000)
Muslim	2.856 (0.092)	2.857 (0.091)	No			1.060 (0.306)	1.085 (0.298)	No		
Christian	0.237 (0.632)	0.269 (0.604)	No			4.190** (0.044)	4.130** (0.042)	No		
Other	0.857 (0.355)	0.863 (0.353)	No					No		

Note: *** 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 chi² 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 C3 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*** (0.000)	16.045*** (0.000)	Yes	136.674*** (0.000)	186.784*** (0.000)	4.450** (0.035)	4.439** (0.035)	Yes	195.453*** (0.000)	294.813*** (0.000)
OBC	15.545*** (0.000)	15.359*** (0.000)	Yes	131.391*** (0.000)	150.122*** (0.000)	21.469*** (0.000)	20.823*** (0.000)	Yes	112.815*** (0.000)	138.742*** (0.000)
SC/ST	6.975*** (0.009)	6.923*** (0.009)	Yes	53.654*** (0.000)	59.584*** (0.000)	4.185** (0.041)	4.176** (0.041)	Yes	125.084*** (0.000)	160.161*** (0.000)
Hindu	49.807*** (0.000)	48.517*** (0.000)	Yes	390.125*** (0.000)	500.577*** (0.000)	23.982*** (0.000)	23.639*** (0.000)	Yes	458.204*** (0.000)	669.608*** (0.000)
Muslim	3.352 (0.068)	3.348 (0.067)	No			12.005*** (0.001)	11.545*** (0.001)	Yes	72.399*** (0.000)	105.580*** (0.000)
Christian										
Other						3.067 (0.088)	3.062 (0.080)	No		

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 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 χ^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 C4 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*** (0.009)	6.816*** (0.009)	Yes	101.100*** (0.000)	153.310*** (0.000)	9.231*** (0.002)	9.142*** (0.003)	Yes	129.315*** (0.000)	161.616*** (0.000)
OBC	11.714*** (0.001)	11.635*** (0.001)	Yes	248.936*** (0.000)	307.756*** (0.000)	1.549 (0.214)	1.553 (0.213)	No		
SC/ST	7.035*** (0.008)	6.992*** (0.008)	Yes	132.926*** (0.000)	165.573*** (0.000)	0.020 (0.887)	0.020 (0.887)	No		
Hindu	22.959*** (0.000)	22.718*** (0.000)	Yes	427.924*** (0.000)	553.097*** (0.000)	6.651*** (0.010)	6.634*** (0.010)	Yes	340.457*** (0.000)	446.315*** (0.000)
Muslim	1.387 (0.240)	1.397 (0.237)	No			0.285 (0.595)	0.292 (0.589)	No		
Christian	0.196 (0.659)	0.201 (0.654)	No							
Other			No			0.386 (0.537)	0.401 (0.527)	No		

Note: *** 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 Western and Southern regions. Endogeneity is tested for using the *Wu-Hausman F test* and the *Durbin-Wu-Hausman chi² 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 C5 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

Note: 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 C6 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

Note: 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.

Table C7 Average incomes among caste and religious groups in each zone

	North	North-East	Central	East	West	South
General	96,161.29	49,948.17	52,411.52	49,180.17	72,654.62	70,454.31
OBC	54,155	70,193.95	31,266.31	27,997.6	48,993.84	53,000.22
SC/ST	46,354.72	60,151.41	25,087.3	28,722.29	40,678.86	41,829.58
Hindu	65,000.71	57,384.93	34,056.33	35,875.77	57,121.57	52,562.32
Muslim	58,775.99	34,766.15	31,220.91	28,958.94	44,082.49	45,859.04
Christian	53,415.68	91,751.24	137,368.5	41,327.35	83,303.32	77,093.01
Other	91,780.63	44,220.46	89,433.44	26,869.54	48,174.57	52,086.02

Note: the table shows the average equivalent income after accounting for sampling weights for each household in the 2011–12 period. The average income is grouped by zone and group (caste and religion separately).