Matching on Height in India

Ajinkya Keskar*

October 26, 2021

Abstract

In India, height is greatly valued in the marriage market, and the child stunting rate remains strikingly high. In this paper, I juxtapose these two seemingly unrelated facts and investigate the role of parents' marital sorting and matching in determining children's height. First, I develop a two-sided matching model of the Indian marriage market to structurally estimate preferences for height. I do so while considering other critical drivers of marital sorting and matching, such as education and family wealth. I find evidence of significant positive assortative matching on height across religion-caste groups. For the majority of religion-caste groups, I also find cross-complementarity in men's education and women's height but not vice-versa. Next, I study changes in complementarities over time, finding a mild increase on average but substantial heterogeneity by caste and religion. Finally, using the model estimates, I simulate parents' counterfactual joint height distribution under several hypothetical scenarios. Based on insights from the medical literature, I also compute children's potential height distribution (and hence their risk of being stunted) given the counterfactual distribution of matches. I find marital sorting and matching to have a limited impact on children's average height, but a significant one on the level of inequality in children's height. Specifically, my analysis indicates that complementarities in height in the marriage market can increase the standard deviation of the distribution of potential height by up to 3% and can increase the prevalence of stunting by up to 4 percentage points.

Keywords: Marriage, Height, Stunting, India, Transferable Utility.

JEL codes: J11, J12, J16

*Rice University, Department of Economics, Houston, TX. E-mail: ajinkya.keskar@rice.edu.

I am grateful to Rossella Calvi, Flávio Cunha, Yinghua He and Isabelle Perrigne for their invaluable guidance and support. I thank Edoardo Ciscato, Hira Farooqi, Amir Kazempour Esmati, Dibya Mishra, Ritika Sethi and Rabia Telli for their suggestions and insight.

All errors are my own.

1 Introduction

Indian children face a strikingly high risk of stunting, with 31% of children between the ages of 0-5 being too short for their age (WHO et al., 2021).[1] India also has one of the world's shortest adult population; according to the latest global height rankings, Indian men and women rank 178 and 180 out of 198 countries, respectively.[2] Contrary to many other developing countries, over the last century, the height of Indians has increased marginally; on average, Indian men and women have grown by just three and five centimeters, respectively (NCD et al., 2016).[3]

Previous work has investigated the role of malnutrition, the disease environment and son preference and women's bargaining power as the main determinants of the low height-for-age of Indian children (Deaton, 2007; Coffey, Deaton, Drèze, Spears, and Tarozzi, 2013a; Coffey, Khera, and Spears, 2013b; Jayachandran and Pande, 2017; Coffey and Spears, 2017; Calvi, Lewbel, and Tommasi, 2021). Since height is a highly heritable trait, parental height also plays a role, and sorting and matching on height in the marriage market in one generation may prove essential to understanding the height distribution in the next.[4] As child height has significant consequences on long-term individual well-being through its association with adult height, cognitive ability, and income (Tanner, Healy, Lockhart, Mackenzie, and Whitehouse, 1956; Case and Paxson, 2008), understanding preferences for height in the marriage market, how they have evolved over time, and how they may affect the height distribution of the next generation is of primary importance.

There is a vast literature from evolutionary biology and sociology studying the importance of anthropometric features in the marriage market.[5] A growing number of research articles in economics have also analyzed marital sorting and matching on physical attributes, such as height, weight, and body mass index. Most of these works, however, have focused on developed countries (Oreffice and Quintana-Domeque, 2010; Chiappori, Oreffice, and Quintana-Domeque, 2012a, 2016).[6]

Common wisdom is that height is greatly valued in the marriage market in India (Banerjee, Duflo, Ghatak, and Lafortune, 2013).[7] In this paper, I study this fact rigorously and estimate preferences for height in the Indian marriage market using a two-sided matching model with transferable utility. Based on the model estimates and insights from the medical literature, I then analyze how parents' marital sorting and matching on height impact children's height distribution and risk of stunting. My analysis indicates that complementarities in height in the marriage market can increase the standard deviation of the distribution of children's height by up to 3% and can increase the prevalence of stunting by up to 4 percentage points.

I start by providing descriptive evidence regarding the relationship between husband's and wife's attributes in the spirit of Oreffice and Quintana-Domeque (2010) using data from the second wave of

[1]A child is defined as stunted if their height-for-age is two standard deviations or more below the world reference population median for a given gender and age.

[2]The average height of Indian men and women is 166 and 155 centimeters, respectively. Source: https://worldpopulationreview.com/country-rankings/average-height-by-country

[3]For example, in contrast, the height of Chinese and Brazilian adults on average has increased by more than ten centimeters.

[4]Evidence from medical literature suggesting that up to 80% of the variation in height can be explained by genetics (Stulp and Barrett, 2016).

[5]See Courtiol, Raymond, Godelle, and Ferdy (2010) for a review of the literature from evolutionary biology about matching on stature. See Stulp, Simons, Grasman, and Pollet (2017) for a meta-analysis, summarizing the literature from sociology on matching on height.

[6]Notable exception is a very recent paper by Chiappori, Ong, Yang, and Zhang (2021) in China.

[7]A common practice in India is to place matrimonial advertisements in newspapers to look for a spouse. Height is mentioned in the ad by over 90% of adults (Banerjee et al., 2013). Another common practice in India is to look for a spouse on matrimonial websites. On most of these websites, individuals looking for a spouse list their height (see the following link for an example: https://www.jeevansathi.com/matrimonials/hindu-matrimonial). From the recent Netflix show, Indian Matchmaking, one of Mumbai's (a large metropolitan city in India) top matchmakers says the following about the Indian marriage market: "In India, we have to see the caste, we have to see the height, we have to see the age."

the Indian Human Development Survey(IHDS). I find strong positive associations between men's and women's height, education, and parental education, indicating positive assortative matching along these attributes. I also document a significant association between men's height and women's fathers' education. Next, I quantify the impact of parental height on children's height by using linear regression analysis.[8] I find that, on average, conditional on household-level determinants and demographic characteristics, a child with both parents above median height rank has 0.7 standard deviations larger height-for-age than a child with both parents below median height rank.[9] Similarly, a child with both parents above median height rank has a 16 percentage point lower probability of stunting than a child with both parents below median height rank. This analysis provides preliminary evidence of some of the relevant dimensions of sorting in India and how sorting on height can impact children's height distribution.

To understand the role of preferences for height in the Indian marriage market, I estimate a structural static frictionless two-sided transferable utility model of the marriage market using the framework of Dupuy and Galichon (2014). In this framework, every man and woman in the marriage market is characterized by a feature vector consisting of height, education, father's education, mother's education, and age.[10] The critical component of the model is the "affinity matrix," which measures the intensity of complementarity/substitutability across features.[11] For example, if the parameter of the affinity matrix corresponding to men's education and women's height is positive, it indicates positive assortative matching on those two features. Since inter-caste or inter-faith marriages are extremely rare in India (Banerjee et al., 2013; Borker, Eeckhout, Luke, Minz, Munshi, and Swaminathan, 2017), I model each religion-caste group, Scheduled Caste, Scheduled Tribe, Other Backward Caste, Upper Caste, and Muslims as a separate marriage market.[12] For each market, I estimate complementarities in height, education, parental education, and age, as well as cross-complementarities between education and height, parental education and height, and parental education and spousal education.

The affinity matrix estimates indicate strong positive assortative matching on height, with taller men and women finding each other mutually attractive in the marriage market in India. For the majority of religion-caste groups, I find significant cross-complementarities in men's education and women's height but not vice-versa, indicating that more educated men and taller women find each other mutually attractive in the marriage market. I also find significant cross-complementarities in men's height and women's father's education in three of the five religion-caste-specific marriage markets, indicating that taller men

[8]The medical literature documents a strong association between parental height, in particular maternal height and offspring stunting hazard across the developing world (Özaltin, Hill, and Subramanian, 2010).

[9]In particular, I regress children's height for age z-score and stunting status on the mother's and father's height rank and other critical determinants of height such as household expenditure, the prevalence of piped water (as a proxy for the child's disease environment (Coffey and Spears, 2017)) and parental education.

[10]I use the father's education as a proxy for the family's wealth status, and the mother's education captures the family's gender attitudes (Dhar, Jain, and Jayachandran, 2019)

[11]A primitive of the transferable utility model is the joint marital surplus. Intuitively, an equilibrium matching is determined by maximizing the sum of marital surplus over all possible matches. The second derivative of the joint marital surplus concerning a man and a woman's characteristics is the complementary/substitutability between the two features. For example, the complementary/substitutability in height is the second derivative of the joint marital surplus regarding the height of men and women. In the model, the affinity matrix is identified from the variation in equilibrium match probability of men and women with different features. See section 4 for details.

[12]According to the IHDS, 95% of the respondents marry someone within their own caste.

Articles 341 and 342 of the Indian constitution define the term Scheduled Caste and Scheduled Tribe precisely, respectively. Individuals belonging to these castes suffer from extreme forms of social, educational, and economic backwardness with much lower living standards than the rest of the population (Gang, Sen, and Yun, 2008). The Other Backward Caste community consists of individuals who are better off than Scheduled Caste and Scheduled Tribe members but less well-to-do than Upper Caste members. According to the 2011 census, 80% of Indians are Hindus, 14% are Muslims, and Scheduled Caste and Scheduled Tribes comprise 16.6% and 8.6%, respectively, of India's population. The Census doesn't provide numbers for the Other Backward Caste. Still, according to the 2015 National Family Health Survey, 44% of individuals belong to the Other Backward Caste community, making it the largest caste group in India.

Ideally, we would like to model district-religion-caste-specific marriage markets, given that most marriages in India are within a given district. The travel time between the marital and natal family is 3-4 hours (Beauchamp, Calvi, and Fulford, 2017; Fulford et al., 2013); however, due to the sample size, this is not feasible. As a result, I consider religion-caste-specific marriage markets instead of district-religion-caste-specific marriage markets.

and women from wealthier natal families find each other mutually attractive. Similarly, the results show significant positive assortative matching between more educated men and women from more affluent families in all the marriage markets. Lastly, I also find evidence of cross-complementarity between women's education and men's mother's education.

Marital sorting and matching patterns have been shown to have important consequences for the intergenerational transmission of inequality (Greenwood, Guner, Kocharkov, and Santos, 2014; Chiappori, Salanié, and Weiss, 2017; Ciscato and Weber, 2020). There is a growing strand of literature studying the changes in marital sorting patterns over time, mainly focusing on sorting on education. I contribute to this literature by analyzing changes in marital sorting on height.[13] I estimate the affinity matrix separately for an old and a young cohort and compute the change in complementarities in the height of men and women in the marriage market. The results indicate significant heterogeneity across religion-caste groups.[14] I estimate the degree of assortativeness on height to have increased over time for couples belonging to the Scheduled Tribe, Muslim, and the Other Backward Caste group. By contrast, I find no change for couples belonging to the Upper Caste or the Scheduled Caste.

Using the model estimates and insights from the medical literature (Tanner, Goldstein, and Whitehouse, 1970), I simulate the effect of a series of counterfactual experiments.[15] In the first experiment, I evaluate children's height distribution under various hypothetical marriage market preferences. The analysis involves two steps. In the first step, I compute the counterfactual joint distribution of height under the hypothetical scenario of no complementarities in height in the marriage market. In the second step, I simulate children's counterfactual potential height (defined as the mean parental height adjusted for the child's gender) distribution, using the computed counterfactual joint distribution of height.[16] Results from the simulation indicate that, although the average potential height of boys and girls does not change, there are substantial effects on their potential height distribution. Under no complementarities, we see a decrease in the prevalence of boys and girls with short stature and a reduction in the prevalence of boys and girls with tall stature. The standard deviation of the potential height distribution decreases, indicating a decrease in the height inequality among children of a given religion-caste group. Lastly, to further understand the relationship between marriage market preferences for height and children's height distribution, I compute children's counterfactual height for age z-score distribution under hypothetical marriage market preferences.[17] Results from the simulation indicate that complementarities in height in the marriage market can increase the standard deviation of children's height distribution by up to 3% and, as a result, can increase the prevalence of stunting by up to 4%. In other words, complemen-

[13] Across the developing world, the evidence regarding the increase in educational homogamy over time is mixed (Anukriti and Dasgupta, 2017). Smits and Park (2009), analyzing data from ten East-Asian societies, find that educational homogamy was higher at higher levels of education. Since 1950 it decreased at all, except the lowest level of schooling since then. In a recent paper studying the changes in assortative matching in Mexico, Hoehn-Velasco and Penglase (2021) find that educational homogamy among college graduates has grown substantially over time. Similarly, Ganguli, Hausmann, and Viarengo (2014) find that assortative matching on education has increased from 1980 to 2000.

A substantial body of research looks at changes in marital patterns over time in developed countries. Fernández and Rogerson (2001); Greenwood et al. (2014); Chiappori et al. (2017); Chiappori, Dias, and Meghir (2020c); Ciscato and Weber (2020) find that preferences for partners with the same education have increased in the United States. Chiappori, Costa-Dias, Crossman, and Meghir (2020b) do not observe a change in assortativeness by education in the United Kingdom between 1945-54 and 1965-74.

[14] In particular, I define two non-overlapping cohorts: the old cohort consists of men born between 1950-1960 and married to women born between 1950-1970. The young cohort consists of men born between 1970-1980 and married to women born between 1970-1990. I expand on the cohort selection procedure in section 4.4.

[15] A child's potential height represents the transmission of height from parents to children. Mid-parental height(average of the mother's and father's height) is frequently used by pediatricians to measure a child's growth potential (Cole, 2000). Tanner's method uses mid-parental height, adjusted for the child's gender to measure a child's growth potential.

[16] According to the WHO Multicentre Growth Reference Study (Garza, Borghi, Onyango, de Onis, and Group, 2013), potential height explains about 21% of the variability in linear growth from birth to 2 years of children in India.

[17] Section 6 details the counterfactual simulation procedure.

tarities in height do not impact children's average height but do have a significant distributional effect and, as a result, can impact the stunting hazard and the inequality in children's height. To assess the magnitude of this result, in comparison, Jayachandran and Pande (2017) find that due to strong eldest son preference in India, relative to their African counterparts, lower birth order children are on average 5.5 percentage points more likely to be stunted compared to their oldest sibling. Lastly, I also simulate the counterfactual scenario of no change in complementarities over time, finding a small change in the joint height distribution for the Scheduled Tribe and Muslim couples, however the impact on children's counterfactual potential height distribution is negligible.

To the best of my knowledge, the two unique contributions of this paper are: one, to estimate preferences for height in the marriage market across religion-caste groups in India. Two, to link the relationship between preferences for height in the marriage market to the realized children's height distribution and child stunting.

The rest of the paper is organized as follows. Section 2 summarizes the existing literature. Section 3 provides descriptive evidence regarding marital sorting on height and quantifies the impact of parental height on children's height. Section 4 describes the model, identification and estimation strategy. Section 5 presents the main results of the paper and robustness checks. Section 6 shows the counterfactual simulation results. Section 7 concludes the paper.

2 Literature Review

This paper is related to the works studying marital matching and sorting on human capital investments in developing countries, focusing on anthropometric features, mechanisms driving child stunting and the changes in marital matching patterns over time.[18]

There is a growing body of work analyzing various aspects of the Indian marriage market.[19] In an insightful paper, Banerjee et al. (2013) develop a non-transferable utility model of the marriage market and estimate preferences for caste and other attributes using a unique dataset of upper-middle-class Indian families who placed matrimonial advertisements in a local newspaper in Kolkata, India.[20] The authors find strong preferences for within-caste marriage and that the preferences for caste are horizontal rather than vertical. The preferences for caste-endogamous are so strong that, bride's family is willing to trade off the difference between no education and a Master's degree in prospective groom to avoid marrying outside caste.[21] The results of their paper lead me to model religion-caste groups as separate marriage markets.

[18] The importance of human capital investments on marital matching in developed countries is extensively analyzed, see Chiappori (2020) for an in-depth review.

[19] Likewise, there is abundant literature studying the importance of human capital investments in the marriage market in other developing countries. Using data from Indonesia and Zambia, Ashraf, Bau, Nunn, and Voena (2020) find that the likelihood of a girl being educated is higher among ethnic groups that practice bride-price and that families from bride-price groups are the most responsive to policies, like school construction, which are aimed at increasing female education. Boulier and Rosenzweig (1984) document positive assortative matching in the couple's educational levels using data from the Philippines. Fafchamps and Quisumbing (2005) report a positive correlation between spousal characteristics along, age, years of schooling, and parental land using data from rural Ethiopia. Boxho, Donald, Goldstein, Montalvao, and Rouanet (2020) find positive assortative matching on cognitive, socio-emotional skills and risk preference in rural Mozambique.

[20] Dugar, Bhattacharya, and Reiley (2012) take a similar approach and design a field experiment in which they place newspaper advertisements of potential grooms by varying caste and income. The authors document strong preferences for within-caste marriage but find that this preference decreases with the income of lower caste males.

Dupuy, Galichon, and Sun (2016) develop a novel methodology to estimate the affinity matrix as in Dupuy and Galichon (2014) when the data is high-dimensional. They apply their methodology to the data collected by Banerjee et al. (2013) and corroborate the finding of same-caste marriage preferences.

[21] Rosenzweig and Stark (1989) rationalize caste-endogamous marriages in India by showing that marriages in rural India are used as a consumption smoothing and risk-sharing device across households.

Differently from their framework, I model marriage markets in a transferable utility framework and focus on estimating preferences for height. Finally, while their data consists of mainly upper-caste families from a particular region of India, I use a nationally representative dataset to estimate marriage market preferences for different religion-caste groups. In line with the paper's finding that upper-middle-class men and women from Kolkata, India, prefer more educated spouses, I find strong positive assortative matching on education across religion-caste groups. Borker et al. (2017) model caste-specific marriage markets in a transferable utility framework and, using data from the rural population of Vellore district in Tamil Nadu, India, find significant positive assortative matching on wealth. Interestingly they find, conditional on wealth, the relation between the bride's education rank within her caste and the groom's education rank within his caste is insignificant. Further, the authors find that matching on wealth operates through an independent channel compared to matching on education. In line with their results, I see significant positive assortative matching on wealth; however, differently from their findings, I find significant cross-complementarities in husband's education and wife's natal family wealth.[22] In a recent paper, Beauchamp et al. (2017) model the Indian marriage market in a dynamic, general equilibrium, two-sided matching with non-transferable utility framework and estimate men's and women's individual-level preferences for education, age, and other attributes such as migration upon marriage, dowries, and preferences for an arranged marriage. Differently from their paper, I model the marriage market in a transferable utility framework and estimate preferences for anthropometric features along with cross-complementarities in education and height. One of the key results of their paper suggests that men's education is valued in the marriage market but not women's. In my framework, I am not able to identify men and women's preferences separately; however, I find significant complementarity in men and women's education across religion-caste groups and cross-complementarity in men's education and women's height, but not vice-versa for the largest caste-religion group in the sample, the Other Backward Caste. Adams and Andrew (2019) illicit parental preferences and subjective beliefs about the importance of education and age in the marriage market in rural Rajasthan, India. They find that conditional on a marital match; parents seldom value their daughter's education. The authors also find that parents believe the probability of receiving a good match increases with the daughter's education. The results from my paper are in line with their main findings, as I find significant sorting on education in the marriage market, with more educated men and women finding each other mutually attractive.

Research focusing on the relevance of anthropometric features in the marriage market in developing countries remains sparse.[23] Chiappori et al. (2021) using a novel dataset from China that tracks clicks on profiles with randomly assigned height and income information from a popular dating website calculate men and women's willingness to pay for mate height. Their results indicate that men prefer taller women, and women prefer taller and higher-income men.[24] Interestingly, they also find that shorter women have

[22] In the primary analysis, I use the father's educational attainment to capture information regarding family wealth since this is the best measure of wealth available for *both* the marital and natal in the IHDS. Using the 2006 wave of the Rural Economic and Demographic Survey, I use family landholding as a measure of family wealth in line with Borker et al. (2017) and find significant positive assortative matching on wealth corroborating their main result.

[23] Comprehensive research from developed countries studying the importance of anthropometric measures in the marriage market exists. In an influential paper, Chiappori, Oreffice, and Quintana-Domeque (2012b) develop a model of the marriage market, in which individual preferences are summarized into a single index. Using data from the United States, the authors find that men compensate 1.3 additional units of body mass index(BMI) with a 1% increase in wages and women compensate two units of BMI with one year of education. Oreffice and Quintana-Domeque (2010) also document positive sorting on BMI, height, and weight using data from the United States. Hitsch, Hortaçsu, and Ariely (2010) estimate mate preferences in a non-transferable utility framework using the Gale-Shapley algorithm from an online dating site data, which includes users from North America and Europe. Their results specific to height indicate that men avoid tall women, whereas women prefer taller men. Dupuy and Galichon (2014) using data from the Netherlands, find positive assortative matching on BMI and height of men and women. Chiappori, Ciscato, Guerriero et al. (2020a) using data from Naples, Italy, document positive assortative matching on spouses' height and BMI.

[24] Banerjee et al. (2013) find that 96% of women and 90% of men placing matrimonial advertisements in a local newspaper in Kolkata, India mention

the highest willingness to pay for height. Similar to their results, I find positive assortative sorting on height; however, I find that taller men and women from wealthier families find each other mutually attractive in the marriage market.

Recent research has analyzed changes in marital matching patterns over time in the developing world concerning education. I contribute to this research by studying changes in marital matching patterns across another key dimension of human capital, height. Evidence across the developing world regarding the increase in educational homogamy over time is mixed (Anukriti and Dasgupta, 2017). Smits and Park (2009), analyzing data from ten East-Asian societies, find that educational homogamy was higher at higher levels of education. Since 1950 it has decreased at all, except for the lowest level of education. In a recent paper studying the changes in assortative matching in Mexico, Hoehn-Velasco and Penglase (2021) find that educational homogamy among college graduates has grown substantially over time. Similarly, Ganguli et al. (2014) find that assortative matching on education has increased from 1980 to 2000 in Latin American countries.[25]

There is a growing strand of literature focusing on the determinants of children's height; my paper contributes to this literature by linking marriage market preferences and children's potential height distribution. Understanding the mechanisms determining children's height is especially salient in India, with a child stunting rate of more than 30%. One of the key drivers of child stunting in India is the child's disease environment and nutritional intake (Coffey et al., 2013a). Spears (2013) documents cross-country variation in sanitation and its role in explaining height differences across the globe. The author shows that a high rate of open defecation in India introduces germs in a child's environment that causes disease and stunts a child's growth. In an influential paper, Jayachandran and Pande (2017) show that the high rate of child stunting in India compared to poorer sub-Saharan countries is due to strong elder son preference. The authors find that higher birth order children, especially girls, have a significantly higher probability of stunting than their eldest brother. In alternate mechanism affecting children's height, Calvi et al. (2021) find that an increase in women's control over household resources positively impacts their children's height in India. A recent paper by Wang, Puentes, Behrman, Cunha et al. (2021) shows that the height parents target concerning their children depends on some reference population height, and parents make nutritional choices regarding their children following this target height. In their framework, reference height is an equilibrium object determined by an earlier cohort of children. Using data from a randomized control trial about a protein intervention in Guatemala, the authors show that 65% of variation in height between the treatment and control group can be accounted for changes in reference points. This fascinating result provides an additional channel through which preferences for height in the marriage market can influence children's height, apart from the direct mechanism that parent's height is a key determinant of children's height, as height is a highly heritable trait (Stulp and Barrett, 2016).[26] If marital sorting on height along with nutritional intake contributes in determining the reference population height distribution, then as parents target height for their children based on this reference population and make nutritional choices accordingly, marriage market preferences for height can indirectly impact

height.

[25] There is a growing strand of literature studying changes in marital sorting on education over time in the developed world. Fernández and Rogerson (2001); Greenwood et al. (2014); Chiappori et al. (2017, 2020c); Ciscato and Weber (2020) find that preferences for partners with the same education have increased in the United States. Chiappori et al. (2020b) do not observe an apparent change in assortativeness by education in the United Kingdom between 1945-54 and 1965-74

[26] In the context of India, from the medical literature, using data from 2015-16, National Family Health Survey Gupta, Cleland, and Sekher (2021) show that maternal height is the strongest predictor of child stunting, followed by paternal height and marital family wealth. Similarly, Özaltin et al. (2010) using data from fifty-four low and middle-income countries, find a strong association between maternal stature and offspring stunting hazard across the developing world.

children's height through their influence on the reference population height distribution.

3 Descriptive Evidence

The first part of this section, in the spirit of Oreffice and Quintana-Domeque (2010), establishes associations between men's and women's height, education, and parental education. The conditional correlations render valuable descriptive evidence regarding the dimensions of marital sorting in India. In the second part, I provide descriptive evidence regarding the impact of parental height on children's height conditional on key household-level variables, such as total expenditure known to impact children's height.

3.1 Determinants of Sorting

Table 1: Sorting on Height

	(1) Husband Height	(2) Wife Height
Wife Height	0.270***	
	(0.026)	
Wife Education	0.046	0.066**
	(0.029)	(0.029)
Husband Education	0.057**	0.042
	(0.021)	(0.027)
Husband Father Education	0.080***	0.027
	(0.024)	(0.033)
Wife Father Education	0.067*	0.004
	(0.036)	(0.023)
Husband Mother Education	-0.072	0.001
	(0.058)	(0.064)
Wife Mother Education	0.043	-0.029
	(0.042)	(0.031)
Husband Height		0.244***
		(0.024)
Region Fixed Effects	Yes	Yes
Year of Marriage Fixed Effects	Yes	Yes
Covariates	Yes	Yes
Observations	4,652	4,652
Mean of Dep. Variable	163.405	152.157

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

NOTES: Covariates include religion, caste, age of the husband and age of the wife. Husband and wife's height is measured in centimeters. Education is measured in number of years.Standard errors are clustered at the state level.

In table 1, I regress the husband's(wife's) height on spousal height controlling for other attributes such as education, parental education, and demographic characteristics such as caste, religion, state of residence, and year of marriage fixed effects, with standard errors clustered at the state level. We find a strong correlation in men and women's height, a one standard deviation increase in the women's height is associated with her having a 0.25 standard deviation taller husband(statistically significant at 1%). Similarly, a one standard deviation increase in men's height is associated with him having a 0.26 standard deviation taller wife(statistically significant at 1%).[27] We also observe a positive correlation between

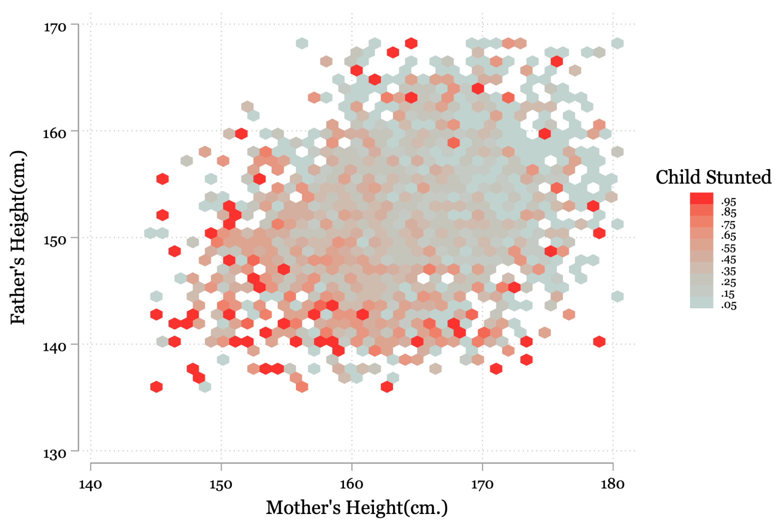
[27]In the appendix, figure A2 I perform a similar analysis, where I regress the husband's(wife's) height on all attributes similar to table 1 excluding

women's height and men's education and vice-versa; however, the coefficient is not statistically significant. Interestingly, we obtain a positive association between the wife's father's education and the husband's height, but the association between the husband's father's education and the wife's height is not statistically significant.

In the appendix, I perform a similar descriptive analysis concerning spousal education and parental education. Results from table A1 indicate that a one standard deviation increase in women's education is associated with her having a husband with 0.52 standard deviations more education(statistically significant at 1%) and, a one standard deviation increase in men's education is associated with him having a wife with 0.4 standard deviations more education(statistically significant at 1%). We also see a positive correlation between a wife's height and a husband's education and vice-versa. A one standard deviation increase in the women's height is associated with her having a husband with 0.02 standard deviations more education. A one standard deviation increase in the men's height is associated with him having a wife with 0.01 standard deviations more education. Results from table A2 and A3 provide suggestive evidence towards marital sorting on father's education and mother's education.

These results provide strong descriptive evidence towards marital sorting on height, education, and parental education in India and the presence of asymmetric cross-complementarities in men and women's marriage market features.

Figure 1: Child Stunting



Note: Figure plots the prevalence of child stunting based on mother's and father's height.

3.2 Children's Height and Parent's Height

The height for age z-score of a child is calculated with respect to the World Health Organization(WHO) reference scale.[28] A child with a height for age z-score two standard deviations or greater below the median world reference population for a given gender and age is considered stunted. Child stunting is widely recognized as a critical marker of child malnutrition (Deaton, 2007), which in turn has long-run consequences on individual well-being since child height is known to influence adult height, cognitive ability, and income (Tanner et al., 1956; Case and Paxson, 2008). The importance of nutrition, disease

spousal height and the plot the residuals of the regression concerning spousal height. Figure A3 shows a similar residual plot concerning education. These plots provide further suggestive evidence regarding sorting on height and education.

[28] In particular, I calculate the height for age z-score for children of age 0 to 19 using 2007 WHO reference chart in STATA using the package *zanthro* (Vidmar, Cole, and Pan, 2013)

environment, son preference and women's bargaining power in explaining child stunting in India has been extensively analyzed in the literature (Spears, 2013; Coffey and Spears, 2017; Jayachandran and Pande, 2017; Calvi et al., 2021). The medical literature has shown height to be a highly heritable trait, with around 80% of the variation in height explained by genetics (Stulp and Barrett, 2016) and has provided evidence for a strong association between parental height and child stunting (Özaltin et al., 2010; Gupta et al., 2021) . Here I provide further descriptive evidence concerning the role of marital matching on height in determining children's height for age z-score.

Figure 1 plots the prevalence of child stunting with respect to parental height among children aged 0-19 born to parents belonging to a particular marital cohort consisting of men born between 1970 and 1980 and women born between 1970 and 1990.[29] The figure suggests that parents with short stature are more likely to have stunted children, and as the mother and father's height increases, their children are less likely to be stunted.

However, parental height is highly correlated with other vital determinants of stunting such as educational attainment and household income.[30] To understand the role of parental height as a determinant of children's height, conditional on other key household-level and demographic determinants of children's height, in table A5, I regress children's height for age z-score and stunting status on whether the parents are above or below median and height, parental education, parental age, household expenditure, whether the household has piped water, the mother's parent's literacy, the father's parent's literacy, child's gender, along with district fixed effects, religion-caste group fixed effects and age fixed effects, with standard errors clustered at the state level. Household expenditure captures information regarding the child's nutritional intake and piped water and district fixed effects proxies a child's disease environment (Spears, 2013), which are known to play a critical role in determining a child's height. Results indicate that *ceteris paribus*, the probability that a child with above the median height rank parents is stunted is 16 percentage points lower than a child with below-median height rank parents. Similarly, a child with above-median height rank parents has a 0.7 standard deviations greater height for age z-score than a child with below-median height rank parents. Figure A23 illustrates the probability of stunting for children conditional on household-level characteristics for different combinations of parental height rank.

These results provide preliminary descriptive evidence regarding the role of parental height in determining the child's height conditional on other essential variables such as household expenditure and parental education. If preferences for height play a role in deciding who marries whom on the marriage market, then preferences for height will also impact children's height for age z-score distribution.

In the next section, I estimate a structural to understand the role of preferences for height on the Indian marriage market. The model allows me to do counterfactual simulations and study how changes in preferences for height on the marriage market can impact children's height distribution.

4 A Structural Analysis of Marital Matching on Height

The descriptive analysis provides a useful starting point for analyzing the relevance of height as a determinant of marital sorting in India. In order to understand the role of preferences for different marriage

[29] Section 4.4 describes in detail the sample selection procedure.

[30] From appendix figure A6 shows a positive correlation between height and education for Indian men. Further, table A24 shows a labor market height premium for Indian men, with the semi-elasticity of height being 0.08%/cm comparable to what Case and Paxson (2008) find in the United States, but smaller than what Vogl (2014) finds in Mexico, ranging from 1.4% to 2.3%

market attributes, I estimate a transferable utility model of the market. In particular, the model helps me identify and estimate complementarities in height, education and parental education as well as cross-complementarities between them. The model is based on the framework of Dupuy and Galichon (2014). The results of the structural estimation will shed light on the features of men and women that are mutually attractive on the marriage market. Finally, using the model estimates, I will perform counterfactual simulations to study the link between between preferences for height on the marriage market and children's potential height and height for age z-score distribution.

4.1 Transferable Utility Model of the Marriage Market

Consider a frictionless transferable utility environment.[31] Let H_x, E_x, FE_x, ME_x and A_x denote the height, educational attainment, father's education and mother's education, and the age of a man in the marriage market. Analogously, let H_y, E_y, FE_y, ME_y and A_y denote the women's characteristics. Therefore, every man in the marriage market is characterized by a feature vector, $x = [H_x, E_x, FE_x, ME_x, A_x] \in \mathbb{R}^5$ and every woman on the marriage market is characterized by a feature vector , $y = [H_y, E_y, FE_y, ME_y, A_y] \in \mathbb{R}^5$. Let $P_x(.)$ and $Q_y(.)$ denote the probability distribution functions of x and y respectively, with $f_x(.)$ and $g_y(.)$ the corresponding probability density functions. A matching is defined as a probability density, $\pi(x, y)$ of observing a couple from the matched population, in which, the man has a feature vector x, and the woman has a feature vector y.

Since we are in the transferable utility environment, a primitive of the problem is the joint surplus function. Define $\Phi(x, y)$ as the joint utility generated when a man with feature vector x marries a woman with feature vector y. Further, let $\Phi(x,\phi)$ denote the utility of a man with feature vector x if he remains single and $\Phi(\phi, y)$ utility of a woman with feature vector y if she remains single. As a result, the marital surplus , $s(x, y)$ is defined as $\Phi(x, y)-\Phi(\phi, y)-\Phi(x,\phi)$. Since singles are not observed in the data, $\Phi(\phi, y)$ and $\Phi(x,\phi)$ are normalized to zero. The main objective of the model is to identify the complementarities between different components of the feature vector x and y, normalizing the utility from remaining single to zero does not impede this.

In the data, we observe two men with the same feature vector x marrying two different women, with different feature vector y; this implies that matching is also occurring along features unobserved to the researcher. This requires the introduction of unobserved heterogeneity into the framework. The seminal paper of Choo and Siow (2006) considers the case when the attributes of men and women are discrete and scalar (for example, education level or age). In their framework, when a man m, of type x marries a woman w, of type y, the total joint utility produced is $\Phi(x, y) + \epsilon_{m|y} + \epsilon_{w|x}$. Let k_1 be the number of types for a man, and k_2 be the number of types for a woman.[32] Let $\epsilon_m = [\epsilon_{m|y_1}, \epsilon_{m|y_2}, .., \epsilon_{m|y_{k_2}}]$ be a $k_2 \times 1$ dimension vector of idiosyncratic shocks , and $\epsilon_f = [\epsilon_{f|x_1}, \epsilon_{f|x_2}, .., \epsilon_{f|x_{k_1}}]$ be a $k_1 \times 1$ dimension vector of idiosyncratic shocks. $\epsilon_{m|y_k}$ is the idiosyncratic shock that a man m draws on matching with a woman y_k on the marriage market. This represents his idiosyncratic preference for a type y_k woman that the researcher doesn't observe. Analogously, $\epsilon_{f|x_k}$ is the idiosyncratic shock that a woman f draws on matching with a man x_k. Choo and Siow (2006) model ϵ_m and ϵ_f as independent and identically Gumbel distributions with scaling parameter $\frac{\sigma}{2}$. As a result, the joint utility, when a man of type x marries a woman of type y

[31] Therefore, we are assuming an environment in which utility can be transferred between partners at a constant rate.

[32] for example, if we consider types by education levels, and each man is either a college graduate or not, and similarly each woman is either a college graduate or not, then $k_1 = 2$ and $k_2 = 2$.

can be divided as $\Phi(x,y) = U(x,y) + V(x,y)$, where the utility of a man m of type x on matching with a woman w of type y is $U(x,y) + \epsilon_{m|y}$ and utility of a woman w of type y on matching with a man m of type x is $V(x,y) + \epsilon_{w|x}$. In this framework, the utility of an individual in equilibrium depends only on their and their potential partner's observable attributes.

In our context, individuals are matching on a multidimensional feature vector, in which some of the features like height are continuous variables. Dupuy and Galichon (2014) extend the framework of Choo and Siow (2006) to continuous multidimensional features. The approach is as follows: Each man, m of type x on the marriage market draws a random, infinite, but countable subset of potential partners. Each potential partner for a man m is defined by a feature vector, $[y_k^m, \epsilon_k^m] \in \mathbb{R}^4 \times \mathbb{R}$, where y_k^m is the vector of observable characteristics and ϵ_k^m is the random sympathy of a randomly drawn woman. Therefore, if a man m, with feature vector x marries a woman with feature vector y_k^m from his randomly drawn subset, he receives a utility $U(x, y_k^m) + \frac{\sigma}{2}\epsilon_k^m$. Each man, with a feature vector x, solves the following discrete choice problem:

$$\max_k \quad U(x, y_k^m) + \frac{\sigma}{2}\epsilon_k^m$$

Finally, we need to specify the random process according to which every man draws his set of potential partners. Assume that the vector $[y_k^m, \epsilon_k^m]$ is an enumeration of a Poisson process on $\mathbb{R}^4 \times \mathbb{R}$ with intensity $dy \times e^{-\epsilon} d\epsilon$. Analogously, each woman, w of type y draws a random subset of infinite but countable subset of potential partners. Each potential partner of a woman is defined by a feature vector,$[x_l^w, \eta_l^w] \in \mathbb{R}^4 \times \mathbb{R}$. If a woman with feature vector y marries a man with feature vector x_l^w from her randomly drawn subset, she receives a utility $V(y, x_l^w) + \frac{\sigma}{2}\eta_l^w$. Each woman, with a feature vector y solves the following discrete choice problem:

$$\max_l \quad U(x_l^w, y) + \frac{\sigma}{2}\eta_l^w$$

As in the case of men, the vector $[x_l^w, \eta_l^w]$ is assumed to be an enumeration of a Poisson process on $\mathbb{R}^4 \times \mathbb{R}$ with intensity $dx \times e^{-\eta} d\eta$. Dupuy and Galichon (2014) show that by assuming partners are drawn randomly from a Poisson process leads to the continuous multinomial logit choice model. Therefore, the probability of a man m with feature vector x, choosing a woman with feature vector y from his randomly drawn set is given by:

$$\pi(y|x) = \frac{e^{[U(x,y)/(\sigma/2)]}}{\int_{\mathbb{R}^4} e^{[U(x,t)/(\sigma/2)]} dt}.$$

Similarly, the probability of a woman, w with a feature vector y choosing a man with feature vector x from her randomly drawn set is given by [33]:

$$\pi(x|y) = \frac{e^{[V(t,y)/(\sigma/2)]}}{\int_{\mathbb{R}^4} e^{[V(t,y)/(\sigma/2)]} dt}.$$

This leads to the key equation, relating the equilibrium match probabilities, $\pi(x,y)$ and the joint utility $\Phi(x,y)$, given by:

$$\log \pi(x,y) = \frac{\Phi(x,y) - a(x) - b(y)}{\sigma}. \tag{1}$$

[33] See appendix section A.2 for the derivation.

The share of the joint utility allocated to the man is given by:

$$U(x,y) = \frac{\Phi(x,y) + a(x) - b(y)}{2},$$

and the share of the joint utility allocated to the woman is given by:

$$V(x,y) = \frac{\Phi(x,y) - a(x) + b(y)}{2}.$$

4.2 Identification of Complementarities

The identification question is as follows: Given the data on the equilibrium match distribution, can the joint utility function, $\Phi(x,y)$ be identified? From equation A1 and equation A2, we can write:

$$\begin{aligned} U(x,y) &= \frac{\sigma}{2}[\log(\pi(y|x) + c(x)] \\ V(x,y) &= \frac{\sigma}{2}[\log(\pi(y|x) + d(y)] \\ \Phi(x,y) &= U(x,y) + V(x,y) = \frac{\sigma}{2}[\log(\pi(y|x) + \log(\pi(y|x) + c(x) + d(y)], \end{aligned}$$

as a result, the joint utility is identified only up to unknown additive functions $c(x)$ and $d(y)$.[34] Therefore, the joint utility function, $\Phi(x,y)$ is not identified from the equilibrium match distribution without additional assumptions. The identified object is the cross-derivative of the joint utility function with respect to x and y. Taking the derivative of equation 1 with respect to x and y gives, $\frac{\partial^2\Phi(x,y)}{\partial x \partial y} = \frac{\partial^2 \log \pi(x,y)}{\partial x \partial y}$.[35] To conclude the identification argument, in the framework of Dupuy and Galichon (2014), the identified object is the cross-derivative of the joint utility function with respect to men's and women's marriage market features. The second derivative of the joint marital utility concerning x and y is precisely the object of interest as it captures the complementarities between any two marriage market features.

4.3 Estimation of Complementarities

Since the feature vectors of men and women in the marriage market consist of continuous variables like height, non-parametric estimation will not be feasible due to a significant missing data problem. Therefore, the joint utility, $\Phi(x,y)$, is parameterized as the following function:

$$\Phi(x,y) = x^T A y \tag{2}$$

$$\Phi(x,y) = \begin{bmatrix} H_x & E_x & FE_x & ME_x & A_x \end{bmatrix}^T \begin{bmatrix} \lambda_{HH} & \lambda_{HE} & \lambda_{HFE} & \lambda_{HME} & \lambda_{HA} \\ \lambda_{EH} & \lambda_{EE} & \lambda_{HFE} & \lambda_{HME} & \lambda_{HA} \\ \lambda_{FEH} & \lambda_{FEE} & \lambda_{FEFE} & \lambda_{FEME} & \lambda_{FEA} \\ \lambda_{MEH} & \lambda_{MEE} & \lambda_{MEFE} & \lambda_{MEME} & \lambda_{MEA} \\ \lambda_{AH} & \lambda_{AE} & \lambda_{AFE} & \lambda_{AME} & \lambda_{AA} \end{bmatrix} \begin{bmatrix} H_y \\ E_y \\ FE_y \\ ME_y \\ A_y \end{bmatrix} \tag{3}$$

, where matrix A is the critical component of interest known as the *affinity matrix*. From the previous section, the identified object is the cross-derivative of $\Phi(x,y)$ with respect to x and y. This object is pre-

[34] Where, $c(x) = \log(\int_{\mathbb{R}^4} e^{[U(x,t)/(\sigma/2)]} dt)$ and $d(y) = \log(\int_{\mathbb{R}^4} e^{[V(t,y)/(\sigma/2)]} dt)$.

[35] Note that $\frac{\partial^2\Phi(x,y)}{\partial x \partial y}$ is a Jacobian, since x and y are vectors.

cisely the affinity matrix. In this context, A is a 5×5 matrix. Recall that each man and each woman on the marriage market is characterized by a 5×1 feature vector, namely, height, education, father's education, mother's education and age. Each element of A, λ_{ij} represents the strength of attractiveness or repulsiveness between the man's feature x_i and the woman's feature y_j. λ_{ij} is the increase in the joint utility when the i^{th} element of man's feature vector x and the j^{th} element of the woman's feature vector is increased by one unit. Therefore, each of the λ_{ij} represents the complementarity or substitutability between two attributes, i and j in the joint utility function. Our main parameters of interest are λ_{HH},λ_{HE} and λ_{EH} representing the complementarities in men and women's height, cross-complementarities in men's height and women's education and cross-complementarities in men's education and women's height.

Galichon and Salanié (2020) and Dupuy and Galichon (2014) show that the equilibrium matching in the presence of unobserved heterogeneity can be written as the following:

$$Z(A,\sigma) = \max_{\pi \in P \times Q} \int\int_{\mathbb{R}^5 \times \mathbb{R}^5} x^T A y \pi(x,y) dx dy - \sigma \int\int_{\mathbb{R}^5 \times \mathbb{R}^5} \log \pi(x,y) \pi(x,y) dx dy. \tag{4}$$

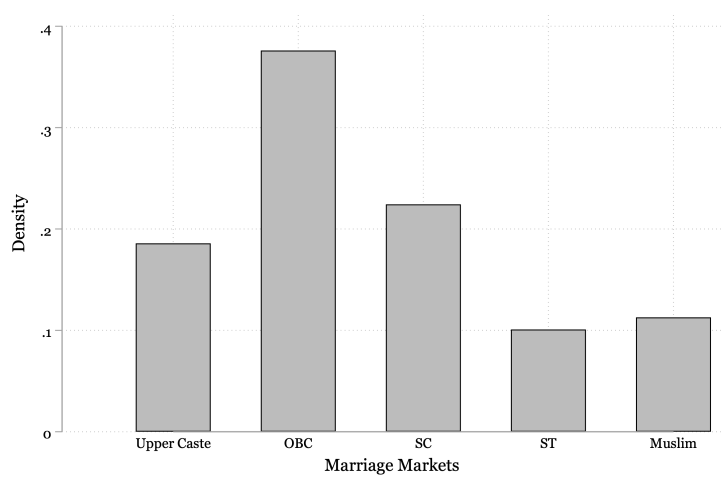
Intuitively, as explained in Galichon and Salanié (2020) the first term explains the part of the surplus that comes from the interaction between observable characteristics. In contrast, the second term represents the unobservable heterogeneity in taste. Maximizing equation 4 is equivalent to maximizing $Z(A/\sigma, 1)$, which leads to the following convex optimization problem, with $B = A/\sigma$:

$$\min_B Z(B,1) - \mathbb{E}_{\hat{\pi}}(x^T B y) \tag{5}$$

where, $\hat{\pi}$ is observed match probabilities in the data. This is a moment matching estimator, in which the moments predicted by the optimal matching function $\pi(x, y)$ are matched to the empirical moments observed in the data. Therefore, B is computed so that, $\mathbb{E}_{\hat{\pi}}[XY^T] = \mathbb{E}_{\pi}[XY^T]$. Since matrix A is identified up to a scale, I normalize σ to one.

4.4 Data

Figure 2: Marriage Market: Distribution



Note:

The Indian Human Development Survey(IHDS) is a nationally representative panel study containing detailed economic and demographic information. The first wave of the IHDS was collected in 2005-2006 and the second wave in 2011-2012. The key feature of the second wave of the IHDS(IHDS II) that makes it ideal for my analysis is that the IHDS II, as far as I know, is the only nationally representative data that contains information on the wife and the husband's height along with details about *both* marital and natal family.[36]

IHDS II is a multi-topic survey of 42,152 households in 1503 villages and 971 urban neighborhoods across India. For every individual in the household, anthropometric features, educational attainment, and age are measured; however, information regarding natal and marital family is measured for a subset of eligible women. In particular, for 39,523 eligible women, information regarding the women's parent's education and her husband's parent's education is measured in the number of years. The data also provides information regarding the natal family's wealth level relative to the marital family for eligible women. I use the father's education as a proxy for family wealth and family status and the mother's education as a proxy for the gender attitudes of the family (Dhar et al., 2019).

Precisely defining marriage markets to study the nature of sorting along various attributes is non-trivial (Chiappori et al., 2017; Ciscato and Weber, 2020). I define a marriage market with respect to a religion-caste group and men's year of the birth cohort. In particular, I define two cohorts; cohort one, which I refer to as the young cohort, consists of men born between 1970 and 1980 and their spouses born between 1970 and 1990, with the age gap ranging from zero to ten.[37] Cohort two, which I refer to as the old cohort, consists of men born between 1950 and 1960 and their spouses born between 1950 and 1970, with the age gap between them ranging from zero to thirteen.[38] Couples belong to one of the five religion-caste groups, namely, Scheduled Caste(SC), Scheduled Tribe(ST), Other Backward Caste(OBC), Upper Caste, or Muslims. Given the fact that marriages in India are primarily endogamous (Banerjee et al., 2013; Borker et al., 2017), I model each caste-religion group as a separate sub-marriage market. In particular, I define ten marriage markets, religion-caste-cohort specific. From the sample of 39,523 eligible women, I select a sample of couples with non-missing information on height, education, parental education, and age, belonging to one of the ten religion-caste-cohort-specific marriage markets. The final sample consists of 7355 couples.

From figure 2, 48% of the sample belongs to the Other Backward Caste, 22% belong to the marginalized Scheduled Caste, 10% belong to Scheduled Tribe, 19% belong to Upper Caste, and 11% are Muslims. Table A4 shows the summary statistics for the selected sample. From table A4, men's average height is 163 centimeters, and women's average height is 152 centimeters, with the average height gap between the husband and the wife being 11 centimeters. Men's median education is seven years, and women's is three years. The average education gap between couples is two years. Parental educational attainment in both marital and natal families is low. The median husband and wife's father and the median husband and wife's mother have zero years of education. In 17% of the couples, the natal family is wealthier than the marital family. Figure A1 shows the raw correlations between the husband and the wife's marriage market features. From A1(A), we see a positive correlation in the husband and the wife's height across religion-caste groups. Similarly, we also observe a positive correlation between the husband and the wife's

[36] The first wave of the IHDS has information on women's height, but information on men's height is missing for the majority of the sample.

The 2005 and 2015 wave of the National Family Health Survey(NFHS) has information on the wife and husband's education, height, and age but doesn't include information regarding the natal family. The 2006 wave of the Rural Economic and Demographic Survey (REDS) is a nationally representative data of *rural* India and also contains information regarding both natal and marital family along with information on height. I use the 2006 wave of REDS for secondary analysis.

[37] 99 percent of the couples belonging to the young cohort have an age gap between zero and ten. Therefore, the spouse's year of birth spans from 1970 to 1990.

[38] 99 percent of the couples belonging to the young cohort have an age gap between zero and thirteen. Therefore, the spouse's year of birth spans from 1950 to 1970.

education, the father's education, and the mother's education.

Next, I show the results for the estimated affinity matrix corresponding to equation 2 for religion-caste-cohort-specific marriage markets.

Table 2: Affinity Matrix Estimates: Other Backward Caste

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.305***	0.022	0.089***	-0.001	0.217***
	(0.03)	(0.036)	(0.034)	(0.033)	(0.051)
Education	0.105***	0.664***	0.14***	0.036	0.111**
	(0.038)	(0.048)	(0.042)	(0.043)	(0.065)
Father Education	0.007	0.063*	0.187***	0.006	0.074*
	(0.034)	(0.041)	(0.035)	(0.035)	(0.057)
Mother Education	0.011	0.195***	-0.048	0.161***	-0.043
	(0.034)	(0.044)	(0.034)	(0.028)	(0.058)
Age	0.159***	0.396***	-0.171	0.085*	2.63***
	(0.05)	(0.062)	(0.057)	(0.054)	(0.113)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Other Backward Caste, with men between the age of 30 and 40 and women between the age of 20 and 40.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table 3: Affinity Matrix Estimates: Upper Caste

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.31***	-0.033	0.087**	-0.044	0.071
	(0.048)	(0.065)	(0.047)	(0.041)	(0.076)
Education	0.081	0.975***	0.132**	0.043	0.082
	(0.064)	(0.098)	(0.064)	(0.06)	(0.103)
Father Education	0.041	0.033	0.283***	-0.038	0.046
	(0.05)	(0.07)	(0.047)	(0.041)	(0.08)
Mother Education	-0.061	0.274***	-0.009	0.107***	0.046
	(0.043)	(0.064)	(0.039)	(0.029)	(0.068)
Age	-0.045	0.262***	-0.159	0.131**	2.195***
	(0.074)	(0.103)	(0.075)	(0.062)	(0.152)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Upper Caste, with men between the age of 30 and 40 and women between the age of 20 and 40.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

5 Estimation Results

Table 2 shows the affinity matrix estimates for the biggest religion-caste group in the sample, Other Backward Caste.[39] The affinity matrix reveals several interesting results. First, we find strong complementarity in men's and women's height. A one standard deviation increase in men's height and a one

[39]The model is estimated for the young cohort and the old cohort separately. All the features are standardized so that they have a mean of zero and a standard deviation of one. Here I describe the results for the young cohort. Results for the old cohort are available in the appendix, table A8, A7, A9, A10 and A11.

Table 4: Affinity Matrix Estimates: Scheduled Tribe

Woman					
Man	Height	Education	Father Education	Mother Education	Age
Height	0.403***	-0.057	-0.012	0.07	-0.03
	(0.061)	(0.065)	(0.083)	(0.096)	(0.091)
Education	0.03	0.778***	0.145*	-0.257	-0.009
	(0.076)	(0.088)	(0.101)	(0.114)	(0.117)
Father Education	0.054	-0.019	0.252**	0.132	0.267**
	(0.099)	(0.11)	(0.113)	(0.113)	(0.152)
Mother Education	0.008	0.114	0.01	0.074	-0.3
	(0.099)	(0.11)	(0.108)	(0.097)	(0.15)
Age	-0.07	0.113	-0.166	0.258**	2.092***
	(0.085)	(0.095)	(0.122)	(0.131)	(0.165)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Scheduled Tribe, with men between the age of 30 and 40 and women between the age of 20 and 40.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table 5: Affinity Matrix Estimates: Scheduled Caste

Woman					
Man	Height	Education	Father Education	Mother Education	Age
Height	0.284***	0.082***	0.078***	-0.1	0.215***
	(0.037)	(0.041)	(0.045)	(0.057)	(0.056)
Education	0.018	0.686***	0.131***	0.084	0.12*
	(0.049)	(0.057)	(0.055)	(0.072)	(0.076)
Father Education	-0.022	0.191***	0.222***	-0.084	-0.012
	(0.046)	(0.052)	(0.046)	(0.061)	(0.07)
Mother Education	0.048	0.077	-0.028	0.267***	-0.056
	(0.06)	(0.071)	(0.057)	(0.051)	(0.091)
Age	0.036	0.225***	-0.137	0.116*	2.29***
	(0.061)	(0.069)	(0.073)	(0.089)	(0.119)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Scheduled Caste, with men between the age of 30 and 40 and women between the age of 20 and 40.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

standard deviation increase in women's height increases the joint utility by 0.31 units(statistically significant at 1%). This indicates that taller men and taller women find each other mutually attractive on the marriage market. Second, we find significant positive assortative matching concerning education; a one standard deviation increase in the husband's education and a one standard deviation increase in the wife's education increases the marital surplus by 0.664 units. Therefore, for the OBC marriage market, the complementarities in height are approximately half the complementarities in education. Third, strong complementarities are also observed concerning parental education. A one standard deviation increase in the wife's father's education and a one standard deviation increase in the husband's father's education increases the joint marital utility by 0.16 units. Similarly, a one standard deviation increase in the wife's mother's education and a one standard deviation increase in the husband's mother's education increases the joint marital utility by 0.13 units. Since the father's education represents family wealth, this result shows assortative matching on wealth in line with Borker et al. (2017), who document positive assortative matching along wealth for caste-specific marriage markets. Finally, complementarities on age have

Table 6: Affinity Matrix Estimates: Muslims

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.434***	-0.009	-0.081	0.101*	0.351***
	(0.059)	(0.068)	(0.071)	(0.074)	(0.09)
Education	0.106*	0.785***	0.105	0.064	0.049
	(0.075)	(0.091)	(0.091)	(0.101)	(0.117)
Father Education	-0.052	0.164**	0.41***	-0.082	-0.19
	(0.071)	(0.083)	(0.076)	(0.083)	(0.112)
Mother Education	-0.014	0.223***	-0.201	0.234***	0.091
	(0.076)	(0.087)	(0.072)	(0.061)	(0.115)
Age	-0.019	0.323***	-0.39	0.133	2.2***
	(0.086)	(0.102)	(0.107)	(0.111)	(0.169)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to Muslims, with men between the age of 30 and 40 and women between the age of 20 and 40.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

the most considerable magnitude. This finding primarily reflects the presence of multiple cohorts in the estimation sample and people marrying individuals from similar cohorts (Chiappori et al., 2020a).

A key point to note is the asymmetry of the affinity matrix. For example, the cross-complementarity in men's education and women's height is 0.105 and statistically significant at 1%, whereas the cross-complementarity in men's height and women's education is statistically insignificant. This result indicates that more educated men and taller women find each other mutually attractive on the marriage market. One of the possible channels driving this result is as follows: The anthropology literature in developing countries has also documented a positive relationship between reproductive success(lower infant mortality rates) and women's height due to the physiological benefits of being tall (Sear, 2006). This relationship becomes especially critical in high disease environments with poor health infrastructure, where the risk of infant mortality is high. In India, high rates of infant mortality have been linked to neighborhoods with poor sanitation and high disease prevalence (Geruso and Spears, 2018). Following Becker (1973), if we consider the joint utility in the transferable utility framework to represent child quality, which is inversely related to infant mortality, then we would expect taller women due to their higher reproductive success to be more preferred on the marriage market.

We also observe strong cross-complementarities in the men's height and the women's father's education, and in men's education and the women's father's education. This result indicates that wealthier natal families and taller and more educated men find each other more attractive on the marriage market. This finding is in line with Chiplunkar and Weaver (2019) who find that more educated men receive a more considerable dowry on the marriage market in India. We also find cross-complementarities in men's father's education and women's education(statistically significant at 10%); however, the cross-complementarity in men's father's education and women's height is statistically insignificant. Interestingly, we also observe cross-complementarity between men's mother's educational attainment and women's educational attainment. Since the mother's education potentially reflects the family's gender attitudes in line with Dhar et al. (2019), this result demonstrates that marital families with more progressive gender attitudes and women with higher educational levels find each other mutually attractive on the marriage market.[40]

[40] A well-documented feature when matching on height is the male-taller norm. The male-taller norm refers to the fact that women prefer their potential partners to be taller than them, and men, in turn, prefer their potential partners to be shorter than them. The male taller norm has been

Tables 4 and 5 show the affinity matrices for the marginalized and economically backward groups in India, the Scheduled Caste and the Scheduled Tribe.[41] Similar to the Other Backward Caste, we find positive assortative matching on height for both the groups. The estimated complementarities in height for the SC sample is 0.28, and for the ST sample is 0.4. For both the groups, we see cross-complementarities in men's education and women's father's education. For the Scheduled caste couples, wealthier marital families and more educated women are mutually attractive on the marriage market. Lastly, more educated and taller men and more affluent natal families also find each other more appealing.

Tables 3 shows the affinity matrix for the Upper Caste group. The estimated complementarities in height are 0.3. Results also indicate that taller and more educated men and wealthier natal families are mutually attractive on the marriage market. Similar to the results for the OBC couples, more educated women and gender progressive marital families are mutually attractive on the marriage market. The estimated coefficient for the cross-complementarity between men's education and women's height is positive but not statistically significant.

From table 6, for the Muslim marriage market, a one standard deviation increase in men's height and a one standard deviation increase in women's height increases the joint marital surplus by 0.403 units(statistically significant at 1%). We also observe positive assortative matching concerning education and parental education. Similar to the Other Backward Caste and Upper Caste couples, we also find cross-complementarities in men's education and women's height, implying more educated men and taller women see each other as mutually attractive. Finally, similar to the Other Backward Caste couples, marital families with gender progressive attitudes and more educated women find each other mutually appealing on the marriage market.

The affinity matrix helps understand the degree of complementarity/ substitutability between various marriage market features. Next, quantify the relative importance of each of these features in explaining the marital surplus using the saliency analysis technique developed by Dupuy and Galichon (2014).[42] From table A14 and A15 for the largest religion-caste-specific marriage market, height explains approximately 8% of the variation in the joint marital surplus; in comparison, education explains about 16% of the joint marital surplus. Therefore, relative to education, height explains about half the variation in the joint marital utility. Similarly, from table A18,A19, A20 and, A21 for the Scheduled Tribe and Muslim couples, height explains about 11% of the observed joint marital utility, which is about half the variation explained by education.[43]

In the appendix, I perform several robustness checks to verify my main results. First, I re-estimate the model using the REDS,2006 dataset and show that the results regarding the strong complementarities in height are qualitatively unchanged. The REDS, 2006 is a nationally representative sample of rural India. Unlike the IHDS, the REDS 2006 measures family wealth in terms of family landholdings.[44] Table A12

documented in several developed countries(Gillis and Avis, 1980; Courtiol et al., 2010). In our sample, we see a strong adherence to the male taller norm. In 95% of the couples, the husband is taller than the man. One of the potential channels leading to the significant positive assortative matching on height is the male-taller-norm. However, the strong cross-complementarities between men's height and natal family wealth and the presence of cross-complementarities between men's education and wife's height indicates a role for stature on the marriage market beyond the male-taller-norm. As a result, it is unlikely that the results are being driven solely by adherence to the male-taller norm.

[41] According to the 2011 census, Scheduled Caste and Scheduled Tribes comprise about 16.6% and 8.6%, respectively, of India's population. Article 341 of the Indian constitution allows the government to define which communities fall under Scheduled Castes. This community historically faced severe discrimination and was subjected to the practice of untouchability.

Similarly, Scheduled Tribes are the other marginalized committee in India characterized by a high degree of poverty. Article 366 of the Indian constitution precisely defines which communities fall under Scheduled Caste and Scheduled Tribe.

[42] The procedure is detailed in the appendix, section A.1.

[43] A16,A17, A22 and, A23 show the saliency analysis results for the Scheduled Caste and Upper Caste couples respectively.

[44] Information regarding the husband and the wife's height and natal and marital family wealth is available only for the household head and his spouse. Data regarding family wealth is missing for a significant portion of the sample; as a result, the sample size of REDS, 2006 is significantly smaller

and A13 show the results for the Other Backward Caste and the Upper Caste sample.

Second, I re-estimate the model for individuals living in rural and urban India separately and show that the results regarding the strong complementarities in height are qualitatively unchanged.[45] Ideally, we would estimate our model concerning religion-caste groups living in a particular district, since a large portion of marriages in India take place within a district, with the average travel time between marital and the natal family home being three hours (Rosenzweig and Stark, 1989; Fulford et al., 2013; Beauchamp et al., 2017). However, the sample size prevents estimation at the district level.

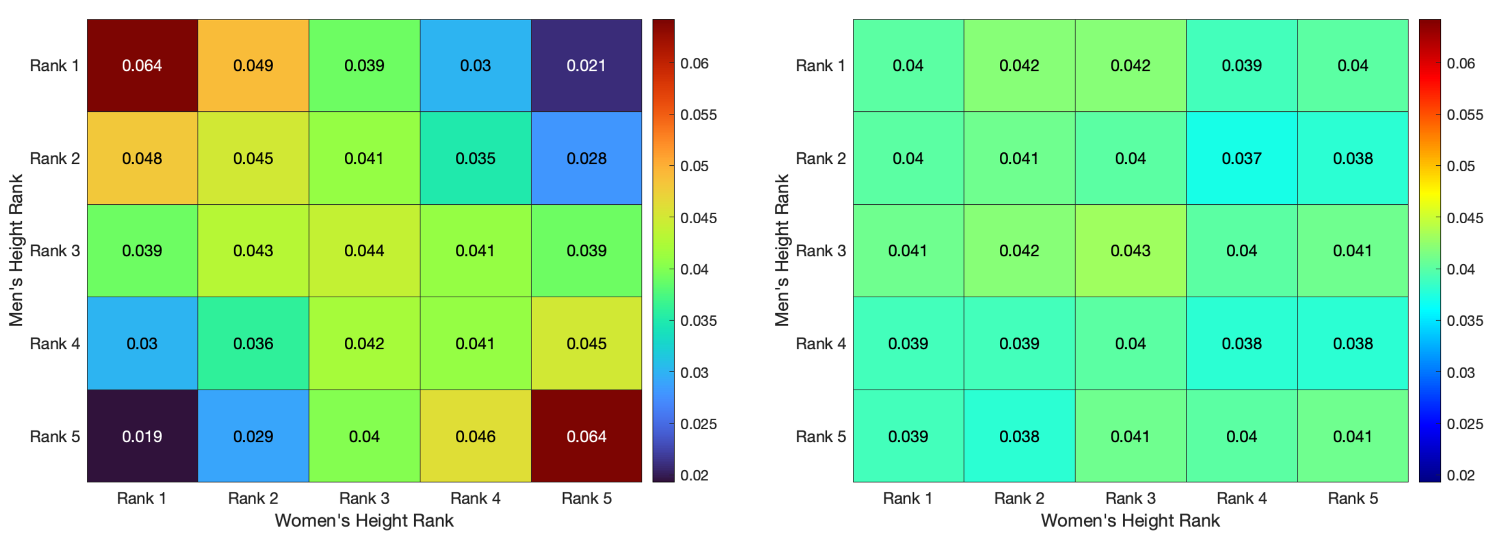
6 Counterfactual Simulations

Using the affinity matrices across religion-caste groups, I simulate various hypothetical preferences for height in the marriage market and compute the couple's joint height distribution. Next, I study how complementarities in height have evolved in India and compute the hypothetical joint height distribution if preferences for height remained unchanged over time.

To understand the implications of matching on height, I rely on the medical literature, which has shown that height is a polygenic trait and is subject to genetic and environmental influences. I explain through my counterfactual simulations how one generation's preferences for height can influence the next generation's height distribution.

6.1 No Complementarities in Height

Figure 3: Matching on Height(Other Backward Caste): No Complementarities



(A) Predicted Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there were no complementarities in height in the marriage market. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

I simulate a hypothetical marriage market in which men and women show no complementarity in height, contrary to what the results in section 5 have shown. I set the parameter capturing the complementarity in height to zero in each religion-caste-specific marriage market and estimate the counterfactual

than the IHDS.

[45] Results not shown for brevity, but are available upon request

equilibrium match distribution. This corresponds to setting λ_{HH} in equation 2 to zero. Note that all other elements of the affinity matrix are unchanged. I plot the couple's joint height distribution under this hypothetical scenario. To illustrate the results more simply, I divide men and women into five height ranks. Rank 1 represents men(women) in the top 20th percentile of the men's(women's) height distribution; similarly, rank 2 represents 20th-40th percentile, rank three 40th-60th percentile, rank four 60th-80th percentile, and rank five the top 20th percentile.

Figure 3 shows the predicted rank-rank matching versus the counterfactual rank-rank matching for the largest religious-caste group in the sample, the Other Backward Caste. From figure 3(A), under the actual equilibrium matching, the probability of a height rank one woman marrying a height rank one man, and the probability of a rank five-man marrying a rank five-woman is six percent. On the other hand, the likelihood of a rank one man marrying a rank five woman and vice-versa is two percent. From figure 3(B), when the complementarity in height parameter is set to zero, the probability of a height rank five-man marrying a height rank five-woman and a height rank one-man marrying a height rank one-woman drops to four percent. Similarly, the likelihood of a rank one man marrying a rank-five woman and vice-versa increases by approximately two percentage points. Overall when the complementarity in the height preference parameter is zero, the couple's joint height distribution is approximately uniform.

In the appendix, figures A4, A5, A6 and A7 show similar results for the remaining religion-caste-specific marriage markets. Further, in the appendix, I simulate a similar counterfactual, wherein I set the complementarities in height parameter equal to the complementarities in age parameter. The affinity matrix estimate corresponding to age is the biggest in magnitude. Therefore, this counterfactual is equivalent to simulating a marriage market with an increase in complementarities in height. Figures A8, A9, A10, A11 and A12 show the results across religion-caste-specific marriage markets. The figures show a substantial increase in positive assortative matching on height. Overall, the results from the two counterfactuals illustrate the role of preferences for height on the marriage market in determining marital sorting on height.

6.2 Preferences for Height Unchanged Over Time

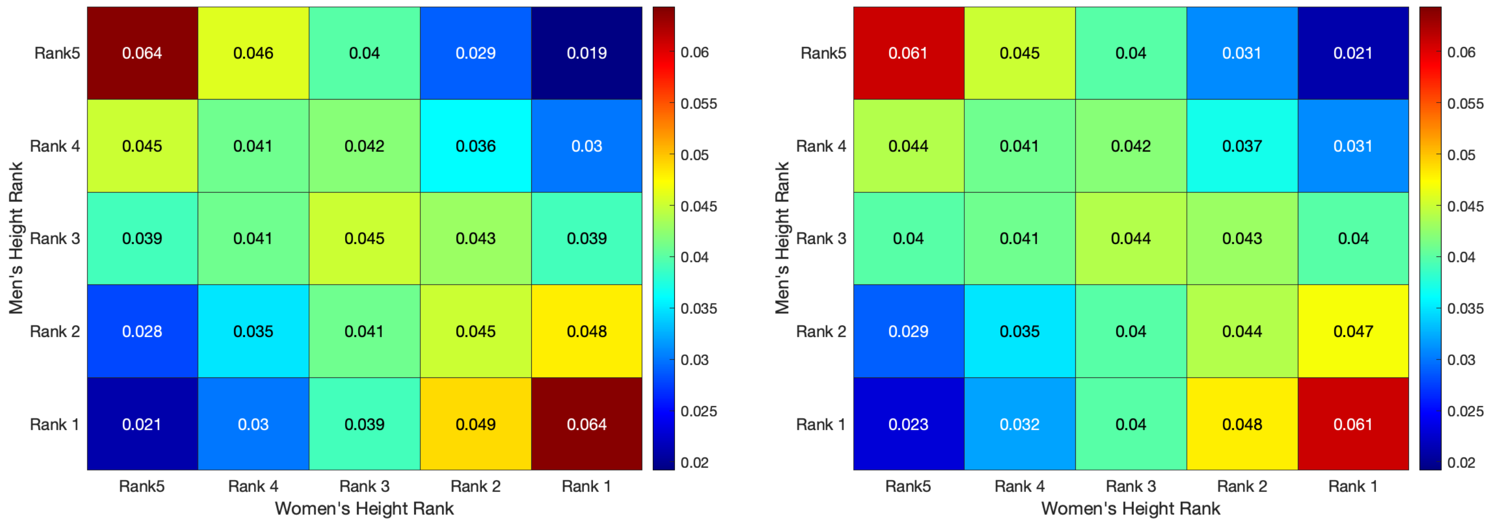
Table 7: Changes in Complementarity in Height: Comparison Tests

	Complementarity in Height Young Cohort	Complementarity in Height Old Cohort	Change in Complementarity	t-statistic(p-value)
Upper Caste	0.31	0.31	-0.0005	-0.0079(0.496)
Other Backward Caste	0.305	0.264	0.038	0.865(0.194)
Scheduled Tribe	0.403	0.232	0.167^{*}	1.5(0.067)
Muslims	0.434	0.272	0.157^{**}	1.69(0.045)
Scheduled Caste	0.284	0.319	-0.033	-0.529(0.298)

Note: The old cohort consists of men born between 1970 and 1980 and their spouses born between 1970 and 1990, with the age gap ranging from zero to ten. Cohort two, which I refer to as the old cohort, consists of men born between 1950 and 1960 and their spouses born between 1950 and 1970, with the age gap between them ranging from zero to thirteen. Standard errors for the difference in complementarity are computed using bootstrap with 500 draws. * $p < 0.1$, ** $p < .05$, *** $p < 0.01$.

The height of Indians irrespective of gender, caste, or religious group has remained relatively constant (Deaton, 2007). To emphasize this puzzling fact, figure A25 compares the trajectory in the development of education and height for three developing countries, China, Mexico, India, and one developed country, the United States. The trend in the educational outcome as measured by the number of primary age children enrolled in primary school follows a similar increasing trajectory across the three developing countries.

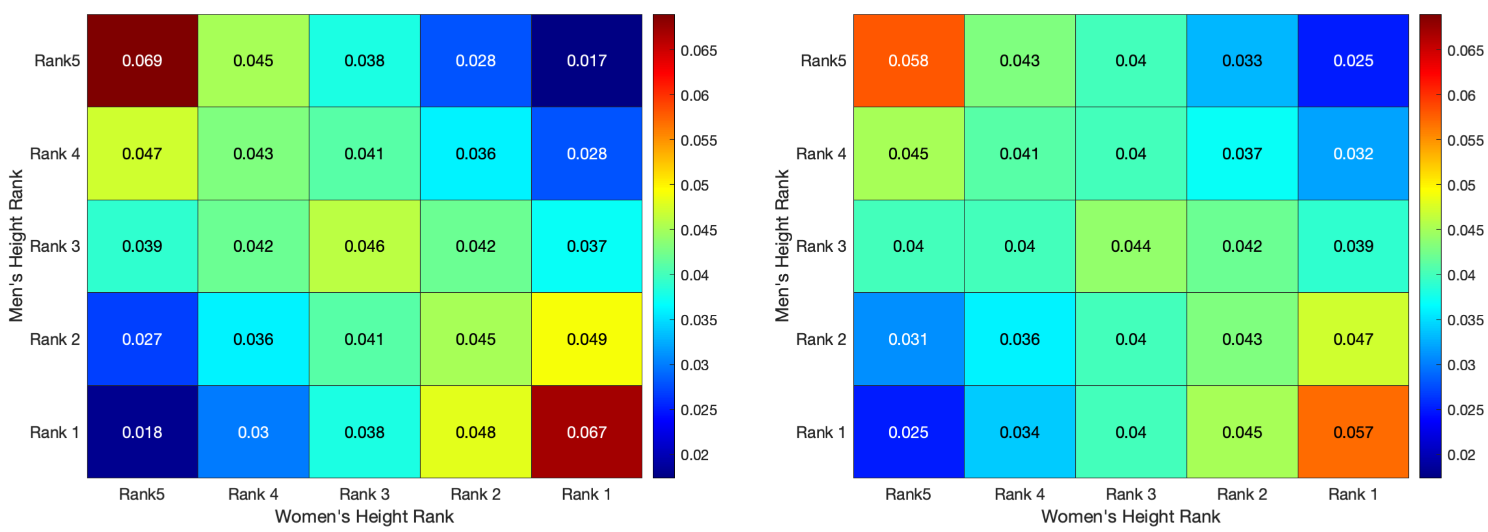
Figure 4: Joint Height Distribution(Other Backward Caste): Preferences Unchanged Over Time



(A) 1970-1980 Preferences(Actual) (B) 1950-1960 Preferences(Counterfactual)

Note: Figure (A) shows actual rank-rank matching for the Other Backward Caste couples belonging to the young cohort. Figure(B) indicates the counterfactual rank-rank matching for couples belonging to the young cohort by imposing the complementarity in height parameter corresponding to the old cohort.

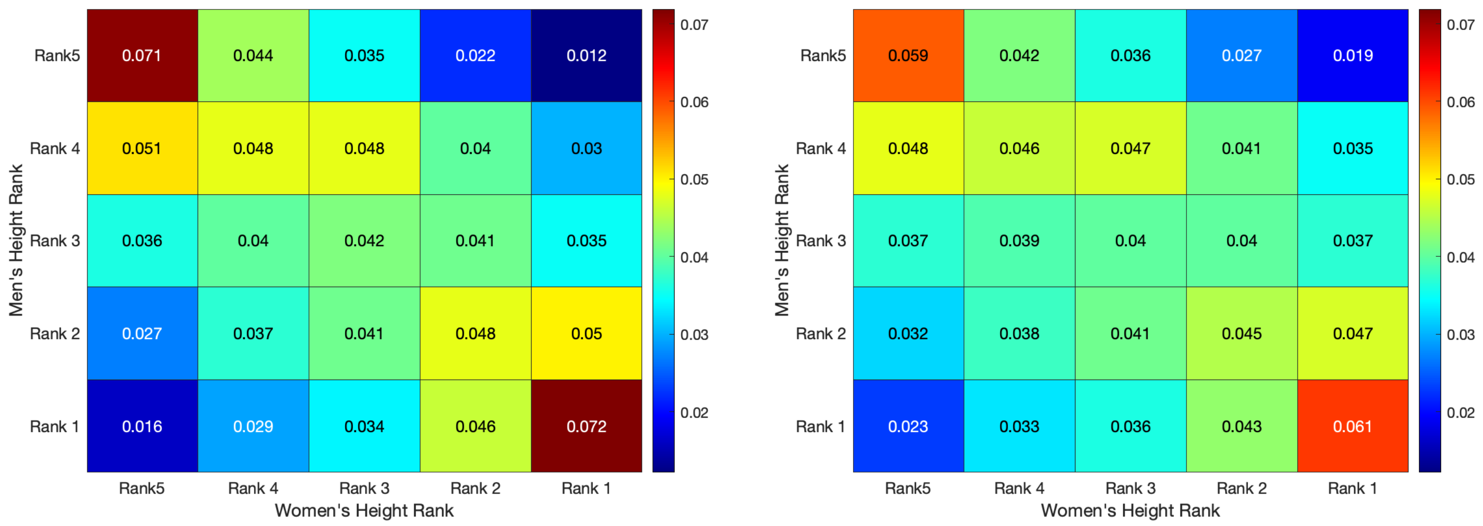
Figure 5: Joint Height Distribution(Scheduled Tribe): Preferences Unchanged Over Time



(A) 1970-1980 Preferences(Actual) (B) 1950-1960 Preferences(Counterfactual)

Note: Figure (A) shows actual rank-rank matching for the Scheduled Tribe couples belonging to the young cohort. Figure(B) indicates the counterfactual rank-rank matching for couples belonging to the young cohort by imposing the complementarity in height parameter corresponding to the old cohort.

Figure 6: Joint Height Distribution(Muslims): Preferences Unchanged Over Time



(A) 1970-1980 Preferences(Actual) (B) 1950-1960 Preferences(Counterfactual)

Note: Figure (A) shows actual rank-rank matching for the Muslim couples belonging to the young cohort. Figure(B) indicates the counterfactual rank-rank matching for couples belonging to the young cohort by imposing the complementarity in height parameter corresponding to the old cohort.

However, when comparing China and Brazil to India, the trend in height paints a different picture. The height of Chinese and Brazilian men has increased by five centimeters between 1940 and 1990, whereas the height of Indian men has increased by only two centimeters. Similarly, the height of Chinese and Brazilian women has risen by four centimeters between 1940 and 1990, whereas the height of Indian women has grown by only two centimeters. One of the key factors driving this result is the differential evolution in the disease environment and health infrastructure across the three countries. High rates of open defecation in India have been linked to child stunting (Spears, 2013; Coffey and Spears, 2017).

Given the evolution of height in India, I first study how preferences for height have evolved in India across religion-caste groups; second, I estimate couples' counterfactual joint height distribution if preferences for height had remained unchanged over time. More precisely, I compare the complementarity parameter in the affinity matrix corresponding to height for the old cohort and the young cohort.[46] Then I impose the height complementarity parameter of the old cohort onto the young cohort's affinity matrix and estimate the counterfactual joint height distribution.

Table 7 shows the complementarity in height parameter for the old cohort and the young cohort, along with the change in the complementarity in height across religion-caste-specific marriage markets. The most significant increase in complementarity in height is observed for the Scheduled Tribe, followed by Muslim couples. In other words, considering the changes in the marginal distribution of men and women's height, taller men and taller women belonging to the young cohort find each other more attractive on the marriage market than taller men and taller women belonging to the old cohort. We see a very marginal increase in the complementarities in height for the Other Backward Caste couples but the difference is not statistically significant. Complementarities in height show no significant change for the Upper Caste and the Scheduled Caste couples. These results indicate a considerable degree of heterogeneity across religion-caste groups in the evolution of complementarities for height in the marriage market in India.

Next, I estimate the couple's counterfactual joint height distribution under the hypothetical scenario

[46] Old and young cohort are precisely defined in the 4.4 section. Full set of affinity matrix estimates for the young cohort are available in table 2 - 6 and full set of affinity matrix estimates for the old cohort are available in table A7 - A11.

that the complementarities in height are unchanged over time. Similar to the previous counterfactual simulation, I divide men and women into five height ranks and compare the actual rank-rank matching on height for the young cohort to the counterfactual rank-rank matching by imposing the complementarities in the height of the old cohort.

Figure 4 shows the results for the Other Backward Caste marriage market. From table 7, complementarities in height have increased only marginally, and this is reflected in the results of the counterfactual simulation. The joint height distribution of couples belonging to the young cohort shows relatively no change when we impose the height preferences of the old cohort. The probability of a height rank-one man marrying a height rank-one woman and a rank five-man marrying a rank five-woman remains constant at around 6%. Similarly, the likelihood of a height rank-one man marrying a height rank- five woman and vice-versa remains stable at around 2%. Therefore we do not see a substantial change in the degree of assortative matching when the complementarity in the height of the old cohort is imposed on the young cohort for the Other Backward Caste couples.

From table 7 the Scheduled Tribe and the Muslim couples show the largest increase in complementarity on height over time. Figure 5 and 6 show the results from the counterfactual simulation for these two groups. From figure 5, the counterfactual rank-rank matching indicates that the probability of a height rank-one marrying a height rank-one woman and a height rank five-man marrying a height rank five-woman falls by one percentage point. Similarly, we see an increase in the likelihood of a height rank-one man marrying a height rank- five woman and vice-versa. The counterfactual simulation indicates that the degree of positive assortative matching on height decreases when the complementarity in height parameter of the older cohort is imposed on the young cohort for the Scheduled Tribe couples. Similarly, from figure 6, we see a change in the degree of positive assortative matching on height for the Muslim couples when we compute the counterfactual rank-rank matching by imposing the height complementarity parameter of the older cohort.

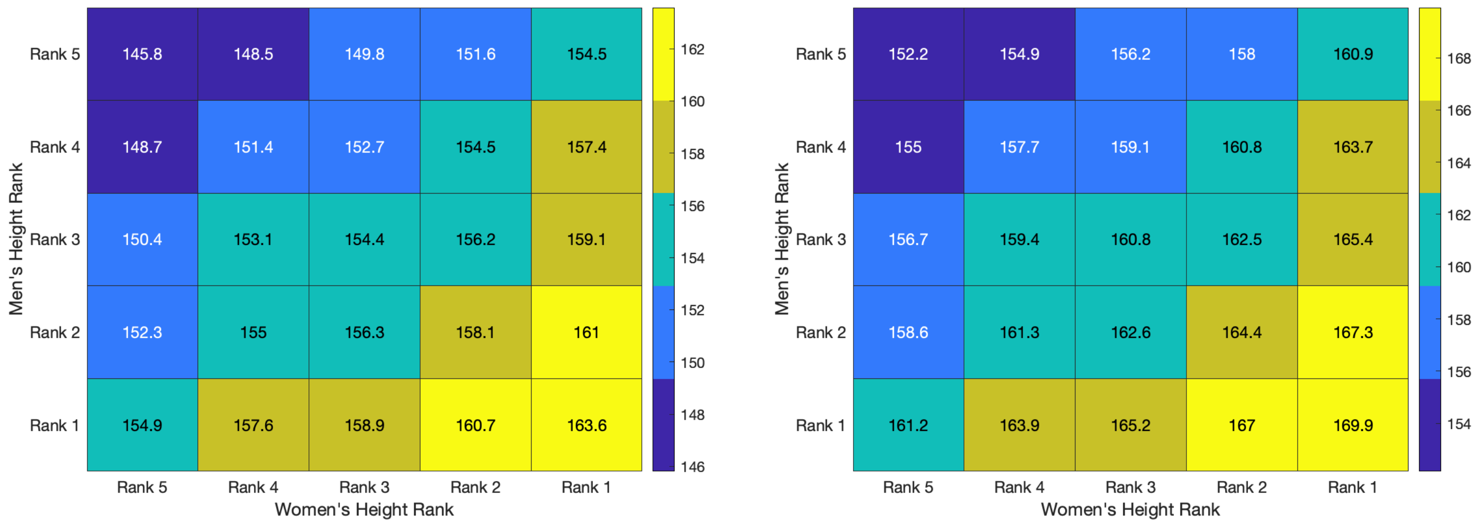
6.3 Children's Potential Height Distribution

Height is a polygenic trait that depends on both genetic and environmental factors. The medical literature suggests that height is a highly heritable trait, with around 80% of the variation in height explained by genetics (Stulp and Barrett, 2016). Pediatricians dating back to the 1800s have calculated a child's target height as a function of parental height. The earliest example of this is the work of Galton (1886) who argued that mid-parental height(average of the mother's and father's height) best represents the transmission of stature from parents to children. Mid-parental height is still considered a strong predictor of children's adult height (Cole, 2000). A modification to mid-parental height by sex introduced by Tanner et al. (1970), known as the Tanner method, is a popular technique used by pediatricians today to predict a child's height potential. According to the Tanner method, son's target height or potential height is given by[47]:

$$(\text{Father Height} + \text{Mother Height} + (\text{mean parental height}))/2,$$

[47]There are several modified versions of Tanner's method, each designed to improve the accuracy of the potential height measure, I do not go into the details of various adjustments to Tanner's method(see Cole (2000) for details), as it's beyond the scope of the paper. I use the original version's of Tanner's method for my analysis.

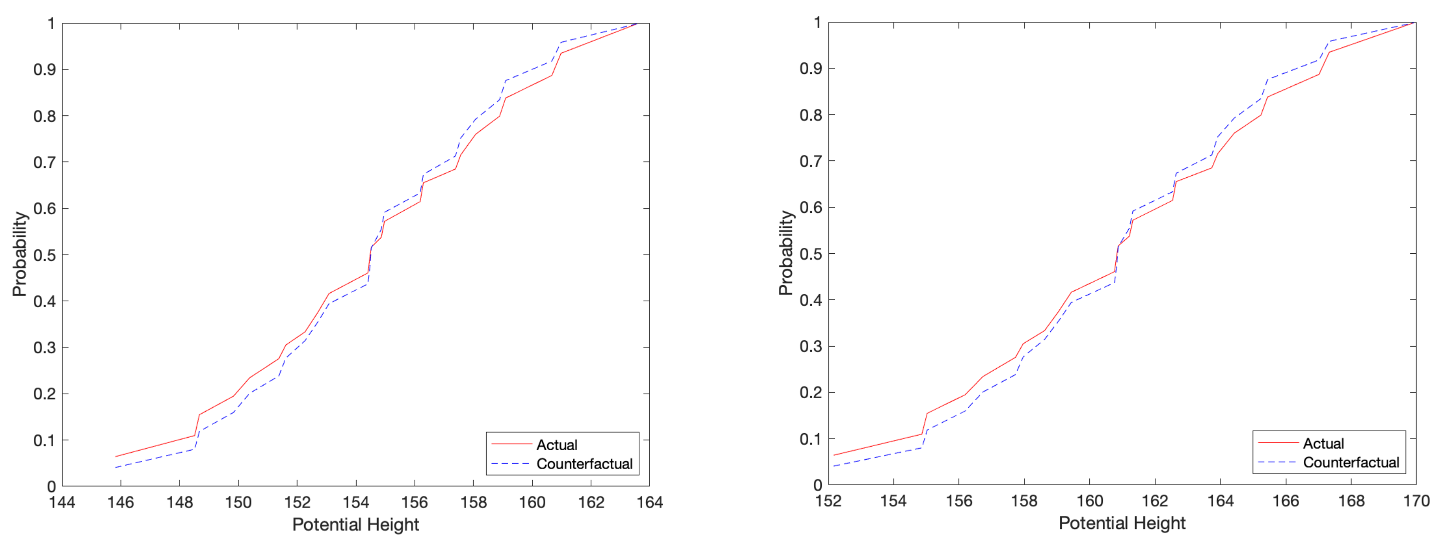
Figure 7: Children's Potential Height:Other Backward Caste



(A) Girl's Potential Height (B) Boy's Potential Height

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1970).

Figure 8: Children's Potential Height Distribution(Other Backward Caste): Actual vs Counterfactual



(A) Girl's Potential Distribution (B) Boy's Potential Distribution

Note: Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1970). Counterfactual refers to the hypothetical scenario with no complementarities in height on the marriage market.

and a daughter's target height is given by:

$$(\text{Father Height} + \text{Mother Height} - (\text{mean parental height}))/2,$$

A crucial caveat in developing countries like India is that Tanner's method can underestimate a child's height potential (Atluri, Bharathidasan, and Sarathi, 2018). However, even in India, mid-parental height plays a vital role in predicting a child's adult height; according to the WHO Multicentre Growth Reference Study (Garza et al., 2013), mid-parental height explains about 21% of the variability in linear growth from birth to 2 years of children in India.

Figure 7 shows the potential height distribution for the children of Other Backward Caste couples by individual height rank calculated using the Tanner method.[48] The average potential height for a girl(boy) with height rank-one father and height rank-one mother is 164(170) centimeters, whereas the average potential height of a girl(boy) with height rank-five father and height rank-five mother is 146(152) centimeters. Figure 7 demonstrates how matching on height can impact the children's potential height distribution.

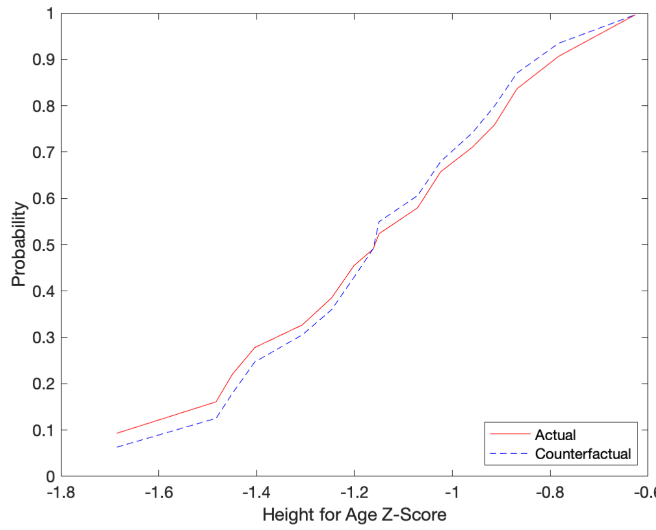
Next, I analyze the impact of changes in the preferences for height on the marriage market on the girl's and boy's potential height distribution. I simulate a hypothetical marriage market with no complementarity in height on the marriage market and compute the potential height cumulative distribution for boys and girls. Figure 8 shows the results from this counterfactual simulation for the Other Backward Caste couples. The results indicate that preferences for height in the marriage market impact the distribution of children's potential height. Although the average potential height of boys and girls remains unchanged, we a shift in the distribution, such that the prevalence of short and tall children decreases. For example, from figure 8(A), under the true equilibrium matching, the probability of a girl's target height is less than 150 centimeters is six percent; this probability falls to four percent when there are no complementarities in height on the marriage market. Similarly, the probability of a girl's target height being greater than 160 centimeters falls by four percent. The standard deviation of the girl's potential height distribution decreases by four percent. This implies that the inequality in potential height decreases as the complementarity in height decreases. An analogous result is observed for boys potential height in figure 8(B), where the probability of a boy's target height being less than 155 centimeters decreases by 4% and the probability of a boy's target height being greater than 165 centimeters also falls by 4%. I simulate the effect of no complementarities in height on children's potential height distribution for other religion-caste groups in the appendix, figures A14, A20, A18 and A16 show the results.

To further illustrate the impact of complementarity in height on the potential height distribution, I simulate an alternate counterfactual, wherein I set the complementarity in height parameter equal to the complementarity in age and compute the hypothetical potential height distribution. Figure A21 compares the actual potential height distribution versus the counterfactual potential height distribution for boys and figure A22 shows the results for girls. We again see significant distributional effects; as complementarity in height increases, the likelihood of children of small stature goes up but so does the likelihood of children of tall stature. The standard deviation of the distribution of potential height under the counterfactual preferences goes up by approximately 3%, which indicates an increase in the height inequality within the religion-caste group.

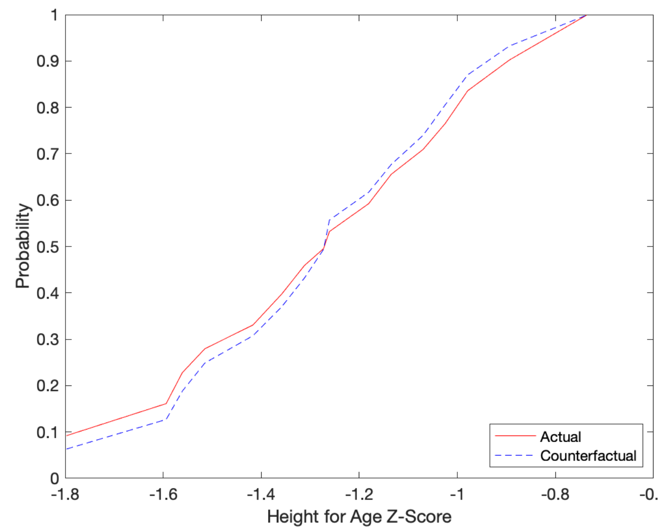
[48] Figures A13, A17, A19 and A15 show the potential distribution for the remaining religion-caste groups

6.4 Children's Height for Age Z-Score

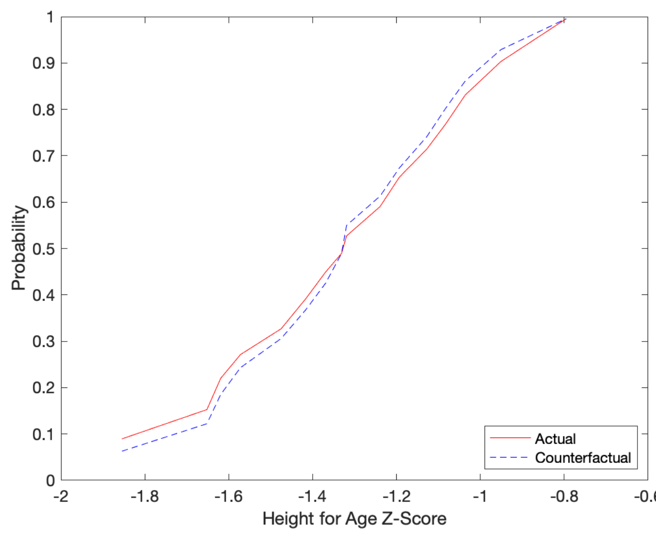
Figure 9: Children's Height For Age Z-Score Distribution: Actual vs Counterfactual



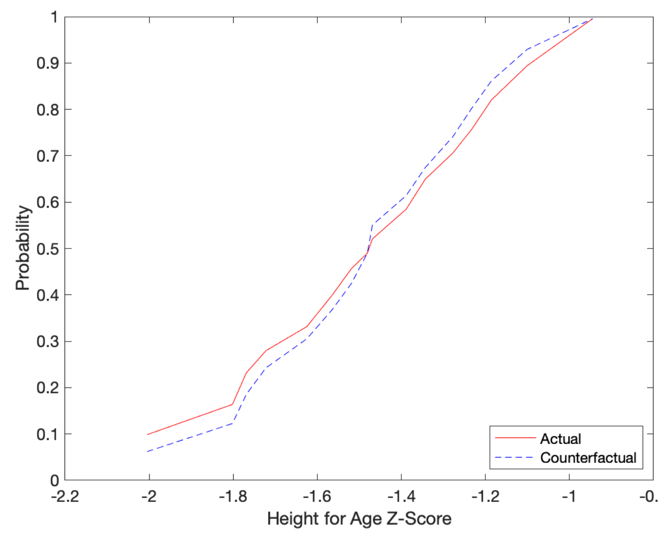
(A) Upper Caste



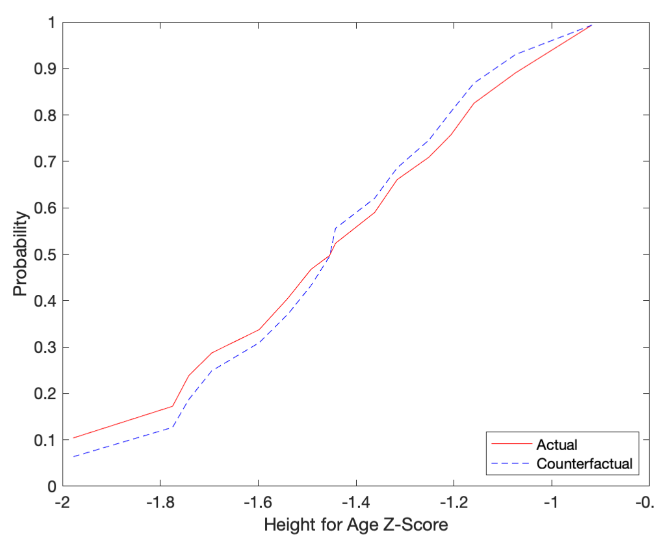
(B) Other Backward Caste



(B) Scheduled Caste



(B) Scheduled Tribe



(B) Muslims

Note: X-axis is the predicted height for age z-score. Counterfactual preferences refer to the hypothetical scenario, in which there are no complementarities in height in the marriage market. Height for age z-score calculated for children between the age of 0 and 19 using the WHO Child Growth Charts.

To further assess the impact of preferences for height in the marriage market on children's stunting hazard, I compare children's height for age z-score distribution under various hypothetical preferences for height to the actual distribution. The analysis involves two steps. In the first step, I run a linear regression of parental height rank on children's height for age z-score, conditional on parental education, parental age, household expenditure, whether the household has piped water, the mother's parent's literacy, the father's parent's literacy, child's gender, along with district fixed effects, religion-caste group fixed effects and age fixed effects. Using the estimates from the regression, I obtain predicted height for age z-score conditional on other parental characteristics and household-level variables for every combination of parental height rank. In the second step, I compute the occurrence of every rank-rank matching under the actual preferences for height and the hypothetical preferences(no complementarities in height in the marriage market). This gives me the final distribution of children's height for age z-score under the actual and hypothetical preferences for height.

Figure 9 shows the impact of complementarities in height in the marriage market on children's height-for-age cumulative distribution across religion-caste groups. From figure 9, although complementarities in height in the marriage market do not impact the average height-for-age, they significantly impact the children's height for age z-score distribution. Under the hypothetical preferences, with no complementarities in height in the marriage market, we find that the standard deviation of the height for age z-score distribution decreases across religion-caste groups by 3 percent. The shift in the cumulative distribution function under the counterfactual reduces the likelihood of children with small height-for-age, as well as the likelihood of tall children. In particular, for the Scheduled Tribe group, we find that complementarities in height in the marriage market can increase the prevalence of stunting (height for age z-score of less than -2) in children by 4 percentage points. To assess the magnitude of the coefficient, Jayachandran and Pande (2017) find that due to strong elder son preference in India, relative to their African counterparts, the hazard of stunting is 5 percentage points and 6 percentage points higher for second birth order and third birth order children with respect to the oldest sibling, respectively.

7 Conclusion

Height plays a crucial role in determining who marries whom in India, given the vital link between parental height and children's height and the long term impacts of height on individual well-being, this paper structurally estimates preferences for height in the Indian marriage market and studies the influence of marital preferences for height on children's height distribution.

Using detailed data on the husband and the wife's individual as well as family characteristics, I estimate a two-sided transferable utility model of the marriage market utilizing the framework of Dupuy and Galichon (2014). I find strong complementarity in men's and women's height across religion-caste groups. Results from the biggest caste group in the data, the Other Backward Caste, indicate significant cross-complementarity in men's education and women's height but not vice-versa, implying that more educated men and taller women find each other mutually attractive on the marriage market. Similarly, I also find strong cross-complementarity between men's height and women's family wealth, implying more educated men and women from wealthier families find each other mutually attractive on the marriage market. Next, I analyze the changes in matching and sorting patterns in height over time; overall I find a mild increase in complementarity in height with substantial heterogeneity across religion-caste groups.

Using the model estimates, I simulate hypothetical marriage markets with counterfactual preferences for height, and compute children's potential height (growth potential) as a function of parental height using a well-known procedure from the medical literature known as Tanner's method (Tanner et al., 1970). Results from the counterfactual simulation suggest that preferences for height significantly impact children's potential height distribution. In particular, I find that complementarities in height increase the standard deviation of the potential height distribution by 3%. Similarly, counterfactual simulations indicate that complementarities in height in the marriage market impact children's height for age z-score distribution. Results from the simulations suggest that although changes in the complementarity do not affect the average height-for-age, they impact the inequality in children's height through their effect on the height distribution. Results also indicate that complementarities in height in the marriage market can increase the hazard of stunting by up to 4 percentage points, comparable to the effect of strong son preference on stunting among lower birth order children in India (Jayachandran and Pande, 2017).

There are several directions along which this paper can be expanded; first, the current framework allows me to estimate complementarities in height but doesn't allow me to estimate preferences concerning height for men and women separately. This could be achieved using a survey design as in Chiappori et al. (2021) to estimate the willingness to pay for height in terms of family income for men and women separately. Second, to explicitly include marital payments, dowries into the framework and study the relationship between dowries and height. In particular, answering the following question: What is the marriage market premium in terms of dowry received for men with respect to height? Future research will attempt to answer these questions.

References

ADAMS, A. AND A. ANDREW (2019): "Preferences and beliefs in the marriage market for young brides," Tech. rep., IFS Working Papers.

ANUKRITI, S. AND S. DASGUPTA (2017): "Marriage markets in developing countries," in *The Oxford Handbook of Women and the Economy*.

ASHRAF, N., N. BAU, N. NUNN, AND A. VOENA (2020): "Bride price and female education," *Journal of Political Economy*, 128, 591–641.

ATLURI, S., K. BHARATHIDASAN, AND V. SARATHI (2018): "Tanner's target height formula underestimates final height in Asian Indians–A cross-sectional observational study," *Indian journal of endocrinology and metabolism*, 22, 441.

BANERJEE, A., E. DUFLO, M. GHATAK, AND J. LAFORTUNE (2013): "Marry for what? Caste and mate selection in modern India," *American Economic Journal: Microeconomics*, 5, 33–72.

BEAUCHAMP, A., R. CALVI, AND S. FULFORD (2017): "Terms of Engagement: Marriage and Migration in India," in *PAA 2017 Annual Meeting*, PAA.

BECKER, G. S. (1973): "A theory of marriage: Part I," *Journal of Political economy*, 81, 813–846.

BORKER, G., J. EECKHOUT, N. LUKE, S. MINZ, K. MUNSHI, AND S. SWAMINATHAN (2017): "Wealth, marriage, and sex selection," Tech. rep., Technical report, Working Paper, Cambridge University.

BOULIER, B. L. AND M. R. ROSENZWEIG (1984): "Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior," *Journal of political Economy*, 92, 712–732.

BOXHO, C., A. DONALD, M. GOLDSTEIN, J. MONTALVAO, AND L. ROUANET (2020): "Assortative matching in Africa: Evidence from rural Mozambique, Côte d'Ivoire, and Malawi," *Economics Letters*, 187, 108924.

CALVI, R., A. LEWBEL, AND D. TOMMASI (2021): "LATE with Missing or Mismeasured Treatment," .

CASE, A. AND C. PAXSON (2008): "Stature and status: Height, ability, and labor market outcomes," *Journal of political Economy*, 116, 499–532.

CHIAPPORI, P.-A. (2020): "The theory and empirics of the marriage market," *Annual Review of Economics*, 12, 547–578.

CHIAPPORI, P.-A., E. CISCATO, C. GUERRIERO, ET AL. (2020a): "Analyzing matching patterns in marriage: theory and application to Italian data," Tech. rep.

CHIAPPORI, P.-A., M. COSTA-DIAS, S. CROSSMAN, AND C. MEGHIR (2020b): "Changes in assortative matching and inequality in income: Evidence for the uk," *Fiscal Studies*, 41, 39–63.

CHIAPPORI, P.-A., M. C. DIAS, AND C. MEGHIR (2020c): "Changes in assortative matching: Theory and evidence for the us," Tech. rep., National Bureau of Economic Research.

CHIAPPORI, P.-A., D. ONG, Y. A. YANG, AND J. ZHANG (2021): "Marrying up: field experimental and household survey evidence of competitive search for a taller and richer spouse," .

CHIAPPORI, P.-A., S. OREFFICE, AND C. QUINTANA-DOMEQUE (2012a): "Fatter attraction: anthropometric and socioeconomic matching on the marriage market," *Journal of Political Economy*, 120, 659–695.

——— (2012b): "Fatter attraction: anthropometric and socioeconomic matching on the marriage market," *Journal of Political Economy*, 120, 659–695.

——— (2016): "Black–White marital matching: race, anthropometrics, and socioeconomics," *Journal of Demographic Economics*, 82, 399–421.

CHIAPPORI, P.-A., B. SALANIÉ, AND Y. WEISS (2017): "Partner choice, investment in children, and the marital college premium," *American Economic Review*, 107, 2109–67.

CHIPLUNKAR, G. AND J. WEAVER (2019): "Marriage markets and the rise of dowry in India," *Unpublished Manuscript*.

CHOO, E. AND A. SIOW (2006): "Who marries whom and why," *Journal of political Economy*, 114, 175–201.

CISCATO, E., A. GALICHON, AND M. GOUSSÉ (2020): "Like attract like? A structural comparison of homogamy across same-sex and different-sex households," *Journal of Political Economy*, 128, 740–781.

CISCATO, E. AND S. WEBER (2020): "The role of evolving marital preferences in growing income inequality," *Journal of Population Economics*, 33, 307–347.

COFFEY, D., A. DEATON, J. DRÈZE, D. SPEARS, AND A. TAROZZI (2013a): "Stunting among children: Facts and implications," *Economic and Political Weekly*, 48, 68–70.

COFFEY, D., R. KHERA, AND D. SPEARS (2013b): "Women's status and children's height in India: Evidence from joint rural households," *Unpublished, Office of Population Research, Wallace Hall, Princeton University, Princeton, NJ*, 8540.

COFFEY, D. AND D. SPEARS (2017): *Where India Goes: Abandoned Toilets, Stunted Development and the Costs of Caste*, Harper Collins.

COLE, T. J. (2000): "Galton's midparent height revisited," *Annals of human biology*, 27, 401–405.

COURTIOL, A., M. RAYMOND, B. GODELLE, AND J.-B. FERDY (2010): "Mate choice and human stature: homogamy as a unified framework for understanding mating preferences," *Evolution: International Journal of Organic Evolution*, 64, 2189–2203.

DEATON, A. (2007): "Height, health, and development," *Proceedings of the national academy of sciences*, 104, 13232–13237.

DHAR, D., T. JAIN, AND S. JAYACHANDRAN (2019): "Intergenerational transmission of gender attitudes: Evidence from India," *The Journal of Development Studies*, 55, 2572–2592.

DUGAR, S., H. BHATTACHARYA, AND D. REILEY (2012): "Can't Buy Me Love? A Field Experiment Exploring the Trade-Off between Income and Caste-Status in an Indian Matrimonial Market," *Economic Inquiry*, 50, 534–550.

DUPUY, A. AND A. GALICHON (2014): "Personality traits and the marriage market," *Journal of Political Economy*, 122, 1271–1319.

DUPUY, A., A. GALICHON, AND Y. SUN (2016): "Estimating matching affinity matrix under low-rank constraints," *arXiv preprint arXiv:1612.09585*.

FAFCHAMPS, M. AND A. R. QUISUMBING (2005): "Marriage, bequest, and assortative matching in rural Ethiopia," *Economic Development and Cultural Change*, 53, 347–380.

FERNÁNDEZ, R. AND R. ROGERSON (2001): "Sorting and long-run inequality," *The Quarterly Journal of Economics*, 116, 1305–1341.

FULFORD, S. ET AL. (2013): "The Puzzle of Marriage Migration in India," Tech. rep., Boston College Department of Economics.

GALICHON, A. AND B. SALANIÉ (2020): "Cupid's invisible hand: Social surplus and identification in matching models," *Available at SSRN 1804623*.

GALTON, F. (1886): "Regression towards mediocrity in hereditary stature." *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246–263.

GANG, I. N., K. SEN, AND M.-S. YUN (2008): "Poverty in rural India: Caste and tribe," *Review of Income and Wealth*, 54, 50–70.

GANGULI, I., R. HAUSMANN, AND M. VIARENGO (2014): "Marriage, education and assortative mating in Latin America," *Applied Economics Letters*, 21, 806–811.

GARZA, C., E. BORGHI, A. W. ONYANGO, M. DE ONIS, AND W. M. G. R. S. GROUP (2013): "Parental height and child growth from birth to 2 years in the WHO M ulticentre G rowth R eference S tudy," *Maternal & Child Nutrition*, 9, 58–68.

GERUSO, M. AND D. SPEARS (2018): "Neighborhood sanitation and infant mortality," *American Economic Journal: Applied Economics*, 10, 125–62.

GILLIS, J. S. AND W. E. AVIS (1980): "The male-taller norm in mate selection," *Personality and Social Psychology Bulletin*, 6, 396–401.

GREENWOOD, J., N. GUNER, G. KOCHARKOV, AND C. SANTOS (2014): "Marry your like: Assortative mating and income inequality," *American Economic Review*, 104, 348–53.

GUPTA, A., J. CLELAND, AND T. SEKHER (2021): "Effects of parental stature on child stunting in India," *Journal of Biosocial Science*, 1–12.

HITSCH, G. J., A. HORTAÇSU, AND D. ARIELY (2010): "Matching and sorting in online dating," *American Economic Review*, 100, 130–63.

HOEHN-VELASCO, L. AND J. PENGLASE (2021): "Changes in Assortative Matching and Educational Inequality: Evidence from Marriage and Birth Records in Mexico," *Available at SSRN 3844840*.

JAYACHANDRAN, S. AND R. PANDE (2017): "Why Are Indian Children So Short? The Role of Birth Order and Son Preference," *American Economic Review*, 107, 2600–2629.

MAX ROSER, C. A. AND H. RITCHIE (2013): "Human Height," *Our World in Data*, https://ourworldindata.org/human-height.

NCD ET AL. (2016): "A century of trends in adult human height," *Elife*, 5, e13410.

OREFFICE, S. AND C. QUINTANA-DOMEQUE (2010): "Anthropometry and socioeconomics among couples: Evidence in the United States," *Economics & Human Biology*, 8, 373–384.

ÖZALTIN, E., K. HILL, AND S. SUBRAMANIAN (2010): "Association of maternal stature with offspring mortality, underweight, and stunting in low-to middle-income countries," *Jama*, 303, 1507–1516.

ROSENZWEIG, M. R. AND O. STARK (1989): "Consumption smoothing, migration, and marriage: Evidence from rural India," *Journal of political Economy*, 97, 905–926.

ROSER, M. AND E. ORTIZ-OSPINA (2013): "Primary and Secondary Education," *Our World in Data*, https://ourworldindata.org/primary-and-secondary-education.

SEAR, R. (2006): "Height and reproductive success," *Human Nature*, 17, 405–418.

SMITS, J. AND H. PARK (2009): "Five decades of educational assortative mating in 10 East Asian societies," *Social Forces*, 88, 227–255.

SPEARS, D. (2013): "How Much Intergenerational Variation in Child Height Can Sanitation Explain?" *Policy Research Working Paper 6351*.

STULP, G. AND L. BARRETT (2016): "Evolutionary perspectives on human height variation," *Biological Reviews*, 91, 206–234.

STULP, G., M. J. SIMONS, S. GRASMAN, AND T. V. POLLET (2017): "Assortative mating for human height: A meta-analysis," *American journal of human biology*, 29, e22917.

TANNER, J., H. GOLDSTEIN, AND R. WHITEHOUSE (1970): "Standards for children's height at ages 2-9 years allowing for height of parents," *Archives of disease in childhood*, 45, 755–762.

TANNER, J. M., M. HEALY, R. LOCKHART, J. MACKENZIE, AND R. WHITEHOUSE (1956): "Aberdeen growth study: I. The prediction of adult body measurements from measurements taken each year from birth to 5 years," *Archives of disease in childhood*, 31, 372.

VIDMAR, S. I., T. J. COLE, AND H. PAN (2013): "Standardizing anthropometric measures in children and adolescents with functions for egen: Update," *The Stata Journal*, 13, 366–378.

VOGL, T. S. (2014): "Height, skills, and labor market outcomes in Mexico," *Journal of Development Economics*, 107, 84–96.

WANG, F., E. PUENTES, J. BEHRMAN, F. CUNHA, ET AL. (2021): "You are What Your Parents Expect: Height and Local Reference Points," Tech. rep.

WHO ET AL. (2021): "Levels and trends in child malnutrition: UNICEF," .

A Appendix

Contents

A.1 Singular Value Decomposition of the Affinity Matrix

The affinity matrix helps us understand the degree of complementarity/ substitutability between these features. Next, we would like to quantify the relative importance of each of these features in explaining the marital surplus. This is achieved using a technique advocated by Dupuy and Galichon (2014) known as Saliency analysis. This technique has been used by Ciscato, Galichon, and Goussé (2020) to compare the relative importance of education, age, and wages across the heterosexual, gay, and lesbian marriage markets in the United States and by Chiappori et al. (2020a) to study the relative importance of education, anthropometric measures, and health-related risk behaviors on the marriage market using data from Italy. Here I summarize the Saliency analysis technique for my particular application.

Saliency analysis involves expressing the original joint utility function,$\Phi(x,y)$, in-terms of orthogonal vectors, $\tilde{x}$ and $\tilde{y}$, such that $\Phi(x,y) = x^T Ay = \tilde{x}^T \Lambda \tilde{y}$, where, Λ is a diagonal matrix. If λ_i is the i^{th} diagonal entry of matrix Λ, then $\frac{\lambda_i}{\sum_i \lambda_i}$ is the proportion of the joint utility explained by the i^{th} pair of the orthogonal vectors, $\tilde{x}$ and $\tilde{y}$. Each element of $\tilde{x}$ and $\tilde{y}$ is a weighted sum of the underlying attributes, namely, age, education, height, and parental education. By analyzing the loading weights of the underlying observables on the independent vectors, we can quantify the relative importance of each of the martial features in explaining the joint utility separately for men and women. This will shed light on questions such as: How much of the joint marital utility is explained by height independently? and What is the relative importance of height in explaining the marital surplus compared to education?

The fundamental technique used to perform Saliency analysis is Singular Value Decomposition(SVD) of the affinity matrix A.[49] Using SVD, We can rewrite A as

$$A = U^T \Lambda V$$

, therefore the orthogonal vectors $\tilde{x}$ and $\tilde{y}$ previously defined can be written as, $\tilde{x} = Ux$ and $\tilde{y} = Vy$.

After estimating the affinity matrix using equation 5, I perform SVD on the estimated affinity matrix for each caste-religion-specific marriage market and quantify the relative importance of education, height, and parental education for men and women in each of the religion-caste-specific marriage markets.

A.2 Derivation: Relationship Between Equilibrium Matching and Joint Utility

Dupuy and Galichon (2014) show that by assuming partners are drawn randomly from a Poisson process leads to the continuous multinomial logit choice model. Therefore, the probability of a man m with

[49]More precisely, before performing SVD, we rescale the the observable vectors, X and Y, so that each element of X and Y is mean zero and has a variance of one.

feature vector x, choosing a woman with feature vector y from his randomly drawn set is given by:

$$\pi(y|x)=\frac{e^{[U(x,y)/(\sigma/2)]}}{\int_{\mathbb{R}^4} e^{[U(x,t)/(\sigma/2)]}dt}. \tag{A1}$$

Similarly, the probability of a woman, w with a feature vector y choosing a man with feature vector x from her randomly drawn set is given by:

$$\pi(x|y)=\frac{e^{[V(t,y)/(\sigma/2)]}}{\int_{\mathbb{R}^4} e^{[V(t,y)/(\sigma/2)]}dt}. \tag{A2}$$

Since, $\pi(y|x)=\pi(x,y)/f_x(x)$ and $\pi(x|y)=\pi(x,y)/f_y(y)$, we can write the following equality:

$$\pi(x,y)=\frac{e^{[V(t,y)/(\sigma/2)]}}{\frac{\int_{\mathbb{R}^4} e^{[V(t,y)/(\sigma/2)]}dt}{f_y(y)}}=\frac{e^{[U(x,y)/(\sigma/2)]}}{\frac{\int_{\mathbb{R}^4} e^{[U(x,t)/(\sigma/2)]}dt}{f_x(x)}},$$

Therefore,

$$(\sigma/2)\log(\pi(x,y))=V-b(y)=U-a(x)$$

where,

$$b(y)=\sigma/2\times\log\left(\frac{\int_{\mathbb{R}^4} e^{[V(t,y)/(\sigma/2)]dt}}{f_y(y)}\right)$$

$$a(x)=\sigma/2\times\log\left(\frac{\int_{\mathbb{R}^4} e^{[U(t,y)/(\sigma/2)]}dt}{f_x(x)}\right)$$

which implies,

$$\log(\pi(x,y))=U+V-a(x)-b(y)$$

Since,

$$\Phi(x,y)=U+V$$

We get the final expressions as follows:

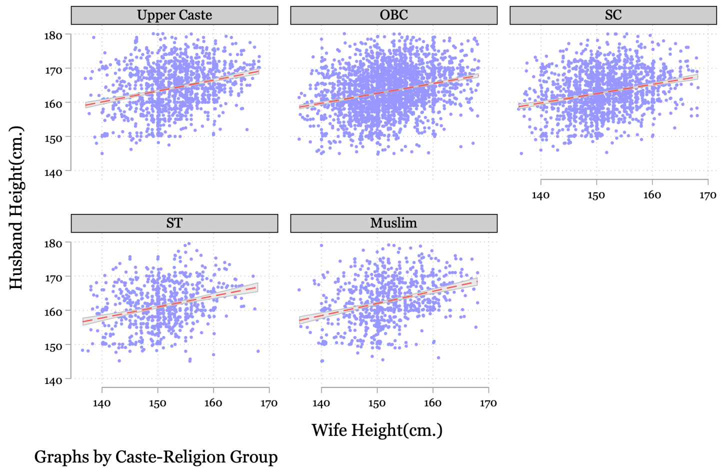
$$\log\pi(x,y)=\frac{\Phi(x,y)-a(x)-b(y)}{\sigma}$$

$$U=\frac{\Phi(x,y)+a(x)-b(y)}{2}$$

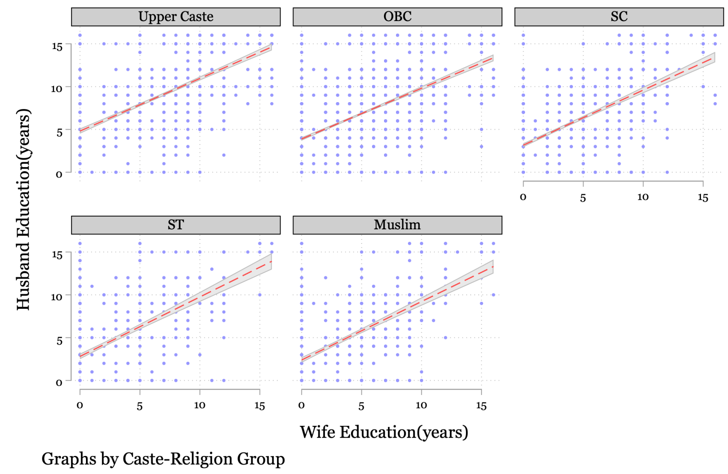
$$V=\frac{\Phi(x,y)-a(x)+b(y)}{2}$$

A.3 Tables and Figures

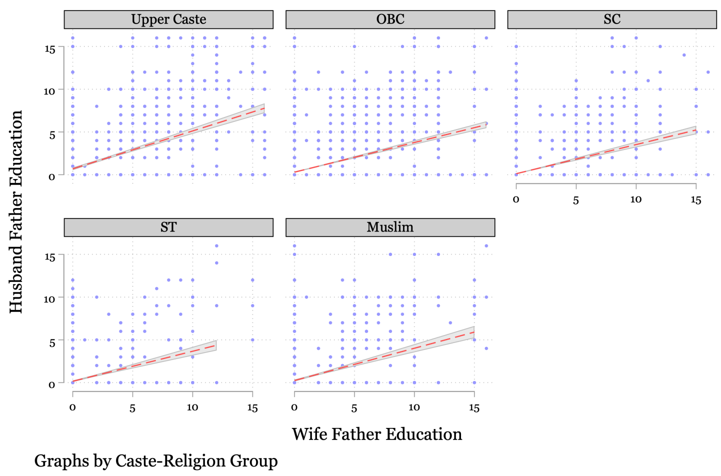
Figure A1: Correlations Between Husband and Wife Marriage Market Features



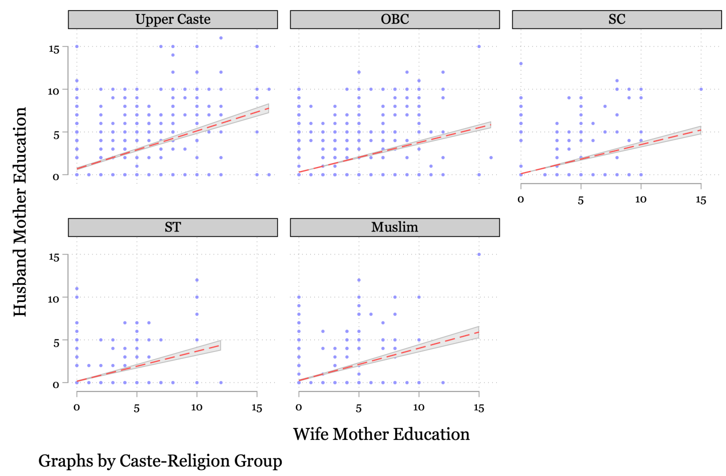
(A) Height



(B) Education



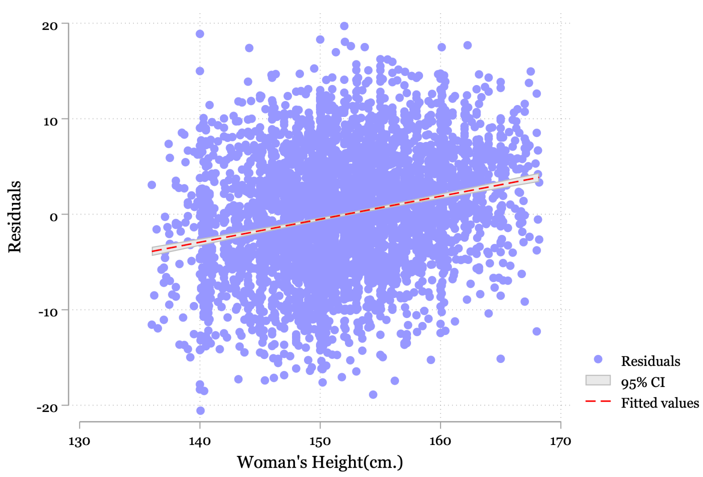
(C) Father Education



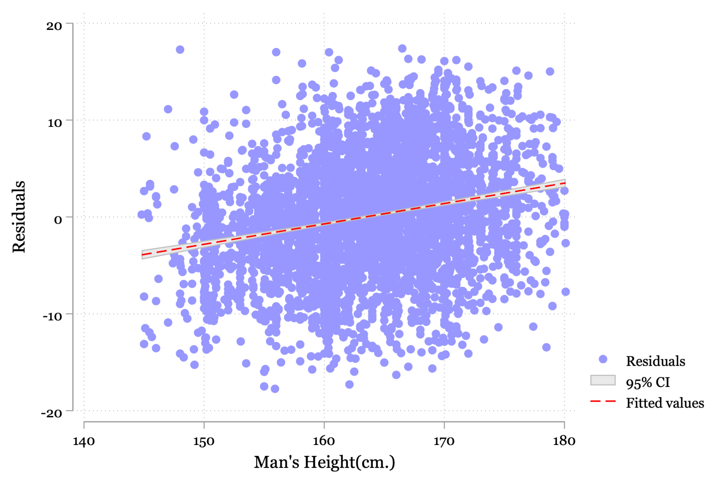
(C) Mother Education

Note: Figure shows the raw correlations between the husband and the wife's marriage market features using the IHDS sample

Figure A2: Height Regression: Residuals



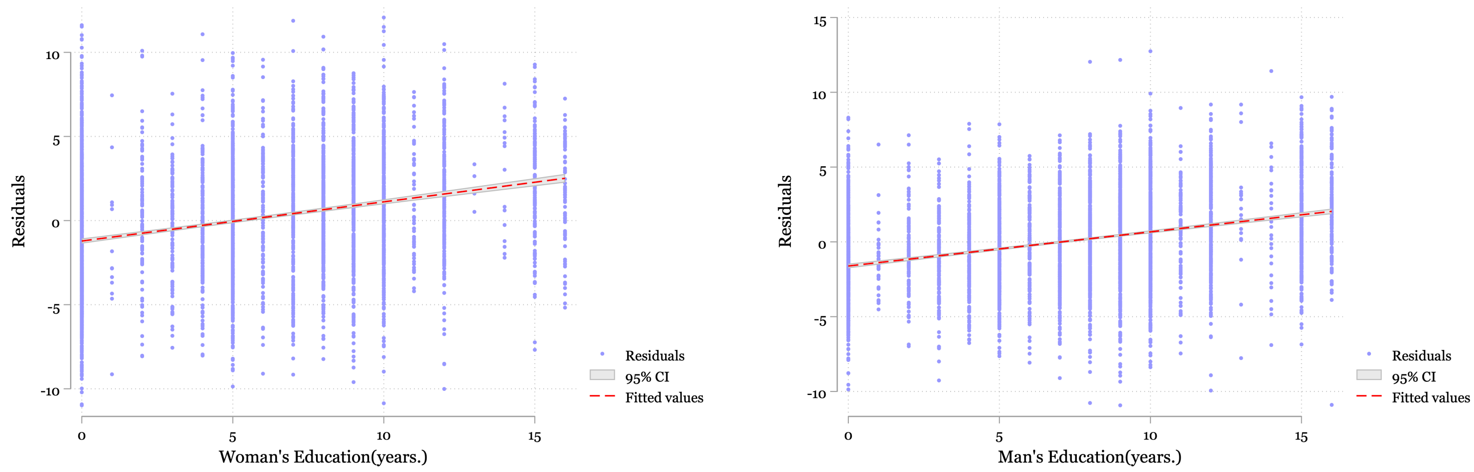
(A) Men's Height
Regression Residuals



(B) Women's Height
Regression Residuals

Note: Graph(A) plots the residuals obtained after regressing men's height on observables excluding women's height. Graph(B) plots the residuals obtained after regressing women's height on observables excluding men's height. Observables include: the husband and the wife's education, father's education, mother's education, year of marriage fixed effects, state fixed effects and caste-religious group fixed effects, with standard errors clustered at the state-level.

Figure A3: Education Regression: Residuals

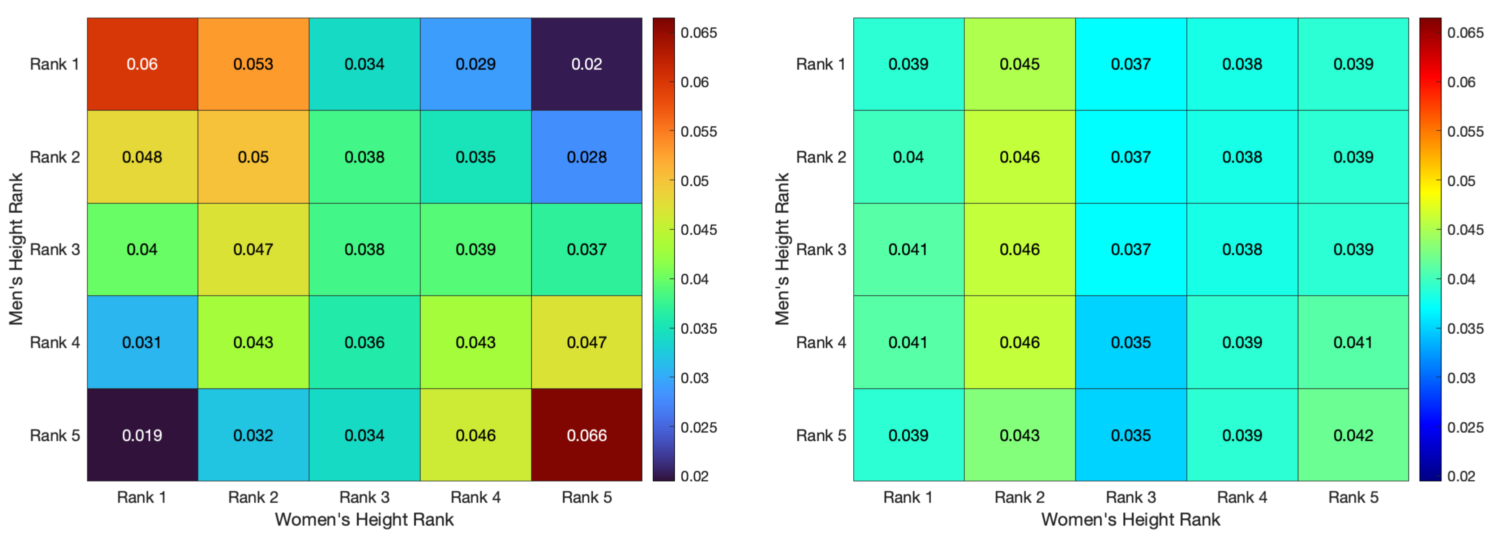


(A) Men's Education Regression Residuals

(B) Women's Education Regression Residuals

Note: Graph(A) plots the residuals obtained after regressing men's education on observables excluding women's education. Graph(B) plots the residuals obtained after regressing women's education on observables excluding men's education. Observables include: the husband and the wife's education, father's education, mother's education, year of marriage fixed effects, state fixed effects and caste-religious group fixed effects, with standard errors clustered at the state-level.

Figure A4: Matching on Height(Upper Caste): No Complementarities

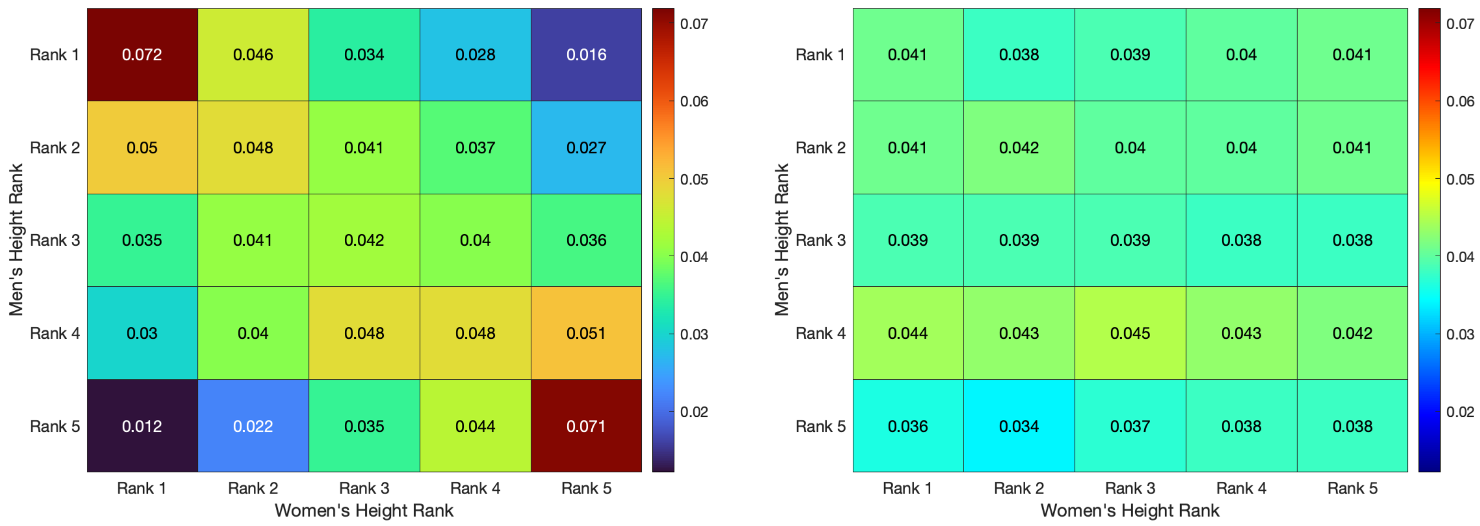


(A) Predicted Rank-Rank Matching

(B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there was no height preference in the marriage market.

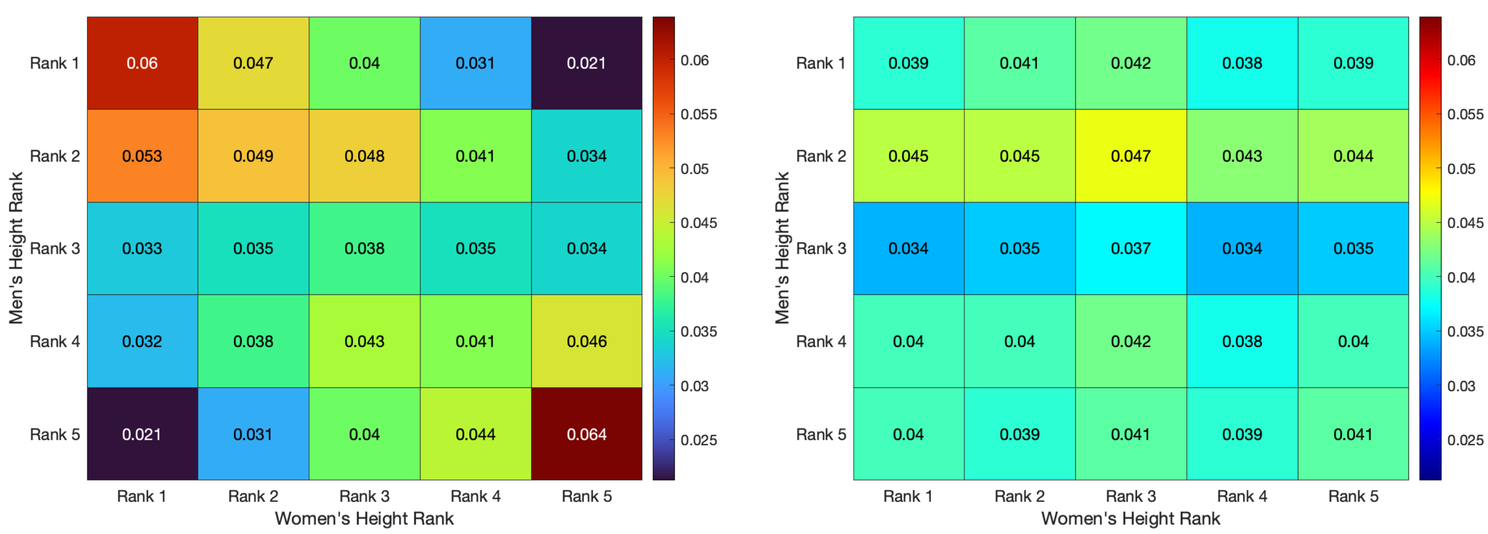
Figure A5: Matching on Height(Muslims): No Complementarities



(A) Predicted Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there were no complementarities in height in the marriage market. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

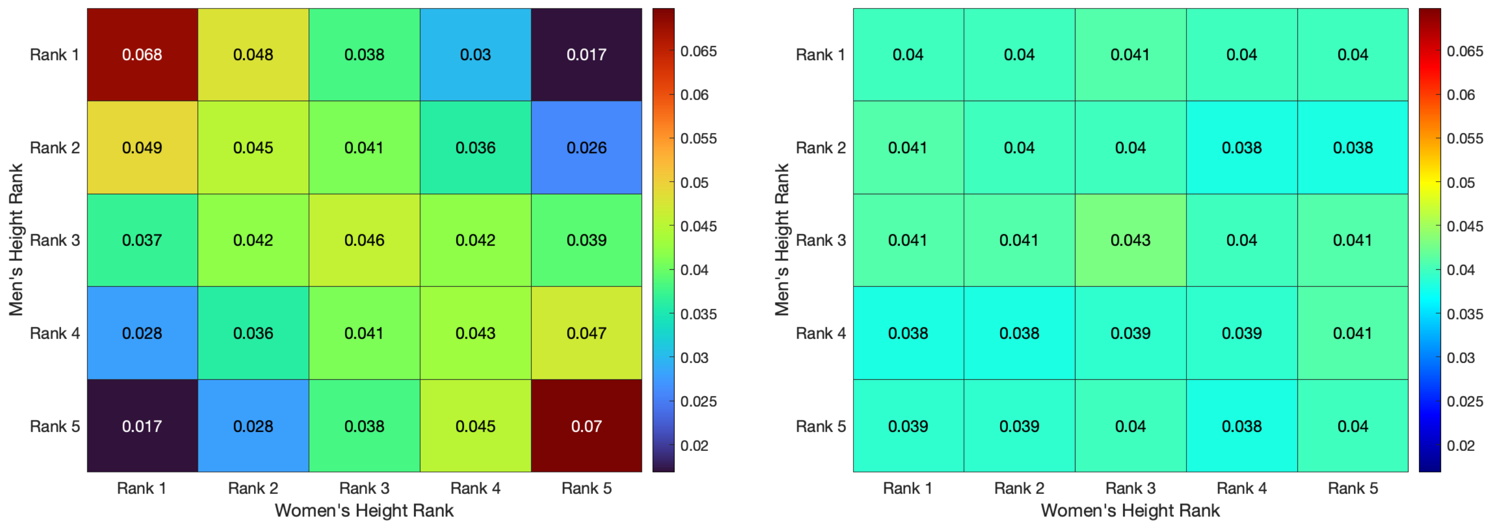
Figure A6: Matching on Height(Scheduled Caste): No Complementarities



(A) Predicted Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there were no complementarities in height in the marriage market. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

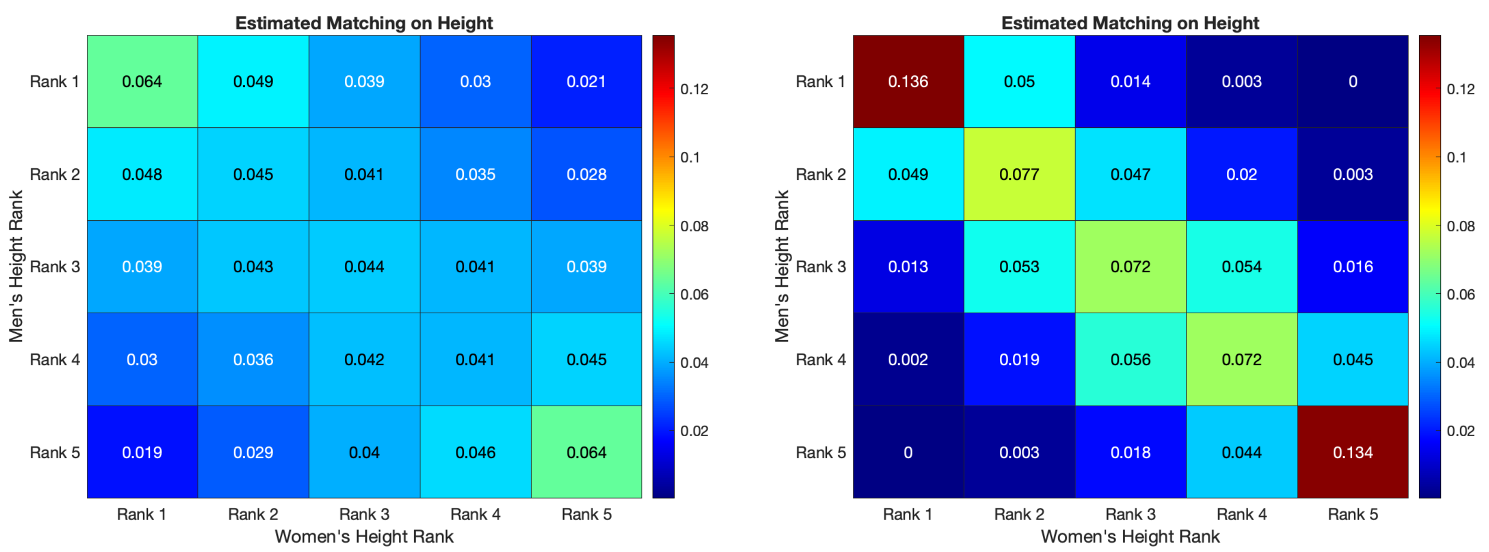
Figure A7: Matching on Height(Scheduled Tribe): No Complementarities



(A) Predicted Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there were no complementarities in height in the marriage market. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

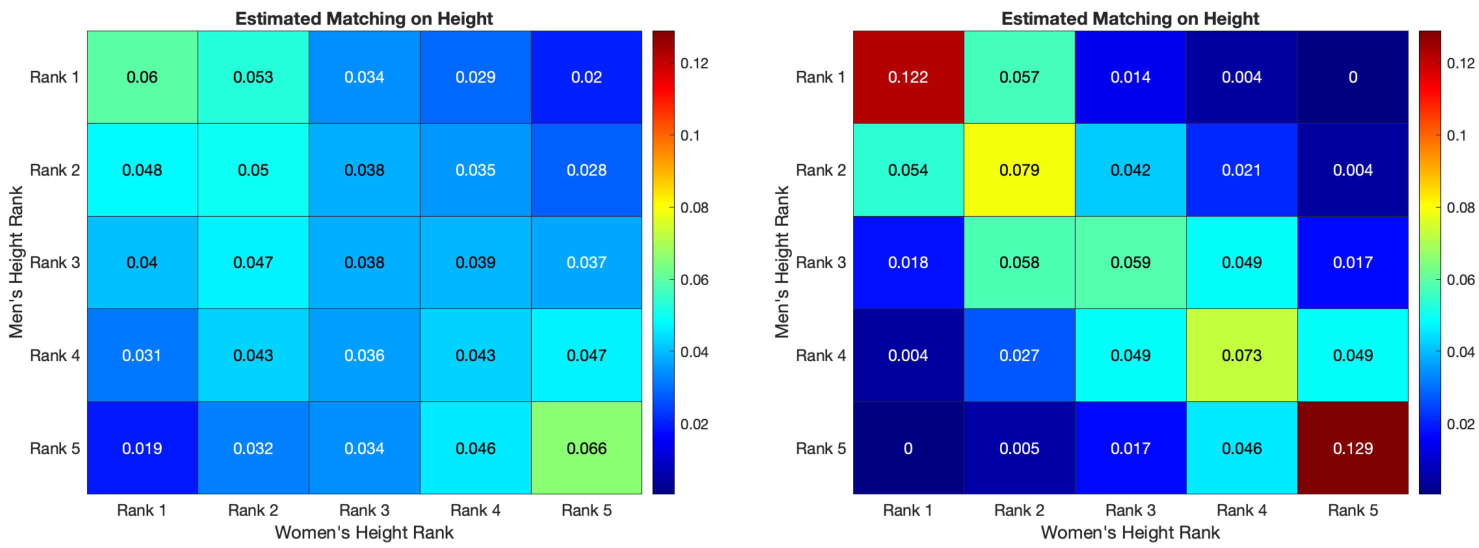
Figure A8: Matching on Height(Other Backward Caste): Same Complementarities as Age



(A) Actual Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there the complementarities in height were the same as complementarities in age. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

Figure A9: Matching on Height(Upper Caste): Same Complementarities as Age

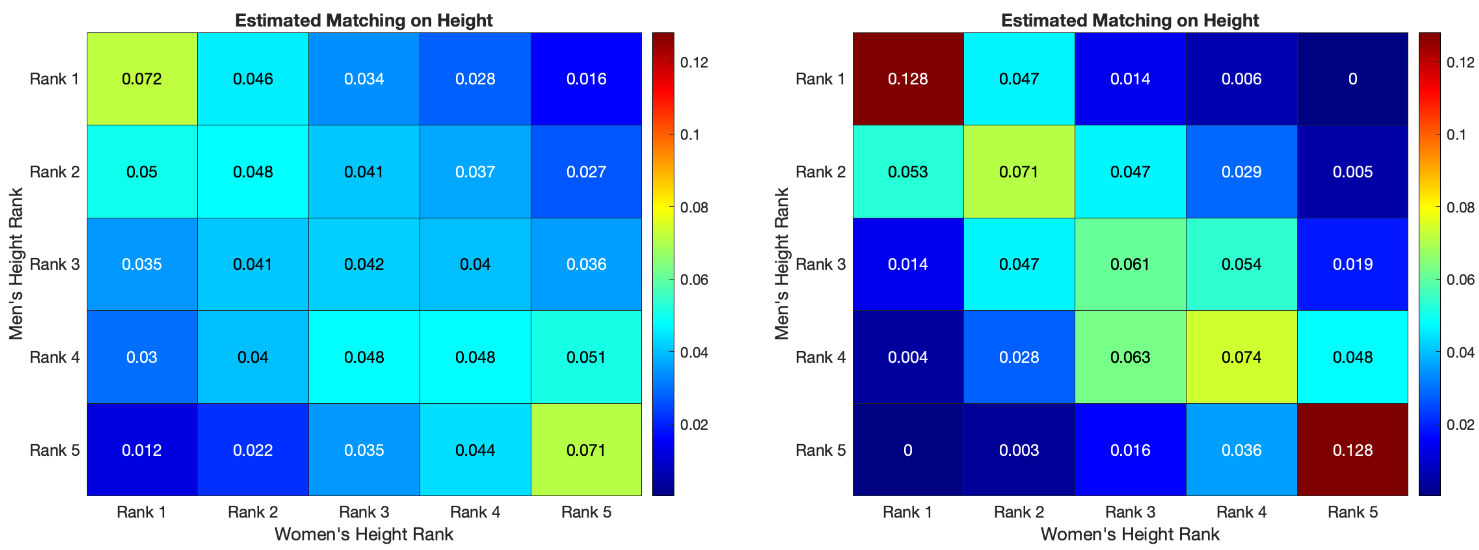


(A) Actual Rank-Rank Matching

(B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there the complementarities in height were the same as complementarities in age. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

Figure A10: Matching on Height(Muslims): Same Complementarities as Age

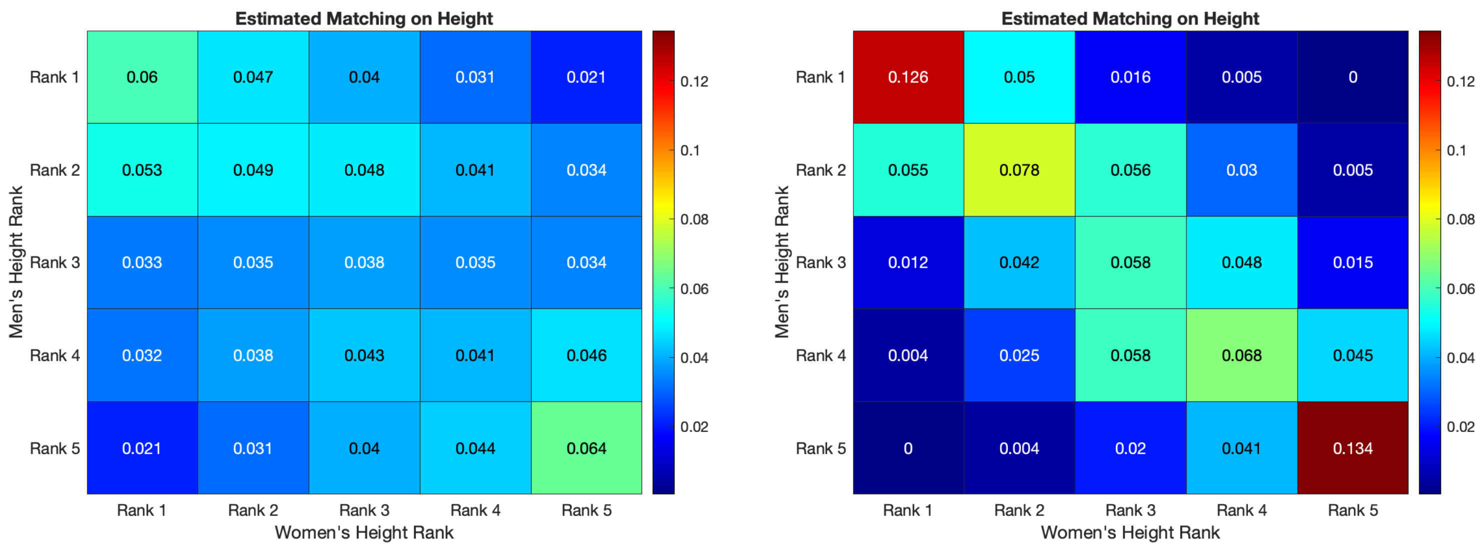


(A) Actual Rank-Rank Matching

(B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there the complementarities in height were the same as complementarities in age. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

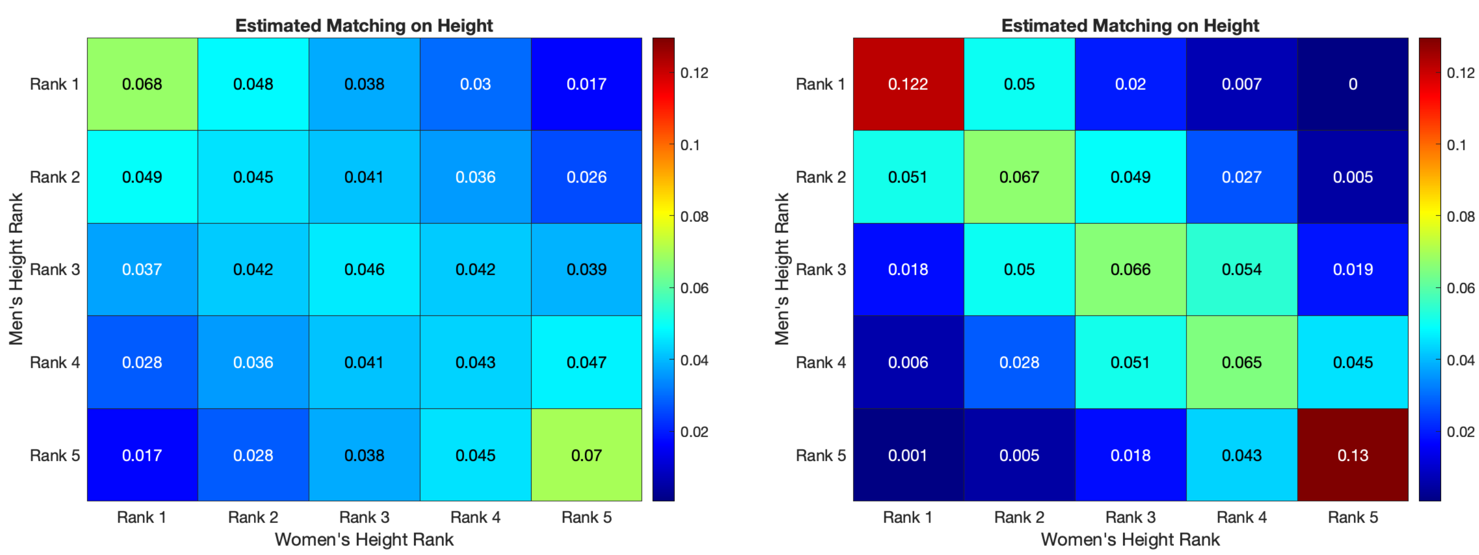
Figure A11: Matching on Height(Scheduled Caste): Same Complementarities as Age



(A) Actual Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there the complementarities in height were the same as complementarities in age. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

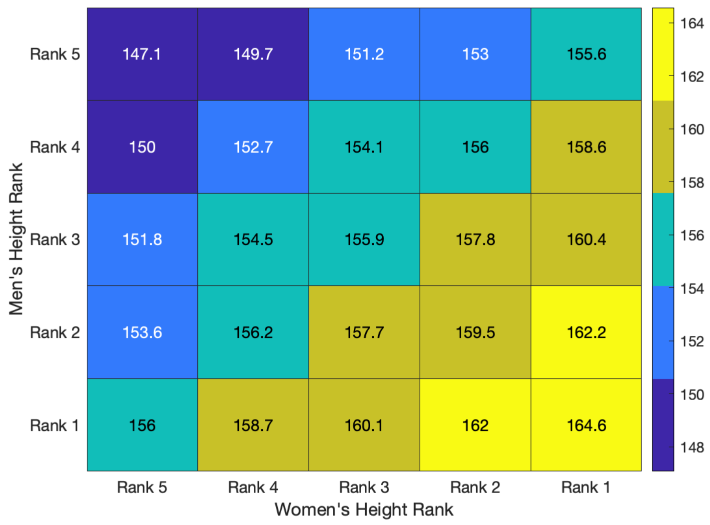
Figure A12: Matching on Height(Scheduled Tribe): Same Complementarities as Age



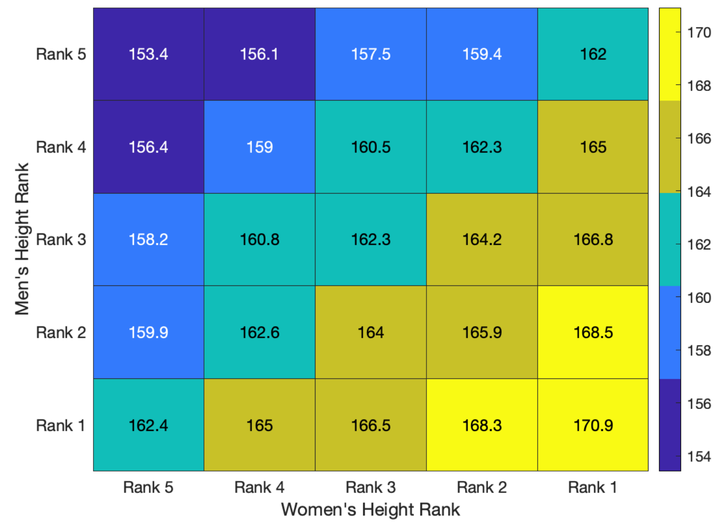
(A) Actual Rank-Rank Matching (B) Counterfactual Rank-Rank Matching

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows the actual Rank-Rank matching and plot (B) shows the counterfactual Rank-Rank matching if there the complementarities in height were the same as complementarities in age. Each value represents the probability mass of a given rank-rank matching. The entries sum to one.

Figure A13: Children's Potential Height:Upper Caste



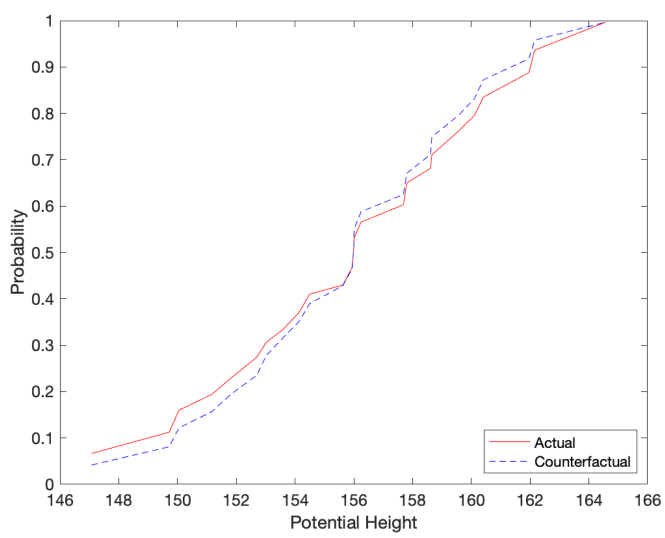
(A) Girl's Potential Height



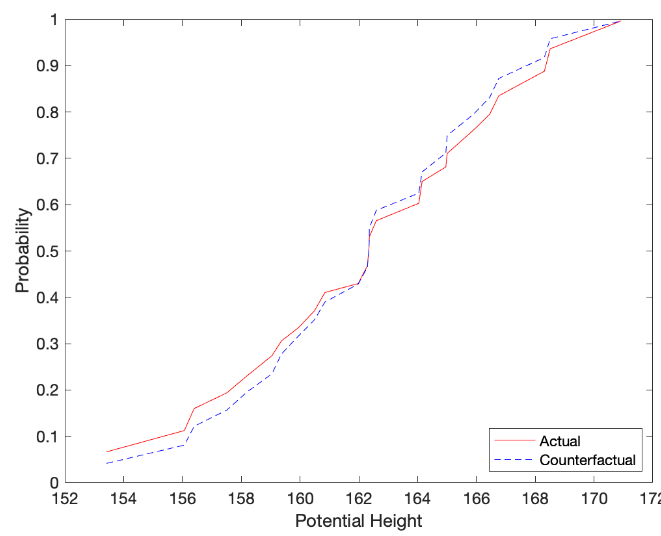
(B) Boy's Potential Height

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1956).

Figure A14: Children's Potential Height Distribution(Upper Caste): Actual vs Counterfactual



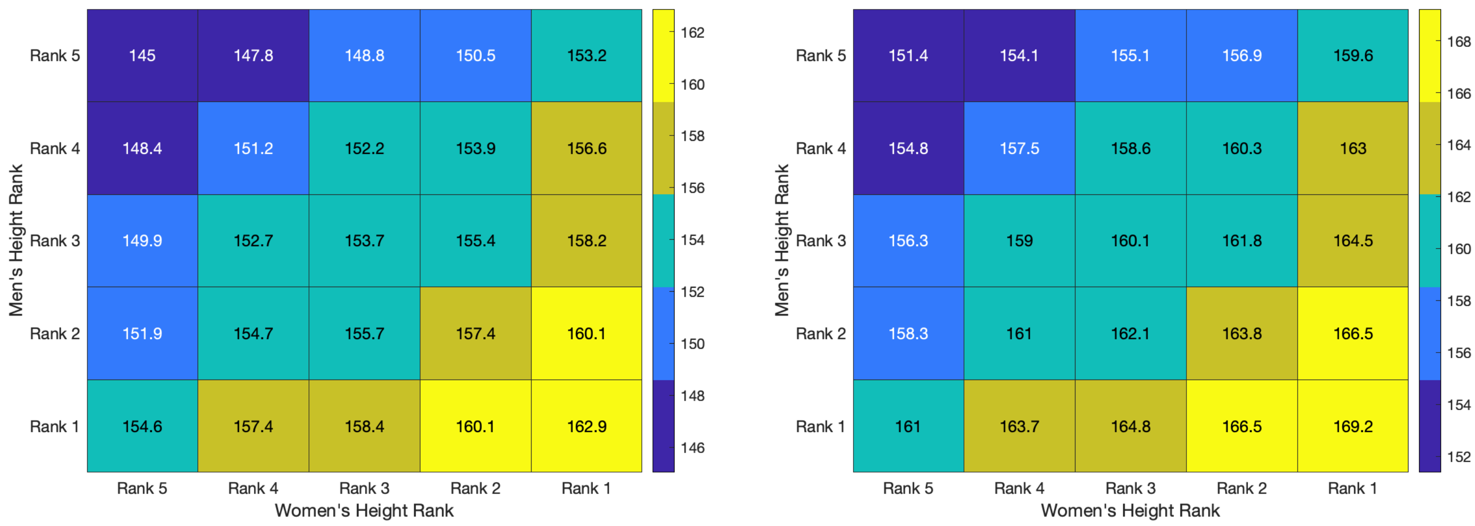
(A) Girl's Potential Distribution



(B) Boy's Potential Distribution

Note: Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method Tanner et al. (1970). Counterfactual potential height distribution is calculated under the hypothetical, no preference for height on the marriage market.

Figure A15: Children's Potential Height:Muslims

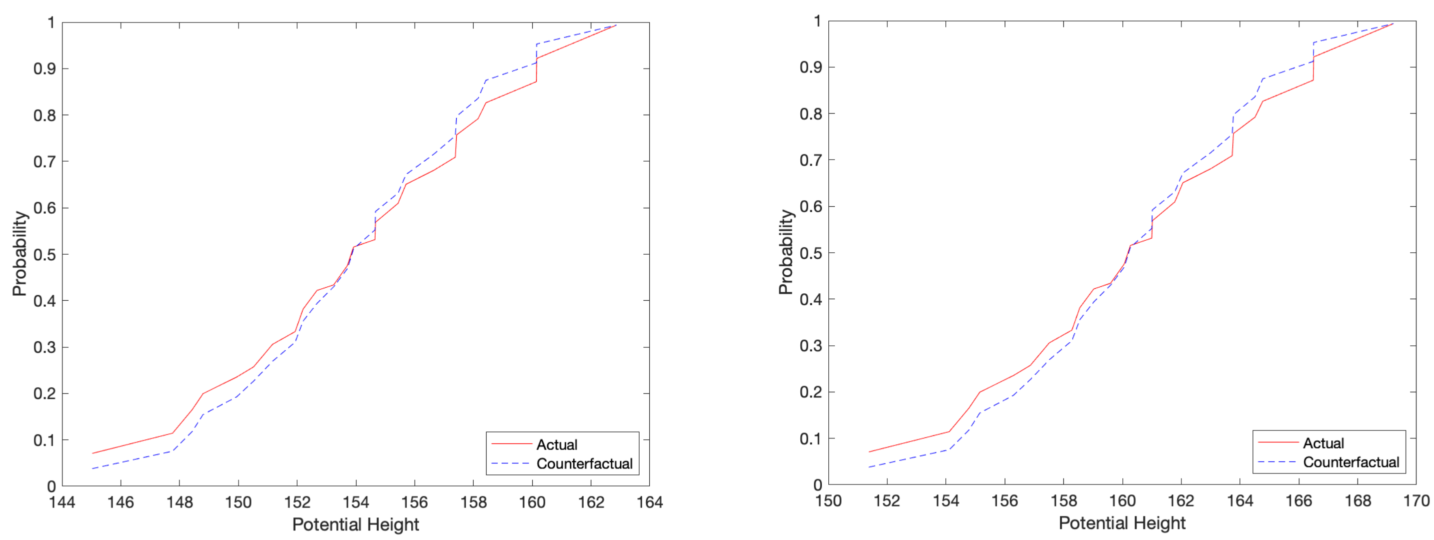


(A) Girl's Potential Height

(B) Boy's Potential Height

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1956).

Figure A16: Children's Potential Height Distribution(Muslims): Actual vs Counterfactual

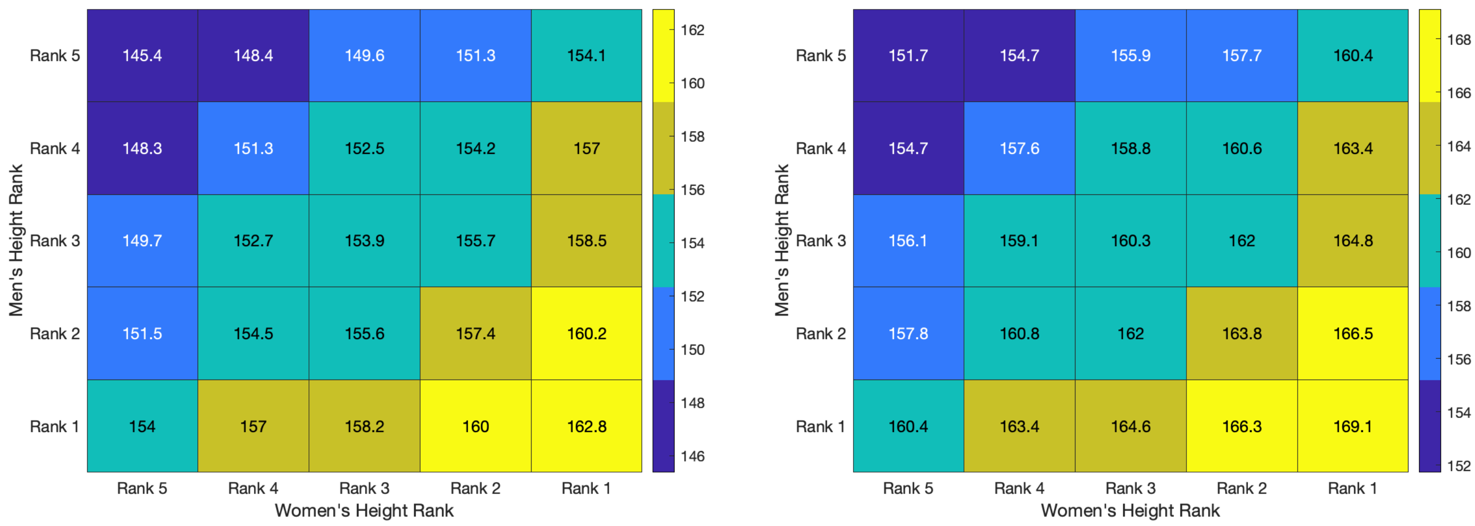


(A) Girl's Potential Distribution

(B) Boy's Potential Distribution

Note: Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method Tanner et al. (1970).

Figure A17: Children's Potential Height:Scheduled Caste

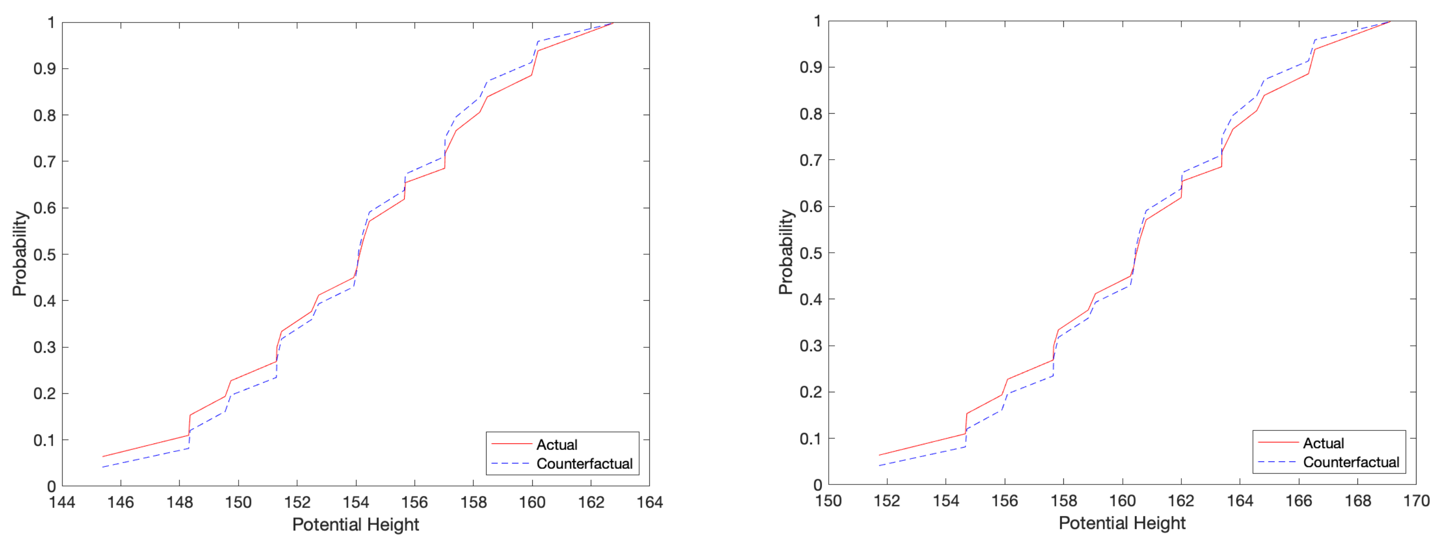


(A) Girl's Potential Height

(B) Boy's Potential Height

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1956).

Figure A18: Children's Potential Height Distribution(Scheduled Caste): Actual vs Counterfactual

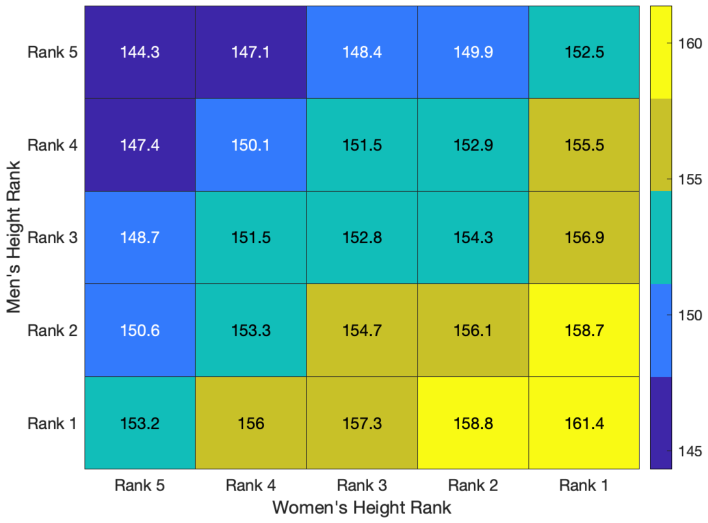


(A) Girl's Potential Distribution

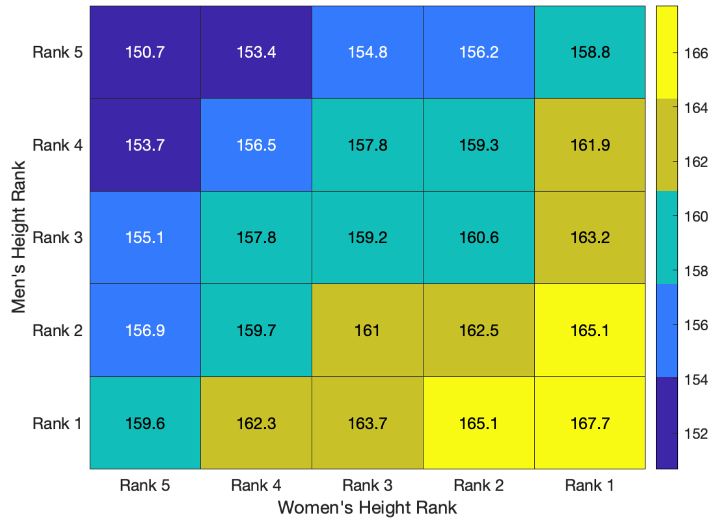
(B) Boy's Potential Distribution

Note: Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1956).

Figure A19: Children's Potential Height:Scheduled Tribe



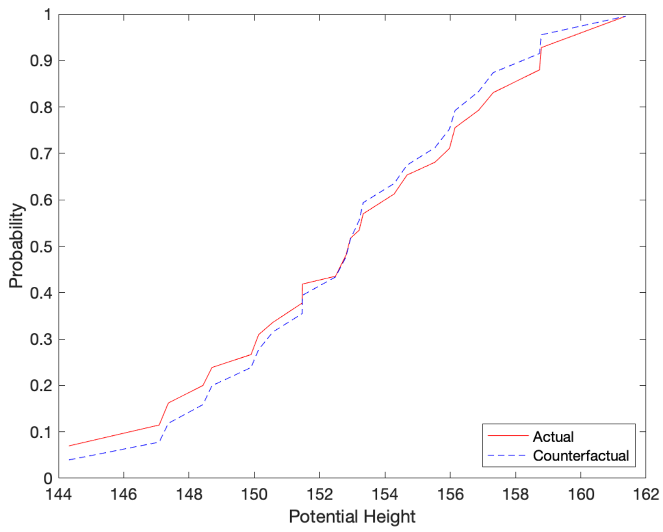
(A) Girl's Potential Height



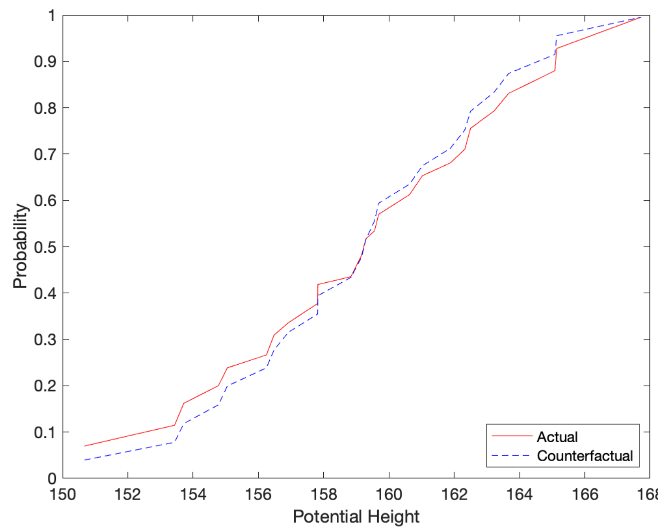
(B) Boy's Potential Height

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1956).

Figure A20: Children's Potential Height Distribution(Scheduled Tribe): Actual vs Counterfactual



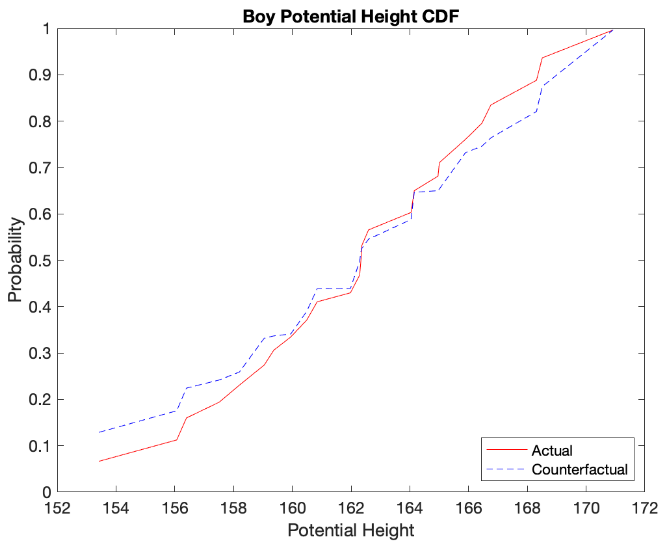
(A) Girl's Potential Distribution



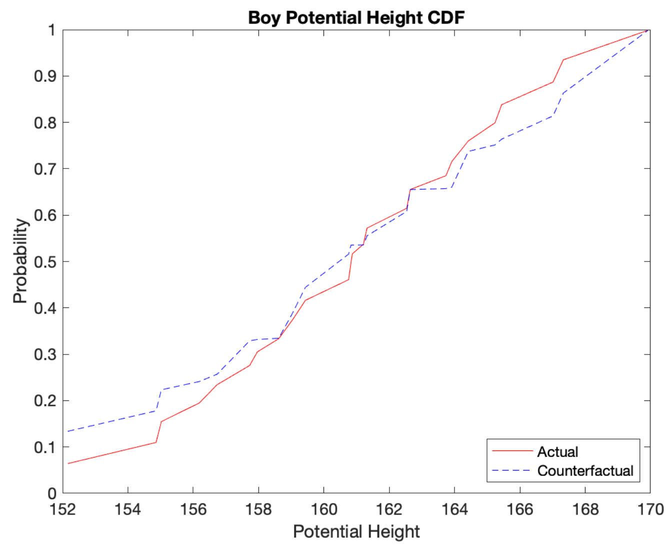
(B) Boy's Potential Distribution

Note: Men's height and women's height is divided into 5 ranks. Rank 1 is top 20th percentile, Rank 2 is 20th to 40th percentile, Rank 3 is 40th to 60th percentile, Rank 4 is 60th to 80th percentile and Rank 5 is bottom 20th percentile. Plot (A) shows girl's potential height with respect to parental height rank (B) shows boy's potential height with respect to parental height rank. Potential height calculated as the mid-parental height using Tanner method (Tanner et al., 1956).

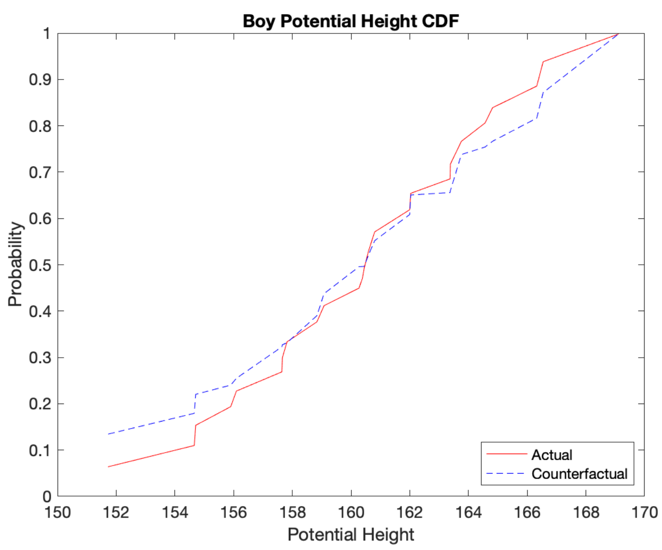
Figure A21: Boy's Potential Height: Actual vs Counterfactual



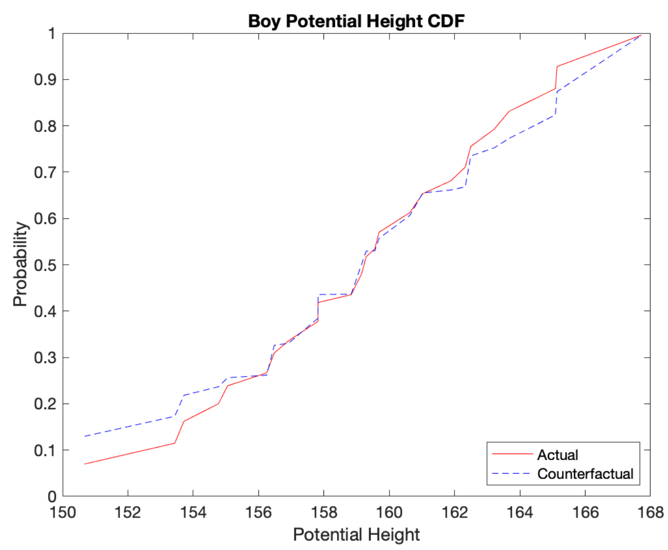
(A) Upper Caste



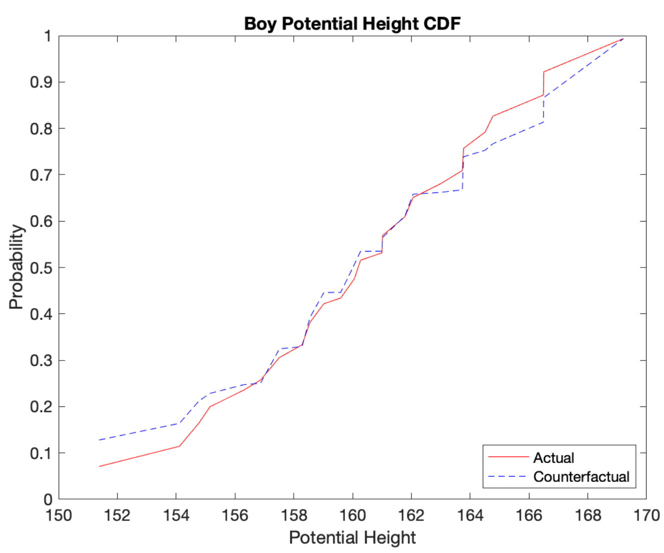
(B) Other Backward Caste



(B) Scheduled Caste



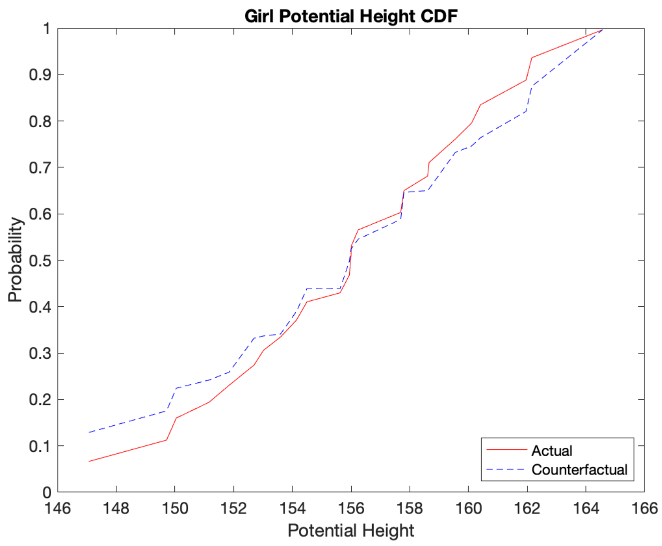
(B) Scheduled Tribe



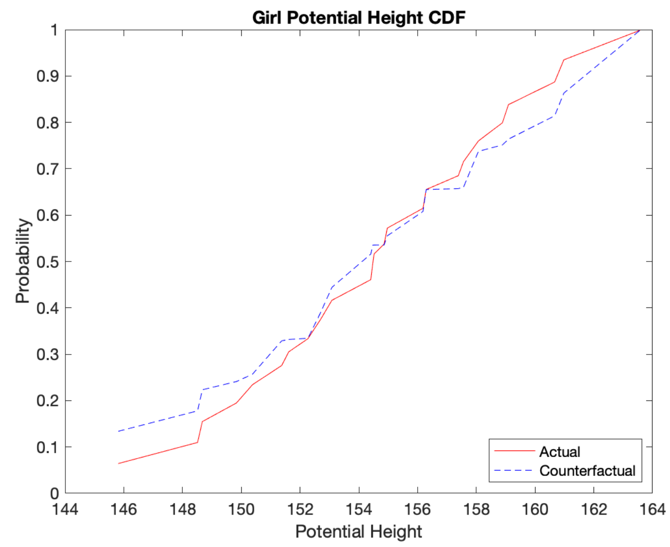
(B) Muslims

Note: The X-axis is the boy's potential height calculated using the Tanner method (Tanner et al., 1970). Counterfactual preferences refer to the hypothetical scenario in which the preferences for height on the marriage market are the same as the preferences for age.

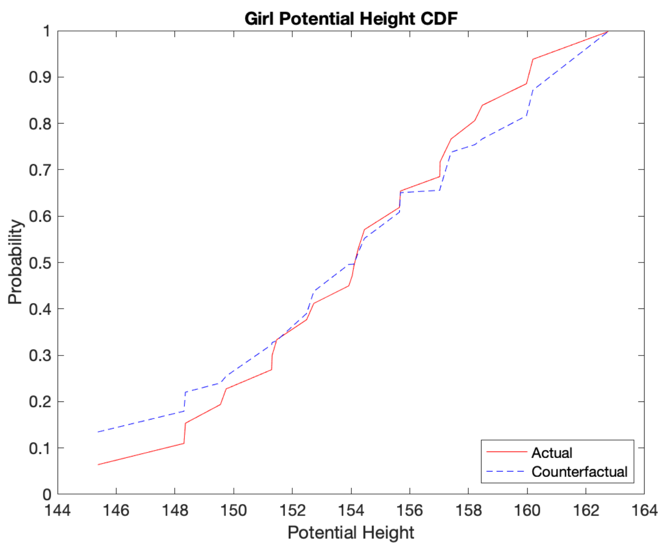
Figure A22: Girl's Potential Height: Actual vs Counterfactual



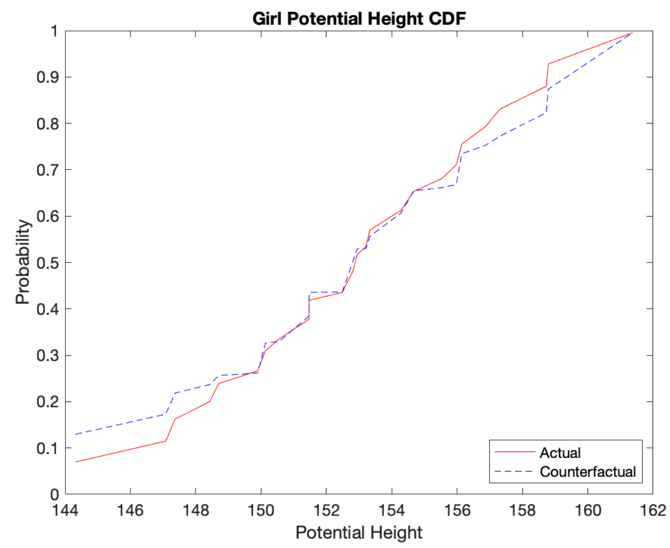
(A) Upper Caste



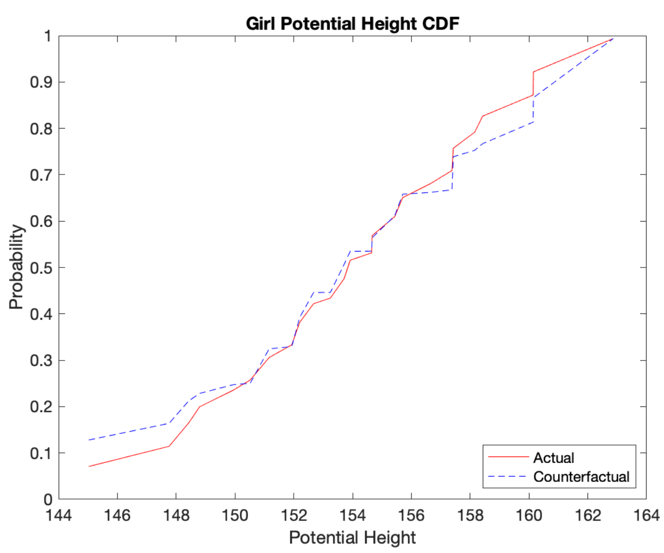
(B) Other Backward Caste



(B) Scheduled Caste



(B) Scheduled Tribe



(B) Muslims

Note: The X-axis is the girl's potential height calculated using the Tanner method (Tanner et al., 1970). Counterfactual preferences refer to the hypothetical scenario in which the preferences for height on the marriage market are the same as the preferences for age.

Table A1: Sorting on Education

	(1) Husband Education	(2) Wife Education
Wife Education	0.505***	
	(0.023)	
Wife Height	0.016	0.018**
	(0.010)	(0.008)
Husband Height	0.019**	0.011
	(0.007)	(0.007)
Husband Father Education	0.174***	0.074***
	(0.021)	(0.018)
Wife Father Education	0.063***	0.203***
	(0.016)	(0.014)
Husband Mother Education	0.017	0.032
	(0.023)	(0.027)
Wife Mother Education	0.013	0.163***
	(0.031)	(0.020)
Husband Education		0.363***
		(0.017)
Region Fixed Effects	Yes	Yes
Year of Marriage Fixed Effects	Yes	Yes
Covariates	Yes	Yes
Observations	4,652	4,652
Mean of Dep. Variable	7.066	5.218

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

NOTES: Covariates include religion, caste, age of the husband and age of the wife. Husband and wife's height is measured in centimeters. Education is measured in number of years.Standard errors are clustered at the state level.

Table A2: Sorting on Father Education

	(1) Husband Father Education	(2) Wife Father Education
Wife Father Education	0.168***	
	(0.026)	
Wife Education	0.080***	0.237***
	(0.020)	(0.023)
Husband Education	0.135***	0.053***
	(0.020)	(0.013)
Wife Height	0.008	0.001
	(0.009)	(0.007)
Husband Height	0.021***	0.019*
	(0.007)	(0.010)
Husband Mother Education	0.584***	-0.025
	(0.031)	(0.029)
Wife Mother Education	0.020	0.551***
	(0.023)	(0.031)
Husband Father Education		0.181***
		(0.029)
Region Fixed Effects	Yes	Yes
Year of Marriage Fixed Effects	Yes	Yes
Covariates	Yes	Yes
Observations	4,652	4,652
Mean of Dep. Variable	2.777	3.332

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

NOTES: Covariates include religion, caste, age of the husband and age of the wife. Husband and wife's height is measured in centimeters. Education is measured in number of years.Standard errors are clustered at the state level.

Table A3: Sorting on Mother Education

	(1) Husband Mother Education	(2) Wife Mother Education
Wife Mother Education	0.233***	
	(0.023)	
Wife Education	0.013	0.085***
	(0.011)	(0.011)
Husband Education	0.005	0.005
	(0.007)	(0.012)
Wife Height	0.000	-0.004
	(0.007)	(0.004)
Husband Height	-0.007	0.005
	(0.005)	(0.005)
Husband Father Education	0.215***	0.009
	(0.022)	(0.011)
Wife Father Education	-0.009	0.246***
	(0.010)	(0.024)
Husband Mother Education		0.307***
		(0.023)
Region Fixed Effects	Yes	Yes
Year of Marriage Fixed Effects	Yes	Yes
Covariates	Yes	Yes
Observations	4,652	4,652
Mean of Dep. Variable	0.941	1.344

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

NOTES: Covariates include religion, caste, age of the husband and age of the wife. Husband and wife's height is measured in centimeters.Standard errors are clustered at the state level.

Table A4: Summary Statistics:IHDS Sample

	Observations	Mean	St.Dev.	Median	Minimum	Maximum
Husband Height	7355	163.075	6.551	163.000	144.800	180.100
Husband Education	7355	6.165	4.774	7.000	0.000	16.000
Husband Father Education	7355	2.287	3.739	0.000	0.000	16.000
Husband Mother Education	7355	0.707	2.076	0.000	0.000	16.000
Husband Age	7355	44.144	9.708	40.000	30.000	60.000
Wife Height	7355	151.908	6.116	151.300	136.000	168.200
Wife Education	7355	4.096	4.525	3.000	0.000	16.000
Wife Father Education	7355	2.712	4.053	0.000	0.000	16.000
Wife Mother Education	7355	1.024	2.524	0.000	0.000	16.000
Wife Age	7355	39.010	9.145	37.000	20.000	60.000
Height Gap	7355	11.167	7.552	10.400	-20.000	40.500
Education Gap	7355	2.069	4.019	1.000	-12.000	16.000
Age Gap	7355	5.134	2.742	5.000	0.000	13.000
Father Education Gap	7355	-0.425	3.984	0.000	-16.000	16.000
Mother Education Gap	7355	-0.317	2.348	0.000	-16.000	15.000
Natal Family Wealthier	7347	0.171	0.377	0.000	0.000	1.000
Upper Caste	7355	0.186	0.389	0.000	0.000	1.000
Other Backward Caste	7355	0.376	0.484	0.000	0.000	1.000
Scheduled Caste	7355	0.224	0.417	0.000	0.000	1.000
Scheduled Tribe	7355	0.101	0.301	0.000	0.000	1.000
Muslims	7355	0.113	0.316	0.000	0.000	1.000
Observations	7355					

NOTES: Age gap is the difference between the husband's age and the wife's age. Education gap is the difference between husband's education and wife's education. Height gap is the difference between husband's height and wife's height. Father education gap is the difference between husband's father's education and wife's father's education and Mother education gap is the difference between husband's mother's education and wife's mother's education.

Figure A23: Child Stunting

Note: Figure plots the prevalence of child stunting based on mother's and father's height rank conditional on observables consist of the father's education, father's age, paternal grandfather's literacy, paternal grandmother's literacy, mother's education, mother's age, wife's maternal grandfather's literacy and maternal grandmother's literacy, religion-caste group fixed effects, district fixed effects and age fixed effects.

Table A5: Child Stunting

	(1) Height Z-Score	(2) Child Stunted
Above Median Rank Father	0.274***	-0.079***
	(0.037)	(0.013)
Above Median Rank Mother	0.388***	-0.082***
	(0.034)	(0.015)
Piped Water	0.100	-0.043*
	(0.060)	(0.022)
Household Expenditure	0.195***	-0.038***
	(0.057)	(0.012)
Female Child	-0.101***	0.020**
	(0.034)	(0.009)
Education Husband	0.002	-0.001
	(0.008)	(0.002)
Education Wife	0.016**	-0.003
	(0.007)	(0.002)
District Fixed Effects	Yes	Yes
Age Fixed Effects	Yes	Yes
Caste-Religion Fixed Effects	Yes	Yes
Observations	8,119	8,119
Mean of Dep. Variable	-1.292	0.315

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

NOTES: A child is defined as stunted if their height-for-age z-score is 2 SD or more below the reference population median for a given age and gender. Observables consist of the father's education, father's age, paternal grandfather's literacy, paternal grandmother's literacy, mother's education, mother's age, wife's maternal grandfather's literacy and maternal grandmother's literacy. Religion-Caste specific groups consist of Scheduled Caste, Scheduled Tribe, Other Backward Caste, Upper Caste and Muslims. Standard errors are clustered at the state-level.

Table A6: Men's Height Premium in India

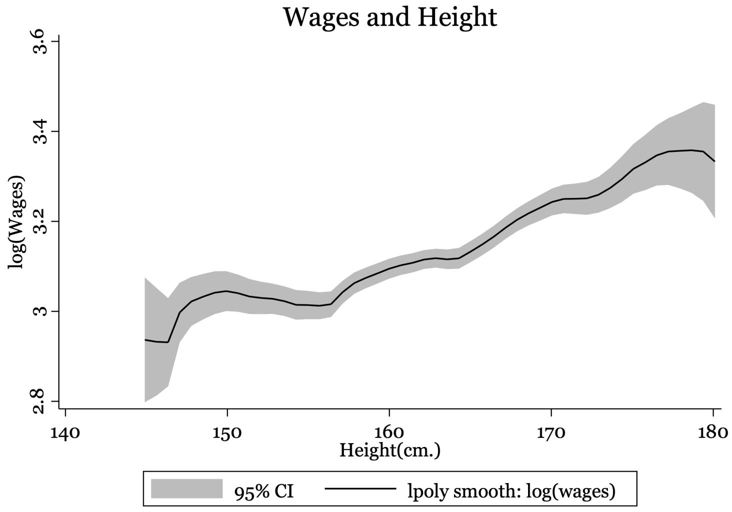
	(1) log(wages)	(2) log(income)	(3) Regular Full-time Job
Height(cm.)	0.0077***	0.0100***	0.0022**
	(0.0013)	(0.0017)	(0.0008)
Education(years)	0.0250***	0.0315***	0.0220***
	(0.0026)	(0.0022)	(0.0013)
State Fixed Effects	Yes	Yes	Yes
Caste-Religion Fixed Effects	Yes	Yes	Yes
Covariates	Yes	Yes	Yes
Observations	9,026	13,329	9,221
Mean of Dep. Variable	3.117	11.119	0.197

Standard errors in parentheses

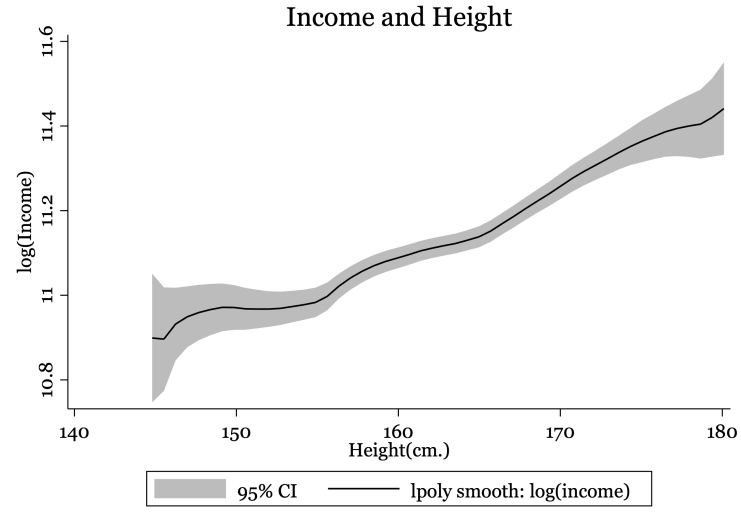
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

NOTES: Sample consists of men who are household heads and in the age group 18 to 65. Wages represent in Covariates consist of individual's english ability, father's education, mother's education, number of brothers and sisters, age and age squared. Caste-Religion groups consist of Upper Caste, Other Backward Caste, Scheduled Caste, Scheduled Tribe and Muslims. Regular Full-time job is 1 if the individual has a Regular/Permanent/Longer contract job and 0 if the individual has a casual daily, casual piecework or a less than 1 year contract job. Standard errors are clustered at the state-level.

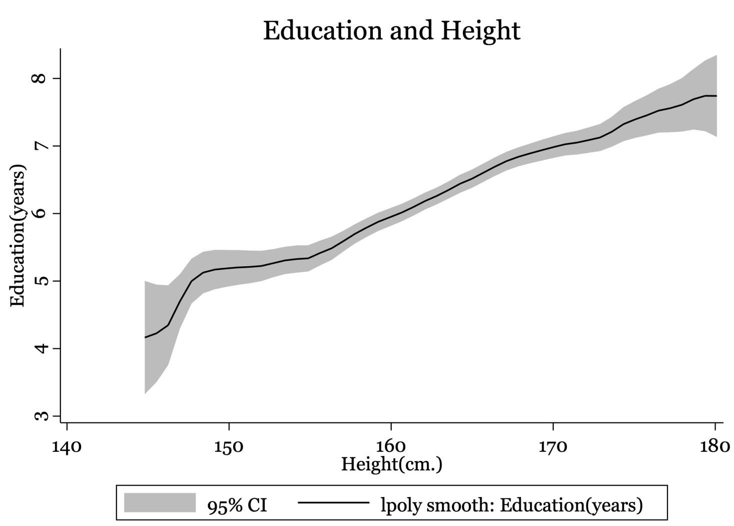
Figure A24: Men's Height, Education, Wages and Income



(A) Men's Wages



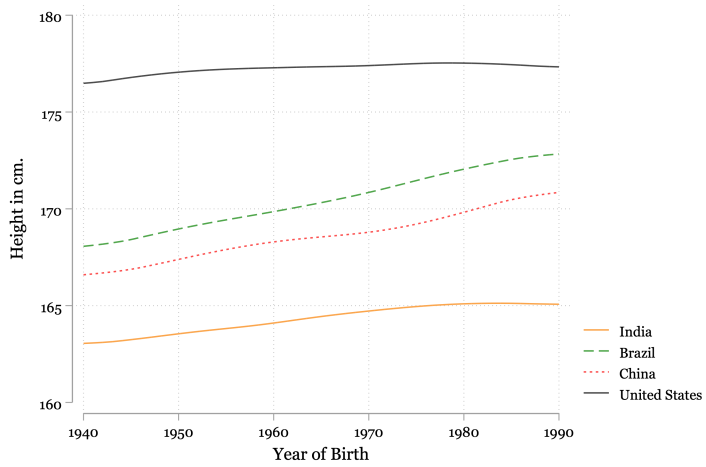
(B) Men's Income



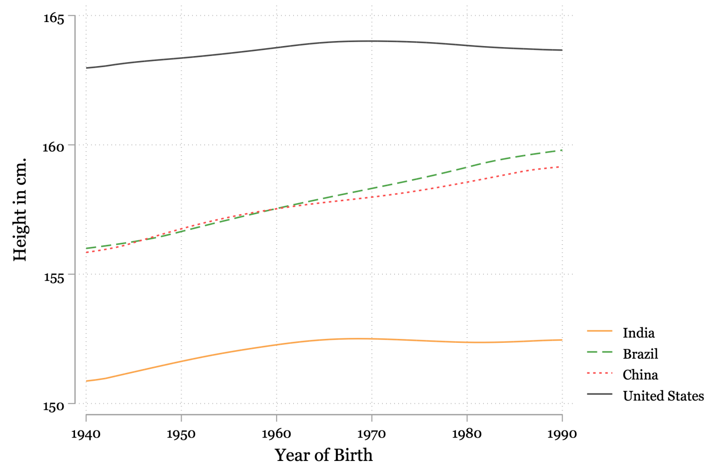
(C) Men's Education

Note: Sample consists of men who are household heads and in the age group 18 to 65.

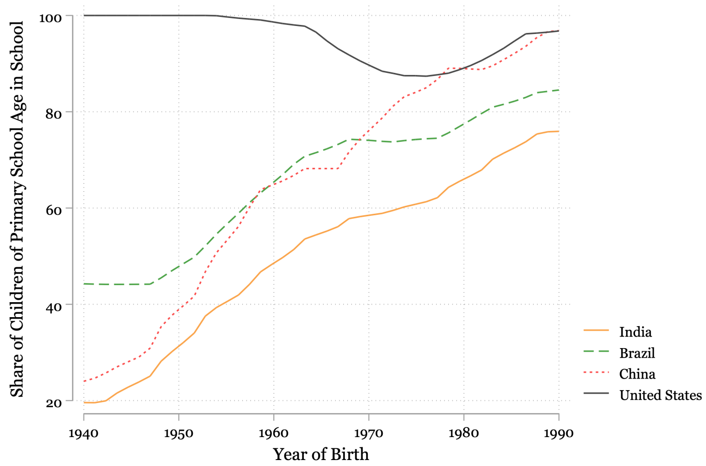
Figure A25: Men and Women's Height and Education over Time



(A) Men's Height



(B) Women's Height



(C) Education

Note: Data Source: Our World in Data:Height Data (Max Roser and Ritchie, 2013) and Education Data (Roser and Ortiz-Ospina, 2013)

Table A7: Affinity Matrix Estimates: Upper Caste

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.31***	-0.009	0.072*	0.025	0.236***
	(0.04)	(0.052)	(0.048)	(0.039)	(0.054)
Education	0.044	0.937***	0.229***	0.011	0.039
	(0.052)	(0.078)	(0.063)	(0.054)	(0.073)
Father Education	-0.047	0.177***	0.286***	0.035	0.105*
	(0.047)	(0.065)	(0.053)	(0.044)	(0.066)
Mother Education	0.014	0.104*	0.038	0.168***	-0.097
	(0.043)	(0.064)	(0.047)	(0.033)	(0.061)
Age	-0.014	-0.002	-0.113	0.062*	1.28***
	(0.049)	(0.067)	(0.06)	(0.048)	(0.083)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Upper Caste, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A8: Affinity Matrix Estimates: Other Backward Caste

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.264***	0.053*	0.005	0.096**	0.103***
	(0.03)	(0.037)	(0.047)	(0.052)	(0.041)
Education	-0.004	0.882***	0.289***	-0.086	0.069
	(0.045)	(0.061)	(0.061)	(0.068)	(0.061)
Father Education	-0.041	0.046	0.416***	-0.009	-0.004
	(0.047)	(0.059)	(0.057)	(0.064)	(0.065)
Mother Education	-0.039	0.341***	-0.106	0.258***	-0.01
	(0.052)	(0.067)	(0.055)	(0.05)	(0.07)
Age	0.048	0.23***	-0.145	0.095*	1.337***
	(0.039)	(0.049)	(0.059)	(0.063)	(0.066)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Other Backward Caste, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A9: Affinity Matrix Estimates: Scheduled Tribe

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.232***	0.081	-0.069	0.05	0.198**
	(0.07)	(0.1)	(0.153)	(0.167)	(0.104)
Education	-0.034	1.052***	0.575***	-0.003	0.13
	(0.13)	(0.166)	(0.215)	(0.244)	(0.183)
Father Education	-0.062	0.354**	0.568***	-0.218	-0.254
	(0.165)	(0.2)	(0.236)	(0.255)	(0.236)
Mother Education	0.065	0.048	-0.071	0.714***	0.265
	(0.174)	(0.252)	(0.257)	(0.242)	(0.225)
Age	-0.044	0.295***	0.058	-0.183	1.555***
	(0.089)	(0.116)	(0.204)	(0.206)	(0.157)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Scheduled Tribe, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A10: Affinity Matrix Estimates: Scheduled Caste

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.319*** (0.047)	-0.079 (0.057)	0.099 (0.091)	0.141* (0.1)	0.271*** (0.061)
Education	0.068 (0.082)	1.017*** (0.102)	0.13 (0.122)	0.241** (0.123)	-0.008 (0.104)
Father Education	-0.072 (0.091)	0.362*** (0.108)	0.472*** (0.12)	-0.149 (0.124)	0.199** (0.116)
Mother Education	0.101 (0.122)	0.044 (0.156)	-0.126 (0.142)	0.315*** (0.104)	-0.113 (0.155)
Age	-0.02 (0.059)	0.073 (0.072)	0.063 (0.116)	0.009 (0.121)	1.331*** (0.095)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to the Scheduled Caste, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A11: Affinity Matrix Estimates: Muslims

	Woman				
Man	Height	Education	Father Education	Mother Education	Age
Height	0.272*** (0.058)	0.047 (0.075)	0.061 (0.092)	0.006 (0.089)	0.097* (0.074)
Education	0.144* (0.097)	0.885*** (0.126)	0.143 (0.131)	0.074 (0.119)	-0.062 (0.128)
Father Education	0.083 (0.103)	0.16 (0.131)	0.441*** (0.118)	-0.169 (0.128)	0.276** (0.133)
Mother Education	-0.132 (0.125)	0.178 (0.157)	-0.11 (0.153)	0.369*** (0.121)	-0.163 (0.169)
Age	0.009 (0.076)	0.003 (0.103)	0.1 (0.129)	-0.173 (0.122)	1.329*** (0.128)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to Muslims, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A12: Affinity Matrix Estimates: Other Backward Caste

	Woman			
Man	Height	Education	Wealth	Age
Height	0.531*** (0.078)	-0.182 (0.061)	-0.002 (0.084)	0.407*** (0.096)
Education	0.007 (0.076)	0.678*** (0.084)	0.295*** (0.103)	0.159* (0.118)
Wealth	-0.02 (0.088)	-0.109 (0.102)	0.298*** (0.082)	0.09 (0.181)
Age	-0.029 (0.115)	0.32*** (0.113)	0.182* (0.141)	2.854*** (0.232)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to Other Backward Caste, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A13: Affinity Matrix Estimates: Upper Caste

	Woman			
Man	Height	Education	Wealth	Age
Height	0.612***	-0.314	0.228***	0.416***
	(0.121)	(0.117)	(0.097)	(0.167)
Education	-0.068	1.064***	0.218***	-0.137
	(0.103)	(0.15)	(0.088)	(0.158)
Wealth	0.083	0.04	0.39***	0.256**
	(0.084)	(0.117)	(0.073)	(0.142)
Age	-0.159	0.037	-0.156	2.33***
	(0.146)	(0.163)	(0.135)	(0.285)

Note: Each element represents the complementarity/ substitutability between two features. Data consists of individuals belonging to Upper Caste, with men between the age of 50 and 60 and women between the age of 40 and 60.
Standard errors are in the parentheses.
* p<0.1, ** p<0.05, *** p<0.01

Table A14: Joint Utility Share Explained: Other Backward Caste

	Index 1	Index 2	Index 3	Index 4	Index 5
Share of joint utility explained	67.63	16.63	7.91	4.89	2.94
Standard deviation of shares	0.81	1.27	0.88	0.76	3.71

NOTES: The table shows the five indices, Index 1 - Index 5, explaining mutually exclusively shares of the joint marital utility using the Saliency analysis technique in Dupuy and Galichon (2014). Standard errors are calculated using 500 bootstrap replications. Coefficient in red shows the share of joint marital utility explained by height of the man and the woman.

Table A15: Loading Matrix:Other Backward Caste

	Index 1		Index 2		Index 3		Index 4		Index 5	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Height	-0.09	-0.07	0.09	0.17	0.92	0.83	-0.36	-0.51	0.00	-0.15
Education	-0.08	-0.16	0.94	0.94	-0.04	-0.24	0.14	-0.01	-0.31	-0.19
Father Education	-0.03	0.05	0.14	0.24	0.24	0.47	0.66	0.70	0.70	0.47
Mother Education	0.00	-0.03	0.30	0.11	-0.28	-0.17	-0.65	-0.49	0.64	0.85
Age	-0.99	-0.98	-0.08	-0.16	-0.08	0.01	0.00	0.09	0.01	0.04

NOTES: The table shows indices of mutual attractiveness (Dupuy and Galichon, 2014) for men and women as a weighted sum of marriage market features. The total share of joint utility explained by an index is given in table A14. Coefficient in red corresponds to the index of mutual attractiveness for men and women with the largest weight corresponding to height.

Table A16: Joint Utility Share Explained: Scheduled Caste

	Index 1	Index 2	Index 3	Index 4	Index 5
Share of joint utility explained	65.22	18.86	8.27	5.22	2.43
Standard deviation of shares	1.09	1.67	1.07	0.84	4.36

NOTES: The table shows the five indices, Index 1 - Index 5, explain mutually exclusively shares of the joint marital utility using the Saliency analysis technique in Dupuy and Galichon (2014). Standard errors are calculated using 500 bootstrap replications. Coefficient in red shows the share of joint marital utility explained by height of the man and the woman.

Table A17: Factor Loadings: Scheduled Caste

	Index 1		Index 2		Index 3		Index 4		Index 5	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Height	-0.09	-0.03	0.13	0.07	0.97	0.88	-0.18	-0.47	-0.03	0.02
Education	-0.08	-0.12	0.94	0.95	-0.16	-0.12	-0.14	-0.09	-0.25	-0.24
Father Education	0.00	0.05	0.29	0.27	0.11	0.28	0.72	0.59	0.62	0.71
Mother Education	0.01	-0.04	0.08	0.05	-0.11	-0.37	-0.65	-0.65	0.75	0.66
Age	-0.99	-0.99	-0.09	-0.10	-0.08	0.02	0.02	0.08	0.03	0.04

NOTES:The table shows indices of mutual attractiveness (Dupuy and Galichon, 2014) for men and women as a weighted sum of marriage market features. The total share of joint utility explained by an index is given in table A16. Coefficient in red corresponds to the index of mutual attractiveness for men and women with the largest weight corresponding to height.

Table A18: Joint Utility Share Explained: Scheduled Tribe

	Index 1	Index 2	Index 3	Index 4	Index 5
Share of joint utility explained	61.07	21.66	10.96	4.29	2.02
Standard deviation of shares	1.81	2.67	1.72	1.31	6.10

NOTES: The table shows the five indices, Index 1 - Index 5, explain mutually exclusively shares of the joint marital utility using the Saliency analysis technique in Dupuy and Galichon (2014). Standard errors are calculated using 500 bootstrap replications. Coefficient in red shows the share of joint marital utility explained by height of the man and the woman.

Table A19: Factor Loadings: Scheduled Tribe

	Index 1		Index 2		Index 3		Index 4		Index 5	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Height	-0.02	-0.03	-0.08	-0.00	0.98	0.99	-0.15	-0.08	0.00	0.09
Education	0.00	0.04	0.99	0.96	0.08	0.03	0.00	-0.06	0.09	-0.26
Father Education	0.09	-0.05	-0.02	0.13	0.15	0.05	0.98	0.95	0.03	0.28
Mother Education	-0.09	0.08	0.09	-0.24	0.02	0.11	0.04	0.30	-0.99	-0.91
Age	0.99	0.99	0.01	-0.02	0.01	0.02	-0.09	0.02	-0.10	0.10

NOTES: The table shows indices of mutual attractiveness (Dupuy and Galichon, 2014) for men and women as a weighted sum of marriage market features. The total share of joint utility explained by an index is given in table A18. Coefficient in red corresponds to the index of mutual attractiveness for men and women with the largest weight corresponding to height.

Table A20: Joint Utility Share Explained: Muslims

	Index 1	Index 2	Index 3	Index 4	Index 5
Share of joint utility explained	58.34	19.56	11.11	8.49	2.50
Standard deviation of shares	1.60	2.40	1.67	1.38	5.50

NOTES: The table shows the five indices, Index 1 - Index 5, explain mutually exclusively shares of the joint marital utility using the Saliency analysis technique in Dupuy and Galichon (2014). Standard errors are calculated using 500 bootstrap replications. Coefficient in red shows the share of joint marital utility explained by height of the man and the woman.

Table A21: Factor Loadings: Muslims

	Index 1		Index 2		Index 3		Index 4		Index 5	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Height	0.16	0.03	0.03	-0.09	0.91	0.92	0.35	0.35	0.13	-0.13
Education	0.06	0.15	-0.94	-0.96	0.09	-0.05	-0.06	-0.16	-0.32	-0.15
Father Education	-0.09	-0.17	-0.28	-0.22	-0.32	-0.27	0.67	0.82	0.60	0.42
Mother Education	0.06	0.07	-0.19	-0.08	0.14	0.26	-0.64	-0.38	0.73	0.88
Age	0.98	0.97	0.04	0.12	-0.19	-0.08	0.05	0.18	0.01	0.04

NOTES:The table shows indices of mutual attractiveness (Dupuy and Galichon, 2014) for men and women as a weighted sum of marriage market features. The total share of joint utility explained by an index is given in table A20. Coefficient in red corresponds to the index of mutual attractiveness for men and women with the largest weight corresponding to height.

Table A22: Joint Utility Share Explained: Upper Caste

	Index 1	Index 2	Index 3	Index 4	Index 5
Share of joint utility explained	52.95	24.78	12.13	6.89	3.26
Standard deviation of shares	1.27	2.82	1.87	1.59	4.65

NOTES: The table shows the five indices, Index 1 - Index 5, explain mutually exclusively shares of the joint marital utility using the Saliency analysis technique in Dupuy and Galichon (2014). Standard errors are calculated using 500 bootstrap replications.

Table A23: Factor Loadings: Upper Caste

	Index 1		Index 2		Index 3		Index 4		Index 5	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Height	0.02	-0.02	-0.02	0.03	-0.53	-0.49	0.72	0.81	-0.44	-0.31
Education	0.11	0.18	0.90	0.96	-0.08	0.11	0.19	0.09	0.37	0.16
Father Education	0.01	-0.07	0.10	0.18	-0.77	-0.80	-0.62	-0.55	-0.08	-0.16
Mother Education	0.08	0.09	0.41	0.11	0.34	0.32	-0.24	-0.17	-0.81	-0.92
Age	0.99	0.98	-0.13	-0.17	0.00	-0.12	-0.01	-0.03	0.03	0.04

NOTES: The table shows indices of mutual attractiveness (Dupuy and Galichon, 2014) for men and women as a weighted sum of marriage market features. The total share of joint utility explained by an index is given in table A22.