I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17366

English Language Premium in a Marriage Market: Experimental Evidence from Delhi

Tomohisa Miyamoto
Arjun S. Bedi

OCTOBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17366

English Language Premium in a Marriage Market: Experimental Evidence from Delhi

Tomohisa Miyamoto
Aflatoun International

Arjun S. Bedi
Erasmus University Rotterdam and IZA

OCTOBER 2024

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

English Language Premium in a Marriage Market: Experimental Evidence from Delhi*

In India, fluency in English remains a status symbol for the wealthy and educated middle class, while creating a divide between those who have strong English language skills and those who do not. This study deploys a correspondence experiment conducted on a matrimonial website, to examine returns to English language skills for females in a marriage market in Delhi. Indicating strong English language skills in a marriage profile increases the number of views and interests received by 20% and 38%, respectively. Rather than relying on a purely passive approach of receiving attention from unknown users, in a more active stage of the experiment, expressions of interest were sent from female profiles to randomly assigned male profiles. While there is no overall effect of English language proficiency on the interest response rate, language fluency makes a difference for some subgroups. For lower caste women the English language premium is 13 percentage points (43%) and while both lower and upper caste males are more likely to respond to lower caste women with strong English language skills, the effect is substantially larger, 14.3 percentage points or 47%, in the case of upper caste males. The results suggest that English language proficiency expands the set of potential marriage partners and may help promote inter-caste alliances.

JEL Classification: J64, J65

Keywords: English, return to language skills, caste, social mobility, marriage market, field experiment

Corresponding author:
Arjun S. Bedi
International Institute of Social Studies (ISS)
Erasmus University Rotterdam
Kortenaerkade 12
2518 AX Den Haag
The Netherlands
E-mail: bedi@iss.nl

* This study was supported by a grant from the ISS and an earlier version was the first author's MA thesis submitted to the ISS. The MA thesis has appeared as ISS Working Paper No. 699. We thank Lorenzo Pellegrini for useful comments and guidance and Ms. Smiti from Jawaharlal Nehru University for providing outstanding research assistance. The study protocol was reviewed and approved by an ethics review board established by Erasmus University Rotterdam (ETH2324-0094). The study was pre-registered at the Wharton Credibility Lab, University of Pennsylvania (#72623).

1. Introduction

In the 2017 film "Hindi Medium", one of the last films of the well-known actor Irrfan Khan, a couple in Delhi disguise themselves as a poor family in an attempt to get their daughter admitted to a posh English medium school, even moving to a poor neighbourhood (Chaudhary, 2017).[1] At all costs, the couple wants to avoid enrolling their daughter in a government school where Hindi is the language of instruction. In a pivotal scene of the film, the wife exclaims *"She won't learn anything in a government school. She'll be terrified if anyone speaks in English. She will be a misfit in society. English isn't just a language in this country, it's a 'class'!*"

In India as well as in other countries, English is considered an essential skill to access better jobs and enhance income, underlining the popularity of English medium private schools. Prior research in India has shown that returns to education in English are higher than returns to education in local languages (Munshi and Rosenzweig, 2006; Chakraborty and Bakshi, 2016), and that there is a strong correlation between subjective English proficiency and a person's income (Azam, Chin and Prakash, 2012). Similar results have been found in other countries and the literature tends to show that knowledge of English and other dominant languages raises earnings.[2,3]

[1] The couple attempt to enroll their child in a private school using the admissions quota reserved for the poor, under the Right to Education Act (RTE). This policy is detailed in Chudgar and Creed (2016), Rao (2019), and Joshi (2020).

[2] There is a large body of work on the economic returns to language skills which shows that knowledge of English and other dominant languages raises one's earnings - English in the US (Chiswick and Miller, 1995; Bleakley and Chin, 2004), UK (Aoki and Santiago, 2015), Israel (Lang and Siniver, 2006), South Africa (Levinsohn, 2004), Kazakh and Russian in Kazakhstan (Aldashev and Danzer, 2014), English, Russian, French, German, and Arabic in Turkey (Di Paolo and Tansel, 2015). Similarly, several studies have found that being taught in a European/foreign language medium translates into higher labor market returns as compared to being taught in the local language in many contexts - French versus Arabic in Morocco (Angrist and Lavy, 1997), English versus Spanish in Puerto Rico (Angrist, Chin and Godoy, 2006), English versus Marathi in India (Munshi and Rosenzweig, 2006), English versus Bengali in India (Chakraborty and Bakshi, 2016), English versus Malay in Malaysia (Parinduri and Ong, 2018).

In contrast to the body of work on the link between English and labor market returns, research on non-monetary returns to English language proficiency is scarce and we were unable to find studies that identify the causal influence of English language skills on social outcomes.[4] Additionally, in the Indian context, where women are less likely to be engaged in wage labor, it is pertinent to examine social returns to English.

To explore social returns to English, this paper focuses on a specific setting, that is, returns to English language skills in the marriage market in Delhi. The questions this study seeks to address are: Does English language proficiency translate into greater attention in the marriage market and whether English proficiency enables matching with higher status (caste) partners? To respond to these questions, the paper conducted a correspondence experiment on a major Indian matrimonial website, which is one of the primary means of finding a marriage partner in India. The experiment consisted of two stages. In the first, a passive phase, a set of fictitious female profiles with different levels of English proficiency (indicating strong proficiency or not), caste (lower or upper caste), educational background (high school, BA, MA), and annual income (four levels of income), were constructed and the number of visitors to the profile and the number of interests received from other profiles were used as outcomes. In the second, the active phase, an expression of interest was sent from the fictitious female profiles to existing (real) male profiles

[3] A related body of work examines the link between language of instruction and educational outcomes and the results are not as clear cut as the language-earnings link. For instance, in his work on Andhra Pradesh, Nair (2015) finds that children taught in Telugu perform better on math achievement tests versus children taught in English. Jain (2017) reaches a similar conclusion – that is, districts where there was a linguistic mismatch (that is, the official language did not match the district's language) recorded lower literacy and lower college graduation rates. In contrast, Parinduri and Ong (2018) find higher educational attainment for those taught in English versus Malay.

[4] Although they do not focus on language skills, a recent study by Ray, Roy and Sahai (2020) examines the link between educational attainment and the incidence of inter-caste marriage. Their analysis shows that the educational levels of the spouses has no effect on the probability of their marriage being inter-caste. However, the educational level of the groom's mother is positively associated with the probability of an inter-caste marriage.

and the acceptance of this interest was used to measure preferences for marriage partners. The experiment was conducted using profiles of middle-class, native Hindi-speakers residing in Delhi. From the information in the matrimonial profiles, the status of potential marriage candidates was observed in terms of caste, annual income, and educational background. Data collection was carried out between August and September 2021 and involved a total of 384 candidates.

The paper's econometric strategy exploits the exogenous nature of English proficiency in the fictitious profiles. For the first stage, the fictitious profiles were assigned three different educational backgrounds, four annual incomes, and two caste statuses, that is, a total of 24 profiles. For each of these 24 profiles another 24 were created which signaled strong English language proficiency. Thus, the profiles with and without English proficiency were, on average, the same with respect to other attributes, and there was no correlation between any of the assigned attributes and English language skills. Therefore, the difference in the number of visitors and interest received by the two sets of profiles (those with and without an indication of English proficiency) may be interpreted as the causal effect of indicating English proficiency in the profile. In the second stage, each of the 48 fictitious female profiles expressed "interest" in eight real male profiles. The female interests were randomly allocated to existing male profiles. Unlike matching based on traits, in the experiment, the traits of the female profile are not correlated with the traits of the male profile that receives the interest. Consequently, the difference in interest acceptance rate may be interpreted as an effect of the feature manipulated in the fictitious profile, in this case English language skills, without being affected by other profile attributes.

The results of the experiment show that English language proficiency translates into a substantial advantage for prospective brides in the marriage market. Female profiles with English

proficiency received 20% more visitors and 38% more interest than female profiles without English proficiency. English proficiency has the effect of attracting a larger number of potential marriage partners and widening the choices faced by prospective brides.

Furthermore, signaling English proficiency increased the probability of matching with high-caste men for low-caste women. Interest acceptance rates were lowest in the caste hypergamy combinations where the female side was low caste, and the male side was high caste. However, the probability of an interest being accepted increased by 14.3 percentage points when English proficiency was mentioned in the female profile. English proficiency may help to undermine caste barriers by increasing the likelihood of inter-caste marriage with high-caste men, compensating for the disadvantages of low-caste status.

The paper unfolds by providing in section 2, contextual information on the use of English in India and the links between caste and the marriage market. Section 3 provides details on the experiment, outlines the hypotheses and the empirical approach. Section 4 describes the data. Section 5 reports the results, and the final section concludes the paper.

2. On language, caste and marriages

This paper deals with the role of language and caste in a marriage market. Before setting out the experimental details, this section focuses on providing relevant contextual information pertaining to three elements, that is, language, caste and the specific features of the marriage market in India.

2.1 English in India

In India, a multilingual society, Hindi is the official language at the federal level as stated in Article 343 of the Indian Constitution. However, since the adoption of the Indian constitution in 1950, English has retained its status as an additional official language (Sandhu, 2014).

Furthermore, as of 2024, the Eighth Schedule to the Constitution of India recognises 22 official languages, although English and Hindi remain the primary means of communication between states and between the states and the federal government.

Since the British colonial period, the language of instruction in higher education institutions has been English, and to date this continues. Consequently, English language skills are perceived as essential for obtaining white-collar jobs in both the private and the public sector. In its education policies, the government of India has emphasised acquisition of English language skills as part of a three language formula, along with Hindi and a state language (Jain, Maitra and Mani, 2019). While emphasizing the importance of a local language in early education (up to Grade 5), the National Education Policy 2020 also recognizes and calls for the continued implementation of the three-language formula, provided that two of the three languages are native to India.

As may be expected, given the importance of English in higher education and in the labour market, alongside Indian languages, the use of English continues to flourish in India albeit as a second or third language.[5] Notwithstanding the importance of English, the percentage of the population which speaks English fluently is perhaps much lower than maybe expected. According to Azam, Chin and Prakash (2012), based on the India Human Development Survey 2005, only 4% of adults aged between 18-65 can converse in fluent English and 14% can converse a little in English. The proportion of people who speak some English is higher for men (26%) than women (14%) and in urban areas (35%) than rural areas (14%). English language proficiency is strongly associated with higher education – while 89% of individuals with a

[5] In IHDS 2005, only 0.2% of respondents listed English as their mother tongue. According to the Indian census conducted in 2011, only 0.02% listed English as their mother tongue. However, this figure rises to 10.6% if one includes English as a first, second, or third language. The corresponding figure for Delhi is 26.5% (Office of the Registrar General & Census Commissioner, India, 2022).

bachelor's degree can speak English the corresponding rate is 56% amongst those who have completed secondary schooling.[6] Albeit based on a non-representative sample of test-takers interested in learning English, according to Education First's English Proficiency Index 2023, on average, English fluency in India is considered moderate with a score of 504 out of 800 and in Delhi the score is 451 out of 800 or low (Education First, 2023).[7] In short, despite its widespread use in India, English language fluency is still likely to be a valued trait in the labour market, and in a variety of other settings (engaging with government officials/government documents, social interactions).

2.2 Caste and marriage

The origins of the caste system, a system of social classification, lie in Hindu scriptures. Caste is characterized by three core elements: (i) it is a hereditary group based on occupation (ii) it is a highly endogamous group and (iii) it has a hierarchical structure (Bidner and Eswaran, 2015). Caste consists of broad groups – *varna* (class), and *jati* (subgroups). In order of hierarchy, *varna* consists of Brahmins (priests), Kshatriyas (warriors), Vaishyas (merchant), and Shudras (laborers). A fifth category, which is considered outside the scope of the caste system are tribal groups (Adivasis) and Adi Shudras or untouchables, often also referred to as Dalits/Scheduled Castes (SC). Within each category, there are numerous, about 4,000, according to Munshi (2019), subgroups based on traditional occupational groups with their own hierarchy. For the purposes of this paper, we distinguish between two broad categories of caste - Brahmins who are referred to as upper castes (UC) and Scheduled Castes referred to as lower caste (LC).

[6] Azam, Chin and Prakash (2012) mention that it is possible to complete secondary school or college education without being able to speak English, as exams may be written in Hindi or other state languages except in science and engineering.

[7] According to the index, a score of low implies that an individual is able to navigate an English-speaking country as a tourist, engage in small talk with colleagues, understand simple emails from colleagues.

While the caste system has implications in terms of access to education and labour markets (occupational choice), the primary concern of this paper is the link between caste and marriage. Unlike the association between caste and education and caste and occupation which have both weakened over time (Munshi and Rosenzweig, 2006; Hnatkovska, Lahiri and Paul, 2012; Asher, Novosad, and Rafkin, 2021), not least because of affirmative action policies, caste endogamy remains strong (Munshi and Rosenzweig, 2006; Banerjee et al., 2013; Munshi, 2019). Caste grouping and sense of belonging to a caste have been reproduced and maintained through adherence to caste endogamy (Allendorf and Pandian, 2016; Titzmann, 2018; Ray, Roy and Sahai, 2020).[8] Consequently, inter-caste marriages are limited. According to data from a National Family Health Survey conducted in 2006, only 11% of married couples were in an inter-caste marriage (Allendorf and Pandian, 2016). Ray et al., (2020) report an even lower percentage of inter-caste marriages (5.82%) based on data from the India Human Development Survey conducted in 2011.

The reason for the traditional abstinence from inter-caste marriage is partly explained by the punishment for inter-caste marriage. Couples in an inter-caste marriage may face persecution and ostracism and may lose community network benefits. However, this punishment is gender asymmetric (Bidner and Eswaran, 2015). When a man and a woman from different castes marry, the woman's caste is usually assumed to change to the man's caste. If the male is of a higher

[8] From an economic perspective, marriages within the same community, including caste endogamy, have been explained as an attempt to maintain networks and as a form of insurance. Rosenzweig and Stark (1989) argue that women's migration through marriage to other villages with different agricultural risks is a mutually insuring contract for both families. They found that when women migrate to a distant village through marriage in rural South India, there is a consumption smoothing effect for both the woman's parents and the groom's household. Further, Munshi and Rosenzweig (2009) pointed out that intra-caste marriages maintain community networks and play a role in risk-sharing, and that is why groups with higher average income are less likely to marry outside the community. From these findings, intra-caste marriage may be viewed as a strategic decision for households.

caste (caste-hypergamy), the woman gains a higher caste status through marriage (Ahuja and Ostermann, 2016; Titzmann, 2018).

In contrast to the actual lack of inter-caste marriages, there appears to be a greater tolerance for inter-caste marriages. Based on data from an Indian National Election Study conducted in 2004, Ahuja & Ostermann (2016) report that while 60% of 27,000 voters supported a ban on inter-caste marriage, the rate was 47% in urban areas. Based on data compiled from 10,000 newspaper matrimonial advertisements they report that the number of advertisements limiting marriage partners to the same caste declined from 30% in the 1970s to 19% in the 2000s. Finally, based on an experimental approach which involved 1,070 female participants, Ahuja and Ostermann (2016) report that 70.7% of SC respondents expressed an interest in inter-caste marriage as compared to 53.9% amongst UC individuals.

2.3 Arranged and love marriages

In addition to caste endogamy, another relevant feature of the Indian marriage market is the high proportion of arranged marriages. This refers to marriages in which parents or other relatives play a predominant role in choosing a partner for their children, taking into account social attributes such as caste and family economic conditions (Banerjee et al., 2013). According to Ray et al. (2020), among married couples sampled in IHDS 2011, 73% responded that their marriages had been arranged by their parents, and 70% of married women responded that they first met their marriage partner on the day of the marriage. Notably, although only 5.82% of the marriages were inter-caste, 63% of these were arranged marriages. Based on more recent data (about 4,000 households) collected in 2016 in Delhi, Reed (2019) displays the persistence of arranged marriage and the almost total absence of the expectation of a love marriage amongst never-married young persons in the sample.

In contrast to arranged marriages, in which parents and relatives play a primary role in the decision-making process, marriages in which couples decide whom to marry are referred to as "love marriages". However, arranged marriages do not mean that the will of the couple is not respected or that romantic feelings do not exist. It implies marriages in which emotional affection is not a prerequisite, and in which "gaining stability from its embedding in the established networks of social and kinship ties" (Polzenhagen and Frey, 2017) is a priority.

Typically, in the context of arranged marriages, individuals or families search for marriage partners from amongst a limited pool of people who belong to the same religion and same or similar caste, considering various pertinent traits (income, educational background, family background, language, appearance, and personality). To mitigate the costs of such a complex search for a partner, various means of finding a partner have been developed. Information on people of marriageable age is compiled into a profile/marriage CV/matrimonial bio, and advertised through community networks, matchmakers in the caste community, marriage bureaus, newspaper advertisement, and/or matrimonial websites. Newspaper advertisement is a traditional method with a history of more than 300 years (Polzenhagen and Frey, 2017), and matrimonial websites are a more recent tool to aid this search process.

While exact numbers are hard to acquire, an early estimate (Trivedi, 2014) suggested that there were more than 20 million users on 150 different matrimonial websites. However, there are perhaps three major English language websites with a pan-Indian reach – these are Jeevansathi.com, Bharatmatrimony.com, and Shaadi.com.[9] On such websites, registration is free and basic services (browsing profiles, shortlisting and sending interest, messaging and chatting) are free while premium services command a price. The registrant creates a profile which

[9] Shaadi.com claims to have "touched more than 50 million lives". BharatMatrimony.com claims to have served 40 million customers. As of September 2024, Jeevansathi.com claimed to have more than 1 million registered profiles.

typically includes details on caste, educational background, mother tongue, occupation, lifestyle, and family background. This profile functions as an advertisement for those who are looking for partners on the same website. Users may browse the profiles of other registrants conditional on various traits, and if they like a profile, they may request to be contacted.

While technology aids the search process, families (parents, relatives) continue to play an important role in finding a marriage partner through online matrimonial websites. Articles in the popular press often point out that such online matchmaking gives the potential bride/groom the privacy to escape the eyes of family and conduct a more individualistic search (Bhandari, 2020). However, like newspaper advertisements, some of the profiles on matrimonial websites are posted by parents, relatives or friends and at times the profile administrator is also someone other than the registered individual.[10] As Bhandari (2020) points out, many of the items in the profile are about family, and often the free text section is also heavily weighted towards statements about family. Furthermore, Titzmann (2018) argues that caste endogamy has been made easier because information about caste, income, and family are clearly marked as items on the websites.

[10] About 22% of the profiles identified in this experiment clearly indicated that the profile administrator was someone other than the individual.

3. The experiment and hypotheses

This experiment was implemented on a major matrimonial website in India. This section provides information on the set up of the matrimonial website, the construction of fictitious profiles, the data collection procedure and finally the hypotheses and empirical specifications.

3.1 Matrimonial website

There are several pan-Indian, English language matrimonial websites operating in the country. Most of these offer similar services and we randomly chose one of these websites. Like other sites, the chosen website offers the possibility of free registration of profiles. Once registered, clients may search and browse profiles according to their desired criteria and may contact a profile by sending an "interest" in the form of a text message, e-mail, or a phone call. While there are no costs for registering on the website and using the basic functions, the site generates revenue through registration fees for premium accounts. Premium accounts get priority customer service, unlimited access to photos of profiles, access to contact information (phone number and email address) and may directly send messages to other profiles. On the website chosen for this study, the primary means for non-premium users to contact other profiles is through sending their "interest". A visitor to a profile may communicate their desire to contact a particular profile by clicking the "Send Interest" button. A user who receives the interest decides whether to "Accept" the interest after reviewing the sender's profile. If the interest is "Accepted", both parties may chat via text messages and exchange contact information. Additional details including the registration process, verification and the details that need to be included in the profiles are provided in Appendix 1.

3.2 Creating fictitious profiles

To create the fictitious female profiles, we began by inputting the same set of values for all profiles, except for English proficiency, caste, income and education level. To elaborate, in all the profiles, gender was set to female, profile manager was the individual, mother tongue was Hindi-Delhi, religion was Hindu, sub-caste was left blank, marriage status was set to "Never Married", height was set to 5 feet, which is about the average height of Indian women above the age of 19, and age was set to 25 years. The place of residence and origin was set to Delhi, working in the private sector, in banking.[11] The family was listed as nuclear also living in Delhi with her father working in the private sector, mother is a housewife, and one married brother. Her preferred marriage partner is a Hindu, but she does not specify a preferred caste (caste no bar). Her lifestyle includes occasional drinking but no smoking. Her date of birth was randomly assigned from dates between September 10, 1995 and August 10, 1996 and pin codes were randomly assigned from a list of Delhi pin codes.[12] For the section of the profile which requires free text/self-introduction, several variations of the standard content were created, and one of these was randomly selected to be included in each profile. An example of a fictitious profile registered on a website is included in Appendix 3. Each profile had a different name, but this name was not revealed to other users on the website. Therefore, the only information regarding a profile's caste status (not specifying subcaste) was transmitted through the item of caste in the profile as discussed below. No profile pictures were included in fictitious profiles, which is common on the website especially for female profiles for privacy reasons.

[11] Although all profiles indicated that women worked in the banking sector there were four income categories, including zero income and up to Rs. 7.5 lakhs per annum. The zero-income category indicates that the woman is working but in an unpaid position which may be viewed as weaker attachment to the labor market.

[12] The dates of birth were randomly selected from dates between September 30, 1995 and August 10, 1996, using the RANDBETWEEN function in Excel, so that the age at the time of the experiment was 25 years - "RANDBETWEEN("1995/9/30","1996/8/10")". Pin code was also randomly assigned from a list of Delhi pin codes (https://dmsouthwest.delhi.gov.in/std-pin-codes/) using Excel's RANDBETWEEN function.

After specifying these common elements, combinations of the four attributes - English proficiency, caste status, income, and education were added to the profiles. The categories for each item are shown in Table 1. First, we assigned two types of castes to the fictitious profiles: upper caste (UC) and lower caste (LC). On the website, if a user selects Hindu in the religion section, one must select a caste (Brahmin, scheduled caste) and one can further specify the subcaste. We did not specify the subcaste to maximize the number of visitors to the profile when other users sort profiles by caste, and to avoid any possible effect based on the status of a specific subcaste. Second, we assigned four categories of income: zero, lower middle income (annual income of Rs. 1-2 lakhs (Euro 2,300)), middle income (Rs. 3-4 lakhs (Euro 4,600)), and upper middle income (Rs. 5-7.5 lakhs (Euro 6,800)).[13] Third, fictitious profiles were assigned one of three educational categories - high school graduate, bachelor's degree, or master's degree and above. Although the platform allows users to choose the level and the subject of the degree (B. Ed, MSc in Engineering and so on), to avoid any effect from the subject of the degree, we only indicate BA or MA degree.

[13] Middle class is a self-identified concept (Bhandari, 2020), and there is no clear definition of its income level. At the all-India level, Shukla (2010) defines middle class income as annual household income of Rs. 2 lakhs to 10 lakhs, while Krishnan and Hatekar (2017) use daily consumption of Rs. 148-742 (annual income of Rs. 0.54-2.7 lakhs) to define middle class. At the all-India level, based on combining data from different sources, Bharti et al. (2024) report that in 2022-23 the average annual income of an Indian adult was Rs. 2.35 lakhs with 90% of the adult population earning less than Rs. 2.91 lakhs. It has been difficult to get a credible idea of Delhi's income distribution. According to the Economic Survey of Delhi, the per capita income in the state was estimated at Rs. 3.54 lakhs in fiscal year 2020-2021 (https://delhiplanning.delhi.gov.in/sites/default/files/Planning/2._state_economy.pdf). Based on this estimate of Delhi's per capita income we settled on an annual income of Rs. 3-4 lakhs as middle income. At the all-India level in 2022-23 (Bharti et al. 2024), the average adult income was Rs. 2.35 lakhs and the threshold for being in the bottom 90% of the population was Rs. 2.91 lakhs. Extrapolating to Delhi – in 2022-23, Delhi had a per capita income of Rs. 4.4 lakhs (https://delhiplanning.delhi.gov.in/planning/economic-survey-delhi-2022-23), and assuming that it had the same distribution as at the national level (ratio of 90th percentile to the mean is 1.24 = Rs. 2.91/2.35 lakhs), it is quite likely that a substantial proportion of Delhi's adult population had an annual income of less than Rs. 5.5 lakhs (1.24 x 4.4). The point is that the earnings of the bulk of the city's adult population may be expected to lie below the highest income bracket.

Finally, explicitly signalling English language skills in the profiles is the most important element of the experiment. In fictitious profiles with English language skills, we included English in the section of "languages known" in the profile, in addition to a local language (Hindi). Further, we mentioned in the self-introduction of the profile that the prospective bride has "outstanding English language skills, both written and spoken". In the profiles with no English skills, we do not mention English language skills in either the language section or the self-introduction. Mother tongue was set as "Hindi-Delhi" in all the profiles, and self-introduction was written in English in all the profiles.

3.3 Data collection

The process of data collection was divided into two stages and took place between August to September 2021. In the first (passive) stage, fictitious profiles were posted on the matrimonial website. When a profile is created it is browsable for all other existing users. Users may readily check how many other users have viewed their profile. The number of visitors to the profile and the number of interests received within a 72-hour period were recorded.[14]

The second (active) part of the experiment involves interaction with existing profiles. "Interest" was sent from fictitious profiles to existing male profiles and the acceptance rate was recorded (see Figure 1). At the beginning of this stage, for each of the 48 fictitious profiles we identified eight existing profiles, or a total of 384 (48 times 8) profiles, which fulfilled certain

[14] These numbers are a measure of the attractiveness of a profile in the online marriage market (see Ong and Wang, 2015).

criteria and recorded relevant information.[15] Key information such as income, educational background, caste, and English language skills were coded in the same manner as in the case of the fictitious profiles that were created. Next, we randomly assigned eight of the identified profiles to each of the 48 fictitious profiles. After assignation, "interests" were sent from each of the 48 fictitious profiles to the 8 assigned profiles. As an outcome variable, we recorded if the identified profiles accepted the "interest". Acceptance of interest indicates a willingness to explore a potential alliance.

For logistical and ethical reasons, we repeated identical waves of the experiment 4 times. To register on the matrimonial website, each profile requires a unique phone number. Therefore, we created 12 fictitious profiles with 12 phone numbers in one week. Each wave of the experiment lasted 5 days. In the first 72 hours, we recorded the number of visits to each profile. At the same time, we identified 96 existing profiles (8 real profiles for each of the 12 fictitious profiles per week) that met our criteria and prepared a dataset with coded attributes of these profiles. After 72 hours, we sent "interests" from each fictitious profile. Then, we waited for 48 hours to see if the interests were accepted. Each real profile received interest only once from a fictitious profile. If the interest was accepted, no further communication took place, and the account was deleted within 48 hours. To avoid any possible biases, each wave began on the same

[15] Male profiles that met our criteria were sampled from the website. Reflecting our research focus and to enhance the credibility of the expression of interest, the criteria used for the search were: age 25 years and above, Hindu, living in Delhi, annual income below Rs. 7.5 lakhs, Hindi as mother tongue, and never married men. We searched using the same criteria for each of the fictitious profiles and identified the top eight search results. If the same profiles appeared as matches for multiple fictitious profiles, the top eight excluding those profiles that had already been identified were used. We worked with eight profiles for pragmatic reasons. While trying to identify male profiles that met the criteria, it was difficult to find more than eight unique profiles for the 12 fictitious profiles per week (overall 8 x 12 = 96 profiles in one week).

day of a week. By repeating this process four times (4 weeks) we covered all 48 fictitious profiles. [16]

3.4 Hypotheses and empirical specifications

As a consequence of the various functions fulfilled by English in a multilingual society such as India, strong English language skills signal higher socioeconomic class (Faust and Nagar, 2001; Jayadeva, 2019), the perception of being intellectually superior as compared to Hindi speakers (Sandhu, 2014) and an educational background – specifically English medium education, which may be attractive in the marriage market.[17] While the literature does show that English language proficiency is rewarded in the labour market, even in a context where female labour force participation is low, it is likely to be perceived positively in terms of the role it may play in (social) interactions, in signalling a modern and international orientation (Bhandari, 2020) and in the transmission of inter-generational language skills.[18]

[16] Similar to other correspondence studies, there may be several ethical concerns. As in other studies of this genre, existing users of the matrimonial website are involved without informed consent and conducting such an experiment based on the construction of fictitious profiles may have a bearing on the credibility of the website. While acknowledging these concerns the experimental design was set up to minimize distortions. First, the resources required from users of the website is minimal (expressing interest is a one-click process). To minimize negative effects, each user received interest only once from a fictitious profile and the fictious profile was deleted within 48 hours of an interest being sent. With real profiles as well, there is no guarantee that communication takes places after interest is expressed – in short it is not unusual to experience a lack of response even after initial interest is expressed. To minimize the impact on the website, only 12 profiles were registered per week and the experiment was spread over four weeks.

[17] Bourdieu (1977) considered language not only as a tool of communication but also as a reflection of power relations. According to Bourdieu (1977) languages, dialects, and accents reflect the social status of the user, and he argued that by speaking a particular language or accent, a speaker can reassert their power and position. Consistent with this view, in their research in Lucknow, Faust and Nagar (2001) comment on how not being able to speak English was a strong complex even for a woman with a master's degree. Jayadeva's (2019) research on English medium education in low-fee private schools in Bangalore, points out that English is preferred as a medium of instruction at the expense of educational attainment because English is seen as a symbol of the middle class, and it is believed that being able to speak English is transformative for a student's future.

[18] In her work on matchmaking in Delhi, Bhandari (2020) uses the term exposure as a desirable trait in the marriage market. Exposure is captured, "*by one's knowledge (and access) to British and American television soaps, music, movies, ease in speaking English, and taste in international cuisines*" (page 97).

Just as other ways of enhancing social status (Srinivas, 1957; Ignatiev, 1995), when the language spoken differs by social class within a hierarchical structure, individuals may enhance their social status by mastering English and emulating the higher social classes (LaDousa, 2007). Consistent with this argument, for potential grooms with weak English language skills, strong English language skills in a potential bride may be viewed as a compensating trait which among other benefits may enhance their social status.[19] At the same time, for grooms with strong English language skills it may be attractive as it signals shared values and norms and potentially linguistic homogamy enables smoother communication leading to a better match (Fu, 2007; Jan, 2019; Bhandari, 2020). In her illuminating study on matchmaking in Delhi, a quote from one of the participants interviewed by Bhandari (2020) underlines the potential importance of linguistic matching which extends not just to the groom but the groom's family.[20] While commenting on the reasons for the lack of a match, a respondent comments (Bhandari, 2020, page 96) on the lack of English language skills of the potential groom's family, *"See, it is not only about money. I am sure they are well-off, but they are the non-urban, Hindi speaking type, and we are the proper Delhi type […] I did not think these things would matter until I encountered this situation and realised that differences come up in little things."* A male respondent who rejected a match articulated, *"I was shocked to know that she had never watched FRIENDS. How can this be? She grew up in Delhi not some village. How can she not know of the most popular series of our time. She simply did not know anything! I was really put off. Also, she did not speak very good*

[19] The "Indian" marriage market also includes the Indian diaspora who are looking for marriage partners of the same caste or ethnic background. A recent news article reported potential brides studying for the International English Language Testing System (IELTS) in order to enhance their chances of marrying men from Canada, the UK or the US, and migrating to these destinations (https://www.dw.com/en/ielts-marriages-indias-ideal-bride-is-proficient-in-english/a-53341947). Matrimonial websites have a section to indicate whether a user is willing to migrate internationally.

[20] Bhandari (2000)'s work on matchmaking is based on more than 100 in depth interviews with individuals in the age group 25 to 32.

English. She was a nice girl, don't get me wrong, but I think she was meant to fit a traditional Marwari setting. I was not looking for that. I wanted someone who can relate to me, who has exposure…" (Bhandari, 2020; page 98)

Motivated by these reasons, our first hypothesis which pertains to the passive stage of the experiment is that:

(*Hypothesis 1*) Female profiles signalling strong English language skills will receive a larger number of visitors and a larger number of expressions of interest than those without.

Since, on average, the two sets of fictitious profiles are the same except for the presence or absence of English proficiency, and there is no statistical association between English and the other attributes, the average difference in the popularity of the profiles may be interpreted as the impact of indicating strong English language skills.

Turning to stage 2, the active stage of the experiment in which expressions of interest are sent out by the fictitious female profiles to selected real male profiles. Based on the same reasons as articulated above, it may be expected that:

(Hypothesis 2) English proficiency increases the probability of receiving a response to an interest.

This hypothesis is examined by estimating the following model:

$$Accept_{ij} = \alpha + \beta_1 English_i + \beta_2 Caste_i + \beta_3 Education_i + \beta_4 Income + \gamma_1 English_j + \gamma_2 Caste_j + \gamma_3 Education_j + \gamma_4 Income_j + \varepsilon_{ij}\,. \quad (1)$$

In (1), $Accept_{ij}$ indicates whether real user j accepted the "interest" sent from a fictitious profile i, $English_i$ is a dummy variable indicating English language fluency, $Caste_i$ is a binary variable indicating UC, while both $Income_i$ and $Education_i$ are categorical variables denoting income and educational levels mentioned in the fictitious profiles. The variables with subscript j index the traits of the male profile receiving the expression of interest. In the set of fictitious profiles,

there is no systematic association between English proficiency and the other variables. Furthermore, the real profiles to which the expression of interest was sent were randomly assigned to the fictitious profiles. Hence, the coefficient β_1 on the independent variable $English_i$ may be interpreted as the impact of indicating English language skills on the probability of reacting to an expression of interest sent from the fictitious profile.

What about the effect of language on inter-caste marriages? To elucidate, in the current setting, if a LC woman with strong English language skills sends an expression of interest to a UC man this may generate greater interest for two reasons. First, due to the possibility of status exchange between class (language) and caste and second, if a potential groom has strong English language skills it may spark greater interest due to the perceived association between a shared language and shared values and norms.[21] On the other hand, if a UC woman signals strong English language skills (class) it reinforces her caste status, which may signal her unattainability and intimidate LC men and lead to less inter-caste interest in such a match. Since our focus is on caste hypergamy, the third hypothesis is that:

(*Hypothesis 3*) English language proficiency increases the probability of inter-caste interest for LC women.

We examine hypothesis 3 by estimating (1) for LC and for UC groups separately and subsequently for the four different caste combinations, that is responses to expressions of interest sent by LC fictitious (female) profiles to LC and UC real (male) profiles and UC fictitious (female) profiles sent to LC and UC real (female) profiles.

[21] A number of authors have explored the role of status exchange between caste and educational attainment in the Indian context (Dhar 2013; Ray et al. 2020) and between caste and income (Dugar, Bhattacharya and Reiley, 2012; Banerjee et al. 2013; Ahuja and Ostermann, 2016; Hortaçsu, Hwang and Mathur, 2019). The empirical work does not find much support for an exchange between caste status and educational background (Banerjee et al. 2013; Ray et al., 2020). There is much stronger support for status exchange between income and caste. Ahuja and Ostermann (2016) find that among high-caste low-income individuals there is greater interest in inter-caste marriages.

4. Data and descriptive statistics

In the data collection stage of the experiment, 384 male profiles were identified and received an interest from the 48 fictitious profiles. Table 3 presents descriptive statistics of the identified profiles and compares differences between LC and UC profiles.

As shown in the table, the mean age of the potential groom in the sample is 28 years. A majority of the sample (54%) has a bachelor's degree and about 29% have a postgraduate degree. Males in the income group Rs. 5-7.5 lakhs account for 31% of the sample. Since Hindi was set as a mother tongue in the search criteria for users, only 5% of the profiles have a mother tongue which is not Hindi.[22] These exceptions were native speakers of Haryanvi or Punjabi. About 30% of the sample identified themselves as English speakers. In 11% of the profiles, the self-introduction field was filled out in Hindi (no profiles were filled out in languages other than Hindi and English). In addition, about 22% of the profiles specified that someone other than the candidate in the profile administers the profile.

Of the 384 user profiles – 169 belonged to the LC group while the remainder were classified as UC. The sample shows inter-caste differences along some attributes, although the same criteria were used to identify the profiles. The UC group displays somewhat higher educational acquisition - 56.7% have a BA degree as compared to 51.5% amongst the LC group and 32.6% have an MA degree as compared to 23.7% amongst the LC group. There are income differences as well, with a larger proportion of UC users falling in the highest income bracket - 35.8% versus 25.4% for the LC group. While these differences match the narrative of UC having

[22] Even though Hindi mother tongue was specified in the search criteria, the algorithm identifies profiles indicating mother tongues/languages that are close to Hindi.

access to higher educational opportunities and higher income these differences should not have a bearing on our analysis.

As shown in Table 4, a consideration of the descriptive statistics conditional on the intervention at hand – that is, comparing means between the intervention (user profiles that received interest from a fictitious profile indicating English proficiency) and the control group (user profiles that received interest from a fictitious profile which does not indicate English proficiency) displays limited differences between the two groups. As is evident from the table there are few (a couple of income classes and profile written in Hindi) statistically significant differences in traits between the two groups. We control for differences in these traits in the regression analysis.

5. Results

5.1 Stage one

The first stage of the experiment recorded the number of views and the number of interests that each fictitious profile received within 72 hours of registration. On average, each of the fictitious profiles received 196 unique views and 55 interests or an interest to view conversion rate of 28.6% (Table 5). Consistent with hypothesis 1, the fictitious profiles that specified English language proficiency received substantially more views, interests, and interests per view compared to the profiles that did not specify such a trait (Table 6). On average, the profiles with English language proficiency attracted 20% (36) more viewers, received 38% (18) more interests and experienced a 16% higher interest to views conversion rate. These differences are substantial and statistically significant and comparable to the differences between the outcomes for upper caste and lower caste profiles – on average, UC profiles received 11% (21) more visitors albeit

not statistically significant, and 51% (23) more interests (statistically significant) as compared to LC profiles (see Table 7).

Table 8 provides regression-based estimates of the effect of English language proficiency and other traits (caste, education, income) on views and interests. These estimates confirm the effect of English language fluency and the effect of UC on views/interests. In addition, the regression estimates show that while higher education does translate into more views (large statistically significant estimates associated with possessing a BA or MA degree as opposed to high school) these views are not converted to interests. There is a clear education penalty of between 14% (MA degree) to 20% (BA degree), in terms of the interest to views conversion rate. Regarding income, while the effect of income on views is ambiguous, what seems to be clearer is that higher incomes are associated with fewer interests/interests per view. The income penalty at income levels above 2 lakhs is about 23-25%.

While the first stage of the experiment clearly shows the effect of indicating English language proficiency on the number of views and interests received by a fictitious profile, the analysis does not yield any information on the traits of the users responding to the fictitious profiles. Furthermore, the algorithms used by websites to recommend profiles for users to browse may influence the number of views and interest received by a fictitious profile. The basic algorithm is to show users the profiles which are most closely matched to the items they are looking for in a potential marriage candidate. For example, even if a user does not specify the desired caste of a marriage partner, the algorithm will automatically recommend profiles of the same or similar caste. The algorithms tend to favor the most recently logged-in users' profiles and place them higher in the search results and the website has a function that displays currently logged in users that match users' preferences for a marriage partner. Therefore, the number of

views and interests may be influenced by how many other users are logged in when the research team logged in to the fictitious profiles in the experiment.

To address these concerns, the research team logged in at the same time on the same day each week to register fictitious profiles. Nevertheless, it is impossible to know everything about these algorithms for the researcher. Therefore, we cannot entirely exclude the possibility that factors other than the profile attributes may have influenced the outcome variables in the first stage of the experiment. To obtain greater control over the experiment we turn to stage two.

5.2 Stage two

One of the aims of this study is to examine whether English proficiency translates into upward mobility in terms of caste in the marriage market and for this purpose it is important to have information on the traits of those responding to the fictitious profiles.

The second stage of the experiment addresses this need by identifying existing profiles, sending interests to randomly allocated profiles, and using their acceptance of this expression of interest as an outcome variable. This method augments the first stage in three ways. First, information about the decision maker is revealed.[23] Second, sending interest to randomly allocated profiles ensures that there is no association between the characteristics of the interest sender and the receiver, and this makes it possible to identify interest in inter-caste

[23] Typically, in correspondence studies, it is difficult to construct a dataset regarding the decision making parties as such studies often rely on information from job sites, company websites and the like (Duflo and Banerjee, 2017; Edo, Jacquemet and Yannelis, 2019). However, in this case, both sides reveal information.

relationships.[24] Third, accepting an interest is a much clearer disclosure of preferences for an alliance as opposed to visiting a profile and sending an interest.

Overall, 384 interests were sent to existing male profiles from 48 fictitious female profiles. Table 9 provides the acceptance rate of interest by English language proficiency (in columns) and other fictitious profile attributes (in rows). Refuting hypothesis 2, there is *no effect* of English language proficiency on acceptance of interest. This is not inconsistent with hypothesis 1 as in this active stage of the experiment interests are sent out by fictitious female profiles to a curated set of male profiles that fulfil certain conditions (see footnote 15) while in the first stage we are unaware of the traits of those who have shown interest in the profiles.

While there is no overall effect, fluency in English does seem to make a difference for some subgroups and there is a clear pattern emerging from Tables 9 and 10. The overall acceptance rate for lower caste profiles is 25% which is lower than the average response rate of 30.2%. However, the interest acceptance rate for LC female profiles indicating strong English language skills is 30.2% as opposed to 19.8% in the case of LC profiles without such an indication. In contrast to this compensatory English language effect for LC, acceptance rates are *lower* when English proficiency is indicated on UC profiles. While the overall acceptance rate (35.4%) is higher for UC profiles, it is 40.6% amongst those who do not indicate English while it is 30.2% amongst UC profiles that indicate such skills. The regression-based estimates (Table 10) also reveal the same picture, that is, there is no effect of English language proficiency for the

[24] For example, if interests were not randomized and sent according to the search results in each fictitious profile, the combinations of female and male profiles are likely to be of the same caste, similar educational and income level. In our experiment, on the other hand, there is no association between the characteristics of the interest sender and the receiver. Although this situation is unlikely in practice, it is useful for determining the interest in inter-caste marriages. The experiment by Ahuja and Ostermann (2016) also employs this strategy to analyze the determinants of interest in inter-caste marriage.

overall sample but there are positive, large (12.9 percentage points) and statistically significant effects for LC women. For UC women the effect is negative but statistically insignificant.

If for a moment, we turn to other traits we observe a similar compensatory pattern. Profiles with BA degrees and English language proficiency exhibit an interest acceptance rate of 37.5% while it is 31.3% amongst BA degree holders that do not indicate English language fluency. In contrast, for profiles with MA degrees and English language fluency the acceptance rate is 23.4% while it is 31.3% amongst those who have an MA degree and do not report English language fluency. Across income categories, the overall acceptance rate is lower than average for profiles with lower middle-income profiles. However, the acceptance rate is substantially higher when the profiles indicate English proficiency (27.1% with and 16.7% without English). In contrast, the acceptance rate is lower when English skills are mentioned for middle-income (33.3% with and 37.5% without) and upper middle-income profiles (31.3% with and 37.5% without). The main message emerging from Table 9 is that indicating English proficiency has a positive effect on the acceptance rate for relatively lower status (lower caste, lower education, lower income) female profiles but has a negative effect for higher status profiles.

The figures in Tables 9 and 10 show that the effects of English language proficiency are asymmetric. To probe this asymmetric pattern, Table 11 examines acceptance rates for LC and UC female profiles conditioning on the caste profile of males (recipients of the interest) and for English language fluency. While both LC and UC males show a greater degree of interest in LC females who are fluent in English, UC males display a substantially larger attraction to English language skills of LC females as opposed to LC males (13.3 percentage point difference and statistically significant versus 7.7 percentage point difference and statistically insignificant). In the case of UC females, both LC and UC males show a lower degree of interest in UC females

who report English fluency, but the negative effect is especially strong amongst LC males – while 21% of LC males are interested in UC females who are fluent, the interest acceptance rate is 38.1% amongst those who do not indicate English language fluency. Amongst, UC males the corresponding figures are 37.7% and 42.6% and the differences are not statistically significant. Regression based estimates (Table 12) paint a similar picture. While both LC and UC males are more likely to respond to LC women with strong English language skills the effect is positive, substantially larger (14.3 percentage points or 47%) and significant in the case of UC males. For LC males the effect is positive but smaller and insignificant. For UC women, indicating English reduces their attraction to both LC males and UC males but the effects are not statistically significant.

The pattern of results for LC women is consistent with the ideas of status exchange of class (language) for caste and the possibility that UC men may be attracted to LC women who are fluent in English as it signals the possibility of shared interests, values and norms. With regard to UC women, notwithstanding the lack of statistical significance, which may be due to power issues, the reduction in the lack of interest for fluent English speakers which emanates mainly from LC men suggests that the assertion of class (language) and caste may be intimidating.[25]

[25] The existing literature on the effects of English language proficiency tends to find large effects. Chakraborty and Bakshi (2016) find that abolishing English as a medium of instruction reduces wages by 26%. Although not the focus of their paper, Rosenzweig and Munshi (2006) find that English contributes to an increase in inter-caste marriages with 31.6 percent of English-educated siblings of the students in their sample marrying outside their *jati,* versus only 9.7 percent of Marathi-educated siblings. An effect of 21.9 percentage points. Azam, Chin, and Prakash (2013) estimate that speaking fluent English is associated with a 34% higher wage rate. Power analysis indicates that a sample of 180 observations is sufficient to detect an effect of 20% (lower end of the existing estimates) with a power of 80% at 5% level of significance.

6. Concluding remarks

Language is an important aspect of human capital, a primary means of accumulating human capital, and an inherent part of our identity. Knowledge of a dominant language, such as English, is not only associated with greater access to white-collar jobs and higher wages, but it also enables access to a wider knowledge base and enhances social interactions. If acquisition of such a language is restricted to the economically advantaged, then the importance such a language will be a source of widening economic and social gaps between those who are fluent in the language versus those are not. In India, English proficiency is a powerful signal that points to a person's educational background, cultural and economic background of the family, and social class.

Despite these effects, the non-economic returns to language proficiency have rarely been studied, especially returns to women's language skills. Motivated by the importance of English and the lack of evidence on its social returns, this study examined returns to English language proficiency in an online marriage market in Delhi, India. To overcome the problem of endogeneity and other socioeconomic variables and to signal language proficiency uniformly, the study used experimental methods and collected data through a major matrimonial website.

There are two key results. First, English proficiency provides women with a potentially wider choice of men. Fictitious female profiles indicating English proficiency attracted 20% more visitors and 38% more interests from real male profiles than profiles without English fluency. Second, English proficiency promotes caste hypergamy. For LC women indicating strong English language skills increased the response rate to their expression of interest from UC males by about 14 percentage points or 47%. In short, strong language skills are able to

compensate for LC status and potentially promote inter-caste alliances between LC women and UC men.

These results provide another rationale for investing in English language skills both at the individual and societal level and have clear policy implications. Currently the Indian government and various state governments promote social integration through inter-caste marriages by providing financial incentives.[26] Rather than relying on a one-time financial incentive which rewards inter-caste marriages, acquisition of English offers a transformative way of enhancing individual choice and achieving social integration.

While the results and implications are clear, it would be amiss to exaggerate the effects. The paper is based on a large (middle-class) although specific segment of the population and in the specific context of an urban metropolitan area. It is unclear whether these patterns would hold in other parts of the country where they may be greater resistance to inter-caste alliances. Nevertheless, at the very least, in the specific context of middle-class Delhi, there are substantial non-economic returns to English in the marriage market.

[26] The Ambedkar Foundation is an independent entity which operates under the Indian government's Ministry of Social Justice and Empowerment. The foundation runs the Dr. Ambedkar Scheme for Social Integration through Inter-Caste Marriages. The scheme provides a financial incentive of Rs. 2.5 lakhs for couples in an inter-caste marriage. As stated on the scheme's website (https://ambedkarfoundation.nic.in/icms.html), it is a reward to *"appreciate the socially bold step, of an inter-caste marriage, taken by the newly wedded couple*". Hortaçsu et al. (2018) examine the effect of such financial incentives and conclude that they have a modest effect in terms of raising the rate of exogamy between SC men and non-SC women in rural India.

References

Ahuja, A. and Ostermann, S.L. (2016) 'Crossing Caste Boundaries in the Modern Indian Marriage Market', *Studies in Comparative International Development*, 51(3), pp. 365–387. doi: 10.1007/s12116-015-9178-2.

Aldashev, A. and Danzer, A.M. (2014) *Economic Returns to Speaking the Right Language(s)? Evidence from Kazakhstan's Shift in State Language and Language of Instruction*. IZA Discussion Paper No. 8624, Available at SSRN: https://ssrn.com/abstract=2529333 or http://dx.doi.org/10.2139/ssrn.2529333

Allendorf, K. and Pandian, R.K. (2016) 'The Decline of Arranged Marriage? Marital Change and Continuity in India', *Population and Development Review*, 42(3), pp. 435–464. doi: 10.1111/j.1728-4457.2016.00149.x.

Angrist, J., Chin, A. and Godoy, R. (2006) 'Is Spanish-Only Schooling Responsible for the Puerto Rican Language Gap?', Working Paper 12005, National Bureau of Economic Research, 33(January).

Angrist, J.D. and Lavy, V. (1997) 'The effect of a change in language of instruction on the returns to schooling in Morocco', *Journal of Labor Economics*, 15(1), pp. 48–76. doi: 10.1086/209856.

Aoki, Y. and Santiago, L. (2015) Education, Health and Fertility of UK Immigrants: The Role of English Language Skills. IZA Discussion Paper No. 9498, Available at SSRN: https://ssrn.com/abstract=2690746 or http://dx.doi.org/10.2139/ssrn.2690746

Aoki, Y. and Santiago, L. (2018) Deprivation, Segregation, and Socioeconomic Class of UK Immigrants: Does English Proficiency Matter? IZA Discussion Paper No. 11368, Available at SSRN: https://ssrn.com/abstract=3137499 or http://dx.doi.org/10.2139/ssrn.3137499

Arceo-Gomez, B.E.O. and Campos-vazquez, R.M. (2014) 'Race and Marriage in the Labor Market : A Discrimination Correspondence Study in a Developing Country', *American Economic Review: Papers & Proceedings*, 104(5), pp. 376–380.

Asher, S., Novosad, P. and Rafkin, C. (2021) *Intergenerational Mobility in India: New Methods and Estimates Across Time, Space, and Communities*. Available online: https://paulnovosad.com/pdf/anr-india-mobility.pdf (Accessed: 24 May 2021).

Azam, M., Chin, A. and Prakash, N. (2012) 'The Returns to English-Language Skills in India', *Economic Development and Cultural Change*, 61(2), pp. 335–367. doi: 10.1086/668277.

Banerjee, A., Bertrand, M., Datta, S., & Mullainathan, S. (2009). Labor market discrimination in Delhi: Evidence from a field experiment. *Journal of Comparative Economics*, *37*(1), 14–27. https://doi.org/10.1016/j.jce.2008.09.002.

Banerjee, A., Duflo, E., Ghatak, M., & Lafortune, J. (2013). Marry for what? Caste and mate selection in modern India. *American Economic Journal: Microeconomics*, *5*(2), 33–72. https://doi.org/10.1257/mic.5.2.33.

Banerjee, B. and Knight, J.B. (1985) 'Caste Discrimination in the Indian Urban Labour Market', *Journal of Development Economics*, 17, pp. 277–307.

Bedi, A., Majilla, T. and Rieger, M. (2018) Gender Norms and the Motherhood Penalty: Experimental Evidence from India. IZA Discussion Paper No. 11360, Available at SSRN: https://ssrn.com/abstract=3137490 or http://dx.doi.org/10.2139/ssrn.3137490

Benson, C. (2016) Addressing language of instruction issues in education: recommendations for documenting progress; Background paper prepared for the 2016 Global education monitoring report, Education for people and planet: creating sustainable futures for all; 2016, ED/GEMR/MRT/2016/P1/6. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000245575 (Accessed: 27 August 2019)

Bertrand, M., Kamenica, E. and Pan, J. (2015) 'Gender Identity and Relative Income Within Households', *The Quarterly Journal of Economics*, 130(2), pp. 571–614. doi: 10.1093/qje/qjv001.Advance.

Bertrand, M. and Mullainathan, S. (2004) 'Are Emily and Greg More Employable than Lakisha and Jamal ? A Field Experiment on Labor Market Discrimination', *American Economic Review*, 94(4), pp. 991–1013.

Bharti, N.K, Chancel, L., Piketty, T. and Somanchi, A. (2024) Income and Wealth Inequality in India, 1922-2023: The Rise of the Billionaire Raj. Working Paper No. 9. World Inequality Lab. https://wid.world/document/income-and-wealth-inequality-in-india-1922-2023-the-rise-of-the-billionaire-raj-world-inequality-lab-working-paper-2024-09/.

Bhandari, P. (2020) *Matchmaking in middle class India: Beyond arranged and love marriage*. Singapore: Springer Nature, doi: 10.1007/978-981-15-1599-6.

Bidner, C. and Eswaran, M. (2015) 'A gender-based theory of the origin of the caste system of India', *Journal of Development Economics*, 114, pp. 142–158. doi: 10.1016/j.jdeveco.2014.12.006.

Bleakley, H. and Chin, A. (2004) 'Language Skills and Earnings: Evidence from Childhood Immigrants', *The Review of Economics and Statistics*, 86(May), pp. 481–496.

Bourdieu, P. (1977) 'The economics of linguistic exchanges', *Social Science Information*, 16(6), pp. 645–668.

Canagarajah, S. (2009) 'The plurilingual tradition and the English language in South Asia', *AILA Review*, 22(January 2009), pp. 5–22. doi: 10.1075/aila.22.02can.

Casale, D. and Posel, D. (2011) 'English language proficiency and earnings in a developing country: The case of South Africa', *Journal of Socio-Economics*, 40(4), pp. 385–393. doi: 10.1016/j.socec.2011.04.009.

Chaudhary, S. (2017) *Hindi Medium*. India: AA Films, Zee Studio.

Chakraborty, T. and S.K. Bakshi (2016) 'English Language Premium: Evidence from a Policy Experiment in India', *Economics of Education Review*, 50, pp. 1–16. doi: 10.1016/j.econedurev.2015.10.004.

Chiswick, B.R. (1998) 'Hebrew language usage: Determinants and effects on earnings among immigrants in Israel', *Journal of Population Economics*, 11(2), pp. 253–271. doi: 10.1007/s001480050068.

Chiswick, B.R. and Miller, P.W. (1995) 'The Endogeneity Between Language and Earnings: International Analyses', *Journal of Labor Economics*, 13(2), pp. 246–288.

Chiswick, B.R. and Miller, P.W. (2003) 'The Complementarity of Language and Other Human Capital: Immigrant Earnings in Canada', *Economics of Education Review*, 22(5), pp. 469–480. doi: 10.1016/S0272-7757(03)00037-2.

Chudgar, A. and Creed, B. (2016) 'How are Private School Enrolment Patterns Changing Across Indian Districts With a Growth in Private School Availability?', *Oxford Review of Education*, 42(5), pp. 543–560. doi: 10.1080/03054985.2016.1217688.

Dallin, A. (2000) 'Review Reviewed Work (s): Identity in Formation : The Russian-Speaking Populations in the Near Abroad', *The Journal of Modern History*, 72(2), pp. 583–585. Available at: https://www.jstor.org/stable/10.1086/316042.

Dhar, R.L. (2013) 'Intercaste Marriage : A Study from the Indian Context', *Marriage & Family Review*, 49(1), pp. 1–25. doi: 10.1080/01494929.2012.714720.

Duflo, E. and Banerjee, A.V. (2017) (Eds) *Handbook of Economic Field Experiments*. North Holland: Elsevier. Available at: https://www.elsevier.com/books/handbook-of-field-experiments/duflo/978-0-444-63324-8.

Dugar, S., Bhattacharya, H. and Reiley, D. (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(2), pp. 534–550. doi: 10.1111/j.1465-7295.2011.00398.x.

Education First (2023). EF English Proficiency Index. https://www.ef.com/wwen/epi/. Accessed on September 14, 2024.

Edgerton, N. (2016) 'Wage Returns to English Skills : A Comparison of Local and Multinational Banks in Thailand', *Southeast Asian Journal of Economics*, 4(June), pp. 113–140.

Edo, A., Jacquemet, N. and Yannelis, C. (2019) 'Language Skills and Homophilous Hiring Discrimination: Evidence From Gender and Racially Differentiated Applications', *Review of Economics of the Household*, 17(1), pp. 349–376. doi: 10.1007/s11150-017-9391-z.

Faust, D. and Nagar, R. (2001) 'Politics of Development in Postcolonial India: English Medium Education and Social Fracturing', *Economic & Political Weekly*, 36(30), pp. 2878–2883.

Fisman, R. *et al.* (2006) 'Gender Differences in Mate Selection: Evidence from a Speed Dating Experiment', *Quarterly Journal of Economics*, 121(2), pp. 673–697. doi: 10.1162/qjec.2006.121.2.673.

Fortin, N. M. (2005) 'Gender Role Attitudes and the Labour-market Outcomes of Women Across OECD Countries', *Oxford Review of Economic Policy*, 21(3), pp. 416–438. doi: 10.1093/oxrep/gri024.

Fu, X. (2007) 'Inter-racial Marriage and Family Socio-economic Status: A Study Among Whites, Filipinos, Japanese, and Hawaiians in Hawaii', *Journal of Comparative Family Studies*, 38(4). doi: 10.3138/jcfs.38.4.533.

Furtado, D. (2012) 'Human Capital and Interethnic Marriage Decisions', *Economic Inquiry*, 50(1), pp. 82–93. doi: 10.1111/j.1465-7295.2010.00345.x.

Ginsburgh, V. and Weber, S. (2020) 'The Economics of Language', *Journal of Economic Literature*, 58(2), pp. 348–404. doi: 10.1257/jel.20191316.

Goldin, C. (2006) 'The Quiet Revolution That Transformed Women's Employment, Education, and Family', *American Economic Review*, 96(2), pp. 1–21. doi: 10.1257/000282806777212350.

Harish, S. (2014) 'Social Strategy Use and Language Learning Contexts: A Case Study of Malayalee Undergraduate Students in India', *System*, 43(1), pp. 64–73. doi: 10.1016/j.system.2014.01.002.

Hnatkovska, V., Lahiri, A. and Paul, S. (2012) 'Castes and Labor Mobility', *American Economic Journal: Applied Economics*, 4(2), pp. 274–307.

Hortaçsu, A., Hwang, S.I.M. and Mathur, D. (2019) 'Monetary Incentives on Inter-caste Marriages in India: Theory and Evidence', *Journal of Development Economics*, 141(October 2018), p. 102371. doi: 10.1016/j.jdeveco.2019.102371.

Ignatiev, N. (1995) *How the Irish became White*. New York: Routledge.

Ito, T. (2009) 'Caste Discrimination and Transaction Costs in the Labor Market : Evidence from Rural North India', *Journal of Development Economics*, 88(2), pp. 292–300. doi: 10.1016/j.jdeveco.2008.06.002.

Jain, T. (2017) 'Common Tongue: The Impact of Language on Educational Outcomes', *Journal of Economic History*, 77(2), pp. 473–510. doi: 10.1017/S0022050717000481.

Jain, T., Maitra, P. and Mani, S. (2019) 'Barriers to Skill Acquisition : Evidence from English Training in India', *World Development*, 114, pp. 314–325. doi: 10.1016/j.worlddev.2018.10.011.

Jan, J. S. (2019) 'Ethnic Language Retention and East Asian Endogamy in the United States', *Marriage and Family Review*, 55(1), pp. 59–77. doi: 10.1080/01494929.2018.1458003.

Jayadeva, S. (2019) 'English-medium: Schooling, Social Mobility, and Inequality in Bangalore, India', *Anthropology & Education Quarterly*, 50, pp. 151–169. doi: 10.1111/aeq.12287.

Joshi, R. (2020) 'Can Social Integration in Schools be Mandated: Evidence from the Right to Education Act in India', *International Journal of Educational Development*, 77 (December 2019), p. 102228. doi: 10.1016/j.ijedudev.2020.102228.

Kaufmann, K.M., Messner, M. and Solis, A. (2013) 'Returns to Elite Higher Education in the Marriage Market : Evidence from Chile'. Available at SSRN: https://ssrn.com/abstract=2313369 or http://dx.doi.org/10.2139/ssrn.2313369

Kijima, Y. (2006) 'Caste and Tribe Inequality : Evidence from India , 1983 – 1999', *Economic Development and Cultural Change*, (2001), pp. 1983–1999.

Krishnan, S. and Hatekar, N. (2017) 'Rise of the New Middle Class in India and its Changing Structure', *Economic and Political Weekly*, 52(22), pp. 40–48.

LaDousa, C. (2007) 'Of Nation and State Language,S, and the Reproduction of Disparity in a North Indian City', *Anthropological Quarterly, 80(4), 925-959.*, 80(4), pp. 925–959.

Laitin, D.D. (1998) *Identity in formation: The Russian-speaking populations in the near abroad*. Ithaca: Cornell University Press.

Laitin, D D. and Ramachandran, R. (2016) 'Language Policy and Human Development', *American Political Science Review*, 110(3), pp. 457–480. doi: 10.1017/s0003055416000265.

Lang, K. (1986) 'A Language Theory of Discrimination', *The Quarterly Journal of Economics*, 101(2), pp. 363–382.

Lang, K., & Siniver, E. (2006). The Return to English in a Non-English Speaking Country: Russian Immigrants and Native Israelis in Israel. *NBER Working Paper*, (w12464).

Leech, T.G.J., Irby-Shasanmi, A. and Mitchell, A.L. (2019) '"Are You Accepting New Patients?" A Pilot Field Experiment on Telephone-based Gatekeeping and Black Patients' Access to Pediatric Care', *Health Services Research*, 54 (January), pp. 234–242. doi: 10.1111/1475-6773.13089.

Levinsohn, J. (2004) *Globalization and the Returns to Speaking English in South Africa*. NBER Working Paper Series. 10985. Cambridge, MA. Available at: https://www.nber.org/papers/w10985

Madheswaran, S. and Attewell, P. (2007) 'Caste Discrimination in the Indian Urban Labour Market : Evidence from the National Sample Survey', *Economic and Political Weekly*, 42(41), pp. 4146–4153.

Madheswaran, S. and Singhari, S. (2017) 'Social Exclusion and Caste Discrimination in Public and Private Sectors in India : A Decomposition Analysis', *The Indian Journal of Labour Economics*, 59(2), pp. 175–201. doi: 10.1007/s41027-017-0053-8.

Munshi, K. (2019) 'Caste and the Indian Economy', *Journal of Economic Literature*, 57(4), pp. 781–834. Available at: https://doi.org/10.1257/jel.20171307%0ACaste.

Munshi, K. and Rosenzweig, M. (2006) 'Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy', *American Economic Review*, 96(4), pp. 1225–1252. doi: 10.1257/aer.96.4.1225.

Munshi, K. and Rosenzweig, M. (2009) Why is Mobility in India So Low? Social Insurance, inequality, and Growth, NBER Working Paper Series. 14850. Cambridge, MA. Available at: http://www.nber.org/papers/w14850.

Munshi, K. and Rosenzweig, M. (2015) Insiders and Outsiders: Local Ethnic Politics and Public Goods Provision. 21720. Cambridge, MA. Available at: http://www.nber.org/papers/w21720.

Nair, S.K. (2015) *Does Medium of Instruction Affect Learning Outcomes?-Evidence Using Young Lives Longitudinal Data of Andhra Pradesh, India*. ESP Working paper Series No. 68, Available at: https://www.younglives.org.uk/publications/does-medium-instruction-affect-learning-outcomes (Accessed 18 April 2021).

Neyt, B., Vandenbulcke, S. and Baert, S. (2019) 'Are Men Intimidated by Highly Educated Women? Undercover on Tinder', *Economics of Education Review*, 73(June), p. 101914. doi: 10.1016/j.econedurev.2019.101914.

Office of the Registrar General & Census Commissioner, India (2022) 'Language Atlas of India, 2011. Available at: https://censusindia.gov.in/nada/index.php/catalog/42561 (Accessed 8 October 2024).

Ong, D. and Wang, J. (2015) 'Income Attraction: An Online Dating Field Experiment, *Journal of Economic Behavior and Organization*, 111, pp. 13–22. doi: 10.1016/j.jebo.2014.12.011.

Di Paolo, A. and Tansel, A. (2015) 'Returns to Foreign Language Skills in a Developing Country: The Case of Turkey', *Journal of Development Studies*, 51(4). doi: 10.1080/00220388.2015.1019482.

Parinduri, R. A. and Ong, K. (2018) *The effects of mediums of instruction on educational-and labor market outcomes: Evidence from Malaysia*. MPRA Paper No. 87560, Available at: https://mpra.ub.uni-muenchen.de/87560/ (Accessed: 15 May 2021).

Polzenhagen, F. and Frey, S. (2017) 'Are Marriages Made in Heaven? A Cultural-Linguistic Case Study on Indian-English Matrimonials', in Sharifian, F. (ed.) *Advances in Cultural Linguistics*. Springer, pp. 573–605.

Ramachandran, R. (2012) 'Medium of Instruction Policies and Human Capital Formation: Theory and Evidence from Africa', *SSRN Electronic Journal*. doi: 10.2139/ssrn.2001331.

Rao, G. (2019) 'Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools', *American Economic Review*, 109(3), pp. 774–809. doi: 10.1257/aer.20180044.

Ray, T., Roy, A. and Sahai, K. (2020) 'Whose Education Matters? An Analysis of Inter Caste Marriages in India', *Journal of Economic Behavior and Organization*, 176, pp. 619–633. doi: 10.1016/j.jebo.2020.02.011.

Reed, Megan N. 2019 "The Persistence of Arranged Marriage in India: New Evidence from the Delhi National Capital Region." Population Association of America (PAA) Annual Meeting. Austin, TX. April. (https://paa2019.populationassociation.org/uploads/193533).

Rosenzweig, M.R. and Stark, O. (1989) 'Consumption Smoothing, Migration, and Marriage: Evidence from Rural India', *Journal of Political Economy*, 97(4), pp. 905–926. Available at: https://www.jstor.org/stable/1832196.

Sandhu, P. (2014) 'The Interactional and Narrative Construction of Normative and Resistant Discourses about Hindi and English', *Applied Linguistics*, 35(1), pp. 29–47. doi: 10.1093/applin/ams075.

Shukla, R. (2010) *How India Earns, Spends and Saves: Unmasking the real India*. SAGE Publications India, New Delhi.

Siddique, Z. (2011) 'Evidence on Caste Based Discrimination', *Labour Economics*, 18, pp. S146–S159. doi: 10.1016/j.labeco.2011.07.002.

Srinivas, M.N. (1957) 'Caste in Modern India', *The Journal of Asian Studies*, 16(4), pp. 529–548.

Titzmann, F. (2018) 'Hindu Religious Identification in India's Online Matrimonial Market', in Zeiler, X. (ed.) *Digital Hinduism*. 1st Editio. Routledge, pp. 35–50. Available at: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315107523-3/hindu-religious-identification-india-online-matrimonial-market-fritzi-marie-titzmann.

Trivedi, I. (2014). India in Love: Marriage and Sexuality in the 21st Century. India: Aleph Book Company.

Global Education Monitoring Report Team. (2016). *If you don't understand, how can you learn?* (24; Global Education Monitoring Report: Policy Paper, Issue February). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000243713.

Yuki, K. (2021) 'Language education and economic outcomes', Available at SSRN: https://ssrn.com/abstract=3846985 or http://dx.doi.org/10.2139/ssrn.3846985.

Tables

Table 1
Categories of key profile attributes

Attributes	Assigned categories	What appears in the profile
English proficiency	Yes	Outstanding English language skills, both written and spoken
	No	-
Caste status	UC	Brahmin
	LC	Scheduled Castes (SCs)
Annual Income	No income	Rs. 0
	Lower-middle income	Rs. 1-2 lakh
	Middle income	Rs. 3-4 lakh
	Upper-middle income	Rs. 5-7.5 lakh
Education	Less than bachelor's degree	High school
	Bachelor's degree	BA
	Postgraduate degree	MA

Table 2
Summary statistics of fictitious profiles

	Without English	With English
Caste	0.5	0.5
Income		
No income	0.25	0.25
Lower-middle income	0.25	0.25
Middle income	0.25	0.25
Upper-middle income	0.25	0.25
Education		
High school	0.333	0.333
BA	0.333	0.333
MA	0.333	0.333
Observations	24	24

Table 3
Summary statistics of identified profiles

	Overall		LC	UC	
	Mean	SD	Mean	Mean	p-value
Age	28.758	2.469	28.811	28.716	0.711
Education					
Lower than bachelor's degree	0.169	0.375	0.249	0.107	0.000***
Bachelor's degree	0.544	0.499	0.515	0.567	0.305
Master's degree and above	0.286	0.453	0.237	0.326	0.056*
Income					
Rs. 0-1 lakh	0.047	0.212	0.053	0.042	0.601
Rs. 1-2 lakh	0.089	0.284	0.142	0.047	0.001***
Rs. 2-3 lakh	0.154	0.361	0.160	0.149	0.769
Rs. 3-4 lakh	0.177	0.382	0.201	0.158	0.274
Rs. 4-5 lakh	0.169	0.375	0.112	0.214	0.008***
Rs. 5-7.5 lakh	0.313	0.464	0.254	0.358	0.030**
Languages					
Mother tongue other than Hindi	0.052	0.222	0.053	0.051	0.927
English	0.297	0.457	0.284	0.307	0.626
Profile written in Hindi	0.112	0.316	0.118	0.107	0.727
Government job	0.133	0.340	0.142	0.126	0.639
Profile managed by others	0.223	0.416	0.165	0.268	0.017**
Accept interest	0.302	0.460	0.302	0.302	0.991
Observations	384		169	215	

Notes: P-values from a means comparison test (t-test) between LC and UC profiles.

Table 4
Randomization check

	Without English (N=192)	**With English** (N=192)	**p-value**
Age	28.703	28.813	0.665
	(0.179)	(0.177)	
Education			
Lower than bachelor's degree	0.135	0.203	0.077
	(0.025)	(0.029)	
Bachelor's degree	0.552	0.536	0.759
	(0.036)	(0.036)	
Master's degree or above	0.313	0.260	0.260
	(0.034)	(0.032)	
Income			
No income	0.052	0.052	1.000
	(0.016)	(0.016)	
Rs. 0-1 Lakh	0.031	0.063	0.148
	(0.013)	(0.018)	
Rs. 1-2 Lakh	0.063	0.115	0.073*
	(0.018)	(0.023)	
Rs. 2-3 Lakh	0.125	0.182	0.120
	(0.024)	(0.028)	
Rs. 3-4 Lakh	0.188	0.167	0.594
	(0.028)	(0.027)	
Rs. 4-5 Lakh	0.172	0.167	0.892
	(0.027)	(0.027)	
Rs. 5-7.5 Lakh	0.370	0.255	0.015**
	(0.035)	(0.032)	
Languages			
Mother tongue not Hindi	0.036	0.068	0.169
	(0.014)	(0.018)	
English	0.333	0.260	0.118
	(0.034)	(0.032)	
Profile written in Hindi	0.156	0.068	0.006***
	(0.026)	(0.018)	
Government Sector Jobs	0.130	0.135	0.881
	(0.024)	0.025	
Profile managed by others	0.240	0.204	0.416
	(0.031)	(0.03)	

Notes: "Without English" is the group that received interest from a fictitious female profile that did not specify English proficiency, and "With English" is the group that received interest from a fictitious female profile that did specify English ability. P-values from a means comparison test (t-test).

Table 5
Stage One - Overall

	Obs.	Mean	Std. dev.	Min	Max
Views	48	196.104	55.932	51	309
Interests received	48	55.396	21.785	21	115
Interests per view	48	0.286	0.080	0.181	0.608

Table 6
Stage One - by English proficiency
(Std. Dev.)

	Without English	With English	p-value
Views	178.125	214.083	0.024**
	(41.614)	(63.138)	
Interests received	46.458	64.333	0.003***
	(12.693)	(25.337)	
Interests per view	0.265	0.307	0.065*
	(0.056)	(0.094)	
Observation	24	24	

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively. P-values from a means comparison test (t-test).

Table 7
Stage One – by Caste Status
(Std. Dev.)

	LC	UC	p-value
Views	185.625	206.583	0.197
	(51.336)	(59.398)	
Interests received	44.083	66.708	0.000***
	(13.906)	(22.568)	
Interests per view	0.249	0.322	0.001***
	(0.087)	(0.050)	
Observations	24	24	

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively.

Table 8
Stage One – The effect of various traits on views and interests
(absolute value of t-statistics)

Variables	Views	Interests	Interests per view
English language proficiency	35.95***	17.87***	0.042**
	(3.25)	(4.95)	(2.43)
Caste (UC = 1)	20.95*	22.62***	0.073***
	(1.89)	(6.26)	(4.21)
BA degree	38.5***	1.56	-0.056**
	(2.84)	(0.35)	(2.66)
MA degree	57.62***	6.37	-0.040*
	(13.56)	(1.44)	(1.88)
Income 1-2 lakhs	26.66*	4.5	-0.035
	(1.70)	(0.88)	(1.45)
Income 3-4 lakhs	-48.75*	-24.16***	-0.070***
	(3.11)	(4.73)	(2.84)
Income 5-7.5 lakhs	15.5	-6.41	-0.067***
	(0.99)	(1.26)	(2.72)
Constant	137.25***	39.02***	0.303***
	(8.76)	(7.63)	(12.29)
Observations	48	48	48
R-squared	0.599	0.718	0.510

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively.

Table 9
Stage Two - Acceptance rate of interest by fictitious profile attributes

		English proficiency		Overall	p-value
		With English	Without English		
Caste	LC	30.2%	19.8%	25%	0.019**
	UC	30.2%	40.6%	35.4%	0.033**
Education	HS	29.7%	28.1%	28.9%	0.730
	BA	37.5%	31.3%	34.4%	0.201
	MA	23.4%	31.3%	27.3%	0.082**
Income	No income	29.2%	29.2%	29.2%	0.983
	Lower middle	27.1%	16.7%	21.9%	0.014**
	Middle	33.3%	37.5%	35.4%	0.389
	Upper middle	31.3%	37.5%	34.4%	0.201
All		30.2%	30.2%	30.2%	0.983
Obs.		192	192	384	

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively. P-values are from statistical tests for differences in response proportion between those with and without English language proficiency.

Table 10
Stage Two – Probability of accepting interest
(absolute value of t-statistics)

Variables	Overall	LC	UC
Female profile			
English Language Proficiency	0.020	0.129**	-0.085
	(0.43)	(2.16)	(1.21)
Caste (UC = 1)	0.123**	.	.
	(2.57)		
BA degree	0.063	-0.013	0.165***
	(1.03)	(0.17)	(1.84)
MA degree	-0.014	-0.049	0.024
	(0.31)	(0.81)	(0.39)
Female income 1-2 Lakhs	-0.100	-0.041	-0.173
	(1.33)	(0.40)	(1.59)
Female income 3-4 Lakhs	0.037	0.090	-0.018
	(0.52)	(1.02)	(0.21)
Female income 5-7.5 Lakhs	0.041	0.026	0.058
	(0.55)	(0.28)	(0.57)
Male profile			
Caste (UC = 1)	-0.008	-0.131**	0.137***
	(0.16)	(2.22)	(1.61)
BA degree	0.084	0.151	-0.013
	(1.07)	(1.32)	(0.12)
MA degree	0.053	0.099	-0.017
	(0.56)	(0.75)	(0.13)
Male income 2-3 lakhs	0.067	-0.12	0.085
	(0.80)	(0.13)	(0.54)
Male income 3-4 lakhs	0.309*	0.196**	0.402**
	(4.02)	(1.97)	(3.29)
Male income 4-5 lakhs	0.136	0.118	0.157
	(1.43)	(0.89)	(1.13)
Male income 5-7.5 lakhs	0.077	0.050	0.090
	(1.05)	(0.55)	(0.84)
Observations	384	192	192
Log-likelihood	-220	-98	-112

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively.

Table 11
Stage Two - Acceptance rate by caste English language proficiency

		English proficiency		Overall	p-value
Female Caste	Male Caste	With English	Without English	*(N)*	
LC	LC	35%	27.3%	31.0% *(84)*	0.446
	UC	26.8%	13.5%	20.4% *(108)*	0.087*
UC	LC	20.9%	38.1%	29.4% *(85)*	0.082*
	UC	37.7%	42.6%	40.2% *(107)*	0.426

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively. P-values are from statistical tests for differences in response proportion between those with and without English language proficiency.

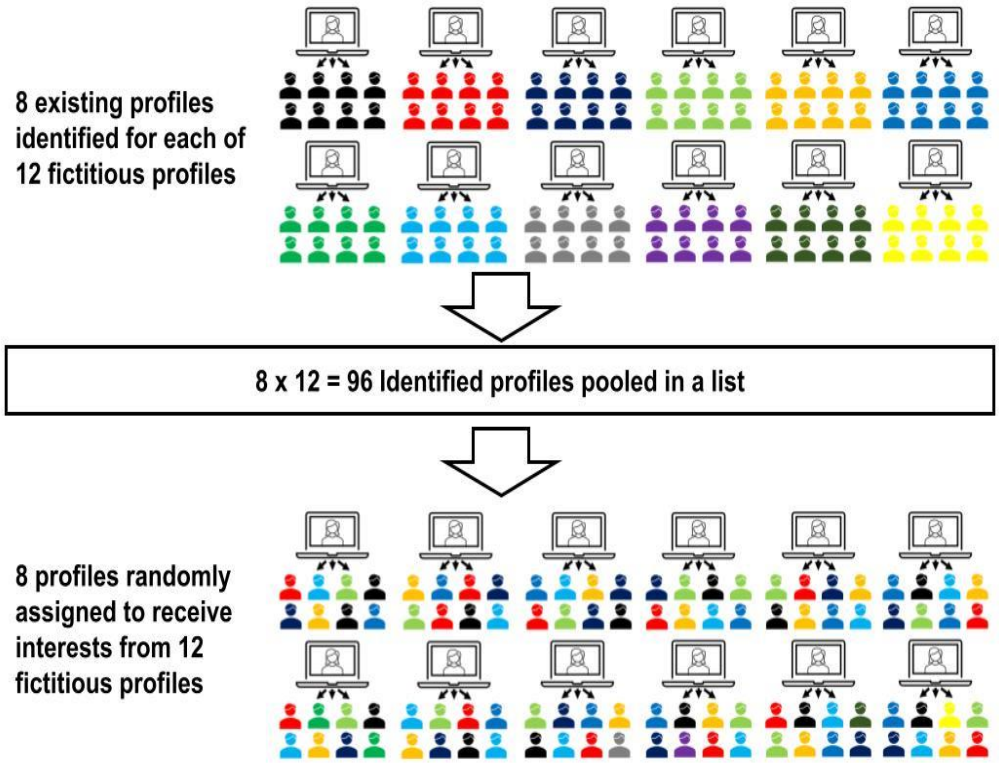
Table 12
Stage Two – Caste specific probability of accepting interest
(absolute value of t-statistics)

Variables	Female LC Male LC	Female LC Male UC	Female UC Male LC	Female UC Male UC
Female profile				
English Language Proficiency	0.088	0.143**	-0.133	-0.060
	(1.04)	(2.11)	(1.33)	(0.70)
BA degree	0.011	-0.052	0.277**	0.139
	(0.10)	(0.54)	(2.07)	(1.45)
MA degree	-0.058	-0.071	-0.022	0.102
	(0.70)	(0.75)	(0.16)	(1.36)
Female income 1-2 Lakhs	0.037	-0.059	-0.251**	-0.020
	(0.40)	(0.56)	(2.20)	(0.17)
Female income 3-4 Lakhs	0.210	-0.001	-0.212*	0.195**
	(1.50)	(0.01)	(1.71)	(1.96)
Female income 5-7.5 Lakhs	0.230**	-0.004	-0.107	0.171*
	(2.30)	(0.04)	(0.77)	(1.79)
Male profile				
BA degree	0.164	0.184*	0.017	-0.018
	(0.88)	(1.84)	(0.12)	(0.08)
MA degree	0.194	0.115	-0.006	-0.002
	(0.88)	(0.91)	(0.05)	(0.01)
Male income 2-3 lakhs	0.437***	-0.204**	0.289	-0.154
	(3.32)	(2.18)	(1.17)	(0.84)
Male income 3-4 lakhs	0.652***	-0.117	0.649***	0.022
	(4.03)	(1.15)	(2.61)	(0.16)
Male income 4-5 lakhs	0.418	-0.094	0.263	-0.097
	(1.64)	(0.83)	(1.18)	(0.55)
Male income 5-7.5 lakhs	0.391	-0.163	0.309*	-0.211
	(2.23)	(1.34)	(1.64)	(1.33)
Observations	84	108	85	107
Log-likelihood	-42	-49	-37	-67

Notes: *, **, *** Statistical significance at 10, 5, 1 percent, respectively.

Figures

Figure 1
Process of identifying profiles, random assignment and sending interests



Appendices

Appendix 1 - Details of matrimonial website

a. Registration Process

The registration process includes the process of identification/authentication and entering personal information. Most of the personal information required at this stage is mandatory for all profiles. The main items with a choice element are summarized in Table A1.

① Sign Up

New registrants to the website must first enter and select the following information.

- Email address
- Mobile phone number
- Password
- Create profile for: (Choose from) Self, Son, Daughter, Brother, Sister, Other. When other is selected, further specify from relative/Friend or client (for marriage bureaus).
- Gender (Male or Female)

② Profile Details

In the next page, users are required to enter key characteristics of the candidate.

- Groom/Bride's Name
- Whether or not to show the name to other accounts
- Date of Birth
- Mother tongue
- Religion
- Caste
- Whether or not to be open to marry for all castes ("Caste no bar")
- Manglik status
- Whether if you prefer horoscope match
- Marital Status
- Height

③ Career details

- Country of residence
- State of residence
- City of residence
- Pin code of the address
- Highest degree
- Sector of employment
- Occupation
- Annual income
- Self-introduction - free-text of at least 50 words.

④ Authentication process

The registrant enters a one-time four-digit password sent by SMS to the mobile phone number they have entered to confirm their identity.

⑤ Lifestyle & Family (this section may be skipped)

- Family type (Joint family, nuclear family, others)
- Father's occupation
- Mother's occupation

- Number of brothers and how many of them are married
- Number of sisters and how many of them are married
- Place of residence of the family
- Native City
- Contact address
- Free description of the family

b. Further information in the profiles

Once a user enters the information above, the profile is complete and awaits verification by an administrator. Verification usually takes place within a few hours. Most of the information, except for date of birth, caste, and gender may be edited afterwards. Additional details such as the following may be added to the profile after registration.

⑥ Education & Career
- Name of school, undergraduate and post graduate college
- Name of the employer
- Interest in settling abroad (Yes, No, Undecided)

⑦ Family details
- Gothra
- Family value (Orthodox, Conservative, Moderate, Liberal)

⑧ Habits
- Dietary Habits (vegetarian, non-vegetarian, Jain, Eggetarian)
- Drinking Habits (Yes, No, Occasionally)
- Smoking Habits (Yes, No, Occasionally)
- Open to Pets? (Yes, No)
- Own a House? (Yes, No)
- Own a Car? (Yes, No)
- Languages I speak (Indian languages + Some European languages such as English, French, German, Portuguese, as well as Asian languages including Arabic, Pushto, Persian, Japanese, and Mandarin)
- Blood Group
- HIV+ (Yes, No)
- Thalassemia (Major, Minor, No)
- Challenged (Physically-from birth, Physically Due to accident, Mentally-from birth, Mentally - due to accident, further choose nature of handicap for physical handicaps)

⑨ Your likes
- Hobbies
- Interests
- Favorite Music
- Favorite Book
- Dress Style
- TV Shows
- Favorite Movies
- Movies
- Sports
- Cuisine
- Food I cook
- Vacation Destination

Table A1
Items with choices

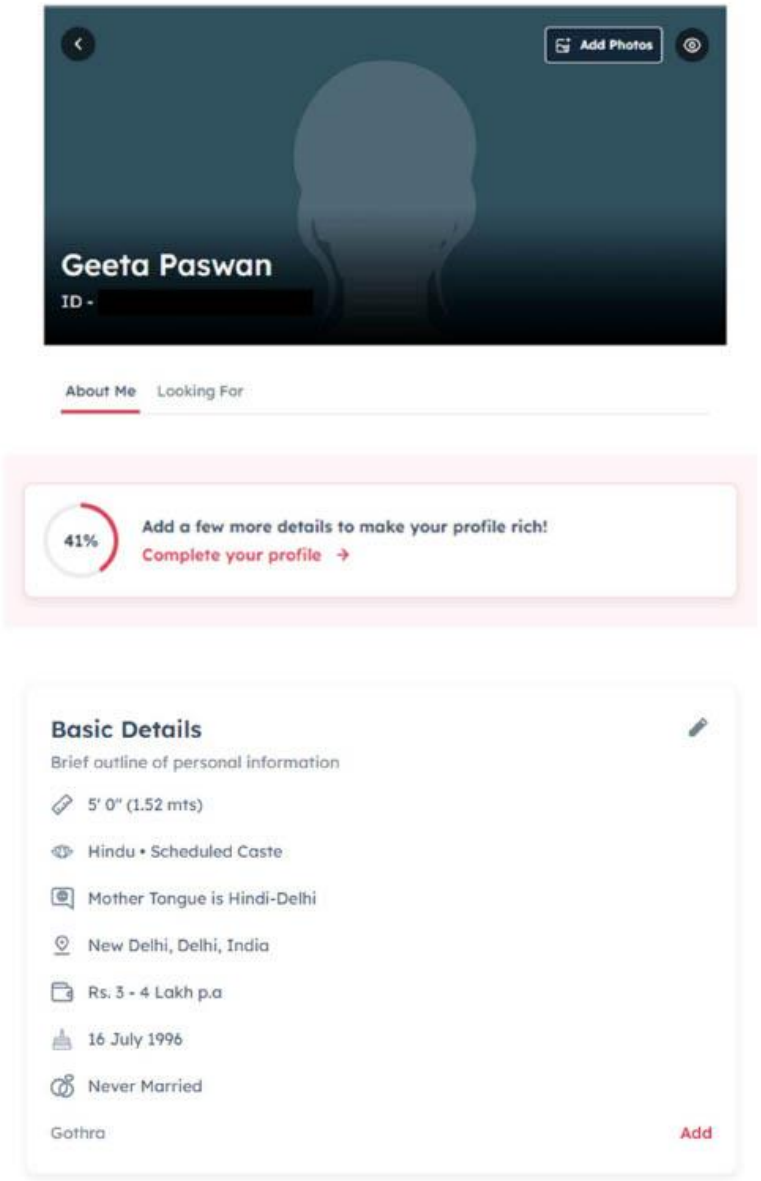
Category	Available choices	
Age	A number between 18 to 70 years old.	
Height	Choice of 4 feet (122 cm) to 7 feet (213 cm) or more in inches	
Religion and Caste	Hindu, Muslim, Sikh, Christian, Buddhist, Jain, Parsi, Jewish, Bahai, Other	Choose from more than 500 subcastes (Group of subcastes: Aggarwal, Arya Vysya, Bania, Banik, Brahmin, Brahmin Andhra, Brahmin Bengali, Brahmin Dravida, Brahmin Gaur, Brahmin Gujarati, Brahmin Kannada, Brahmin Kanyakubj, Brahmin Karnataka, Brahmin Maharastra, Brahmin Rajasthani, Brahmin Saraswat, Brahmin Utkal, Chettiar, Kayastha, Khatri, Kapu, Kshatriya, Maratham Marwari, Mudaliar, Nair, Patel, Patil, Rajput, Scheduled Caste, Scheduled Tribe, Sindhi, Teli, Vaishnav, Vellalar)
	Muslim	Select "sect" from Shia or Sunni, Select Caste from 47 categories
	Sikh	Select one from 19 caste categories
	Christian	Select one from 30 "sect" categories
	Buddhist	(No choice of subgroup available)
	Jain	Select from 3 caste categories
	Parsi	(No choice of subgroup available)
	Jewish	(No choice of subgroup available)
	Bahai	(No choice of subgroup available)
	Other	(No choice of subgroup available)
Mother Tongue	Hindi, Hindi-Delhi, Hindi-MP/CG, Hindi-UP/UK, Punjabi, Hindi-Bihar/Jharkhand, Hindi-Rajasthan, Haryanvi, Himachali, Kashmiri, Sindhi, Urdu, Marathi, Gujarati, Kutchi, Konkani, Tamil, Telugu, Kannada, Malayalam, Tulu, Bengali, Oriya, Assamese, Sikkim/Nepali, English	
Country	Select one from all countries	
City / State	(When India is selected for country)	
Annual Income	No Income, Rs. 0-1 Lakh, Rs. 1-2 Lakh, Rs. 2-3 Lakh, Rs. 3-4 Lakh, Rs. 4-5 Lakh, Rs. 5-7.5 Lakh, Rs. 7.5-10 Lakh, Rs. 10-15 Lakh, Rs. 15-20 Lakh, Rs. 25-35 Lakh, Rs. 35-50 Lakh, Rs. 50-70 Lakh, Rs. 70 Lakh-1 Crore, Above Rs. 1 Crore	
Marital status	Never Married, Awaiting Divorce, Divorced, Widowed, Annulled	
Have children	"No", "Yes, living together", "Yes, living separately"	
Astro	Manglik Status	Non Manglik, Manglik, Angshik (partial manglik)
	Horoscope	Date of Birth, Time of births

Appendix 2 - Example of fictitious profiles

All fictitious profiles contained the same information in all categories except English proficiency, caste, educational background, and income.

In all the profiles, gender was set to female, profile manager was the individual, mother tongue was Hindi-Delhi, religion was Hindu, sub-caste was left blank, marriage status was set to Never Married, height was set to 5 feet, which is about the average height for Indian women above 19, and age was set to 25 years. Her place of residence and origin was set to Delhi, working in the private sector, particularly in banking. Her family is nuclear and lives in Delhi, her father works in the private sector, her mother is a housewife, and she has one married brother. Her preferred marriage partner is a Hindu, but she does not specify her caste (caste no bar). Her lifestyle includes occasional drinking, and she does not smoke. Her date of birth and zip code were randomly assigned from a list. For the self-introduction, several variations of the standard content were created, and one of these was randomly selected to be included in each profile. An example of a fictitious profile registered on a website is included in Appendix 3.

Appendix 3 – Marriage profiles with and without indication of strong English proficiency



About Me

Describe yourself in a few words

I'm a banking professional who truly loves my job and enjoys tackling financial challenges. I have a passion for continuous learning and exploring new ideas, which keeps my work both engaging and fulfilling. I consider myself an easy-going, flexible, and energetic person with a love for cricket!

In a partner, I value mutual respect for each other's individuality and the importance of supporting one another. Building a happy life together, while nurturing our families, is a priority for me.

Outstanding English language skills, both written and spoken.

Profile managed by Self

Disability, Thalassemia, HIV+ Add

Education

Showcase your educational qualification

B.A - Undergraduate Degree
UG College

School Name, UG College Add

Write about your education...

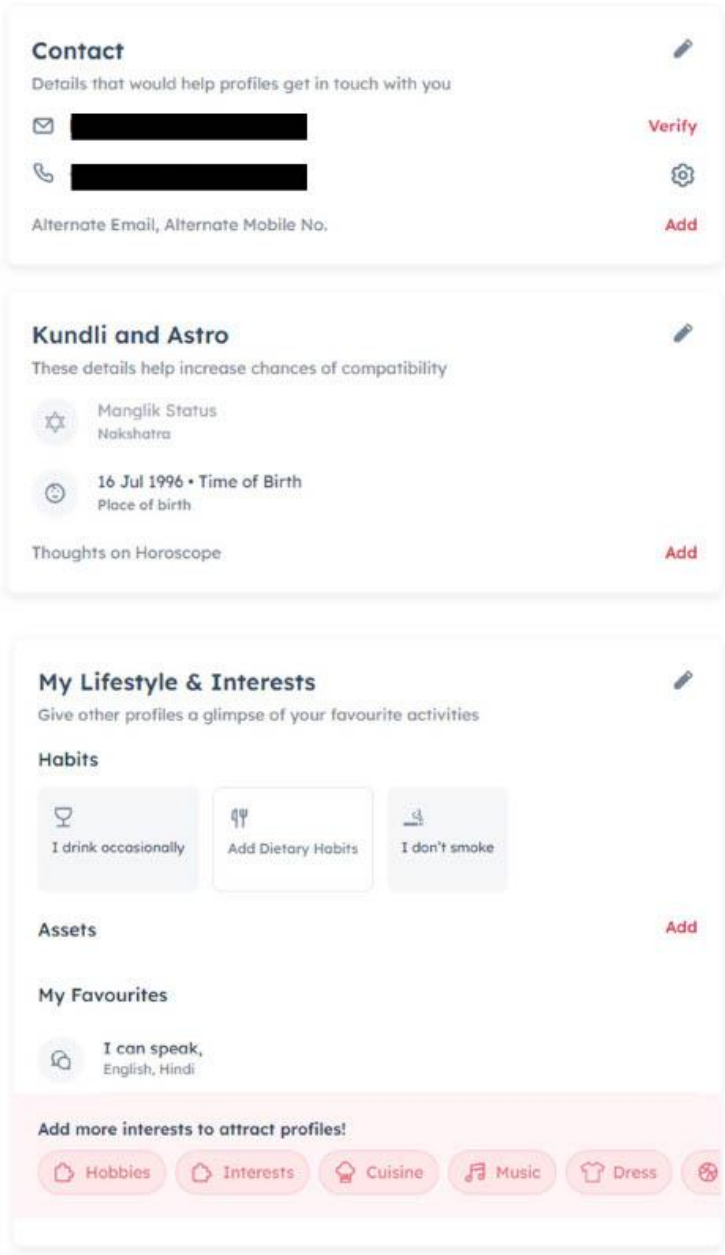
Career

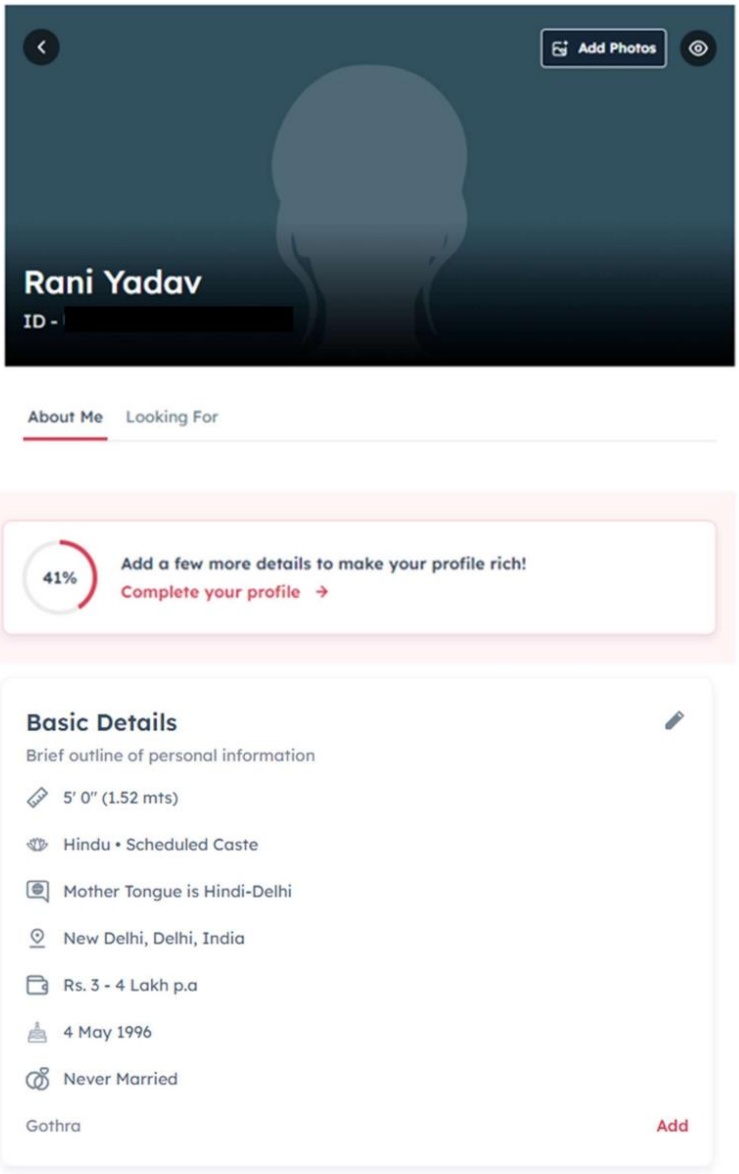
Give a glimpse of your professional life

Banking Professional
Private Sector

Organisation Name, Thoughts on settling abroad Add

Write about your career...


Contact
Details that would help profiles get in touch with you
[redacted]
Verify
[redacted]
Alternate Email, Alternate Mobile No.
Add
Kundli and Astro
These details help increase chances of compatibility
Manglik Status
Nakshatra
16 Jul 1996 • Time of Birth
Place of birth
Thoughts on Horoscope
Add
My Lifestyle & Interests
Give other profiles a glimpse of your favourite activities
Habits
I drink occasionally
Add Dietary Habits
I don't smoke
Assets
Add
My Favourites
I can speak,
English, Hindi
Add more interests to attract profiles!
Hobbies
Interests
Cuisine
Music
Dress


Add Photos
Rani Yadav
ID -
About Me
Looking For
41%
Add a few more details to make your profile rich!
Complete your profile →
Basic Details
Brief outline of personal information
5' 0" (1.52 mts)
Hindu • Scheduled Caste
Mother Tongue is Hindi-Delhi
New Delhi, Delhi, India
Rs. 3 - 4 Lakh p.a
4 May 1996
Never Married
Gothra
Add

About Me

Describe yourself in a few words

I'm a dedicated banking professional who finds great joy in my work and thrives on addressing financial challenges. My passion for continuous learning and exploring fresh ideas makes my career both exciting and rewarding. I'd describe myself as easy-going, adaptable, and full of energy, with a strong enthusiasm for cricket!

When it comes to a partner, I deeply value mutual respect for our individuality and the significance of supporting each other. Creating a joyful life together while caring for our families is very important to me.

Profile managed by Self

Disability, Thalassemia, HIV+ Add

Education

Showcase your educational qualification

B.A - Undergraduate Degree
UG College

School Name, UG College Add

Write about your education...

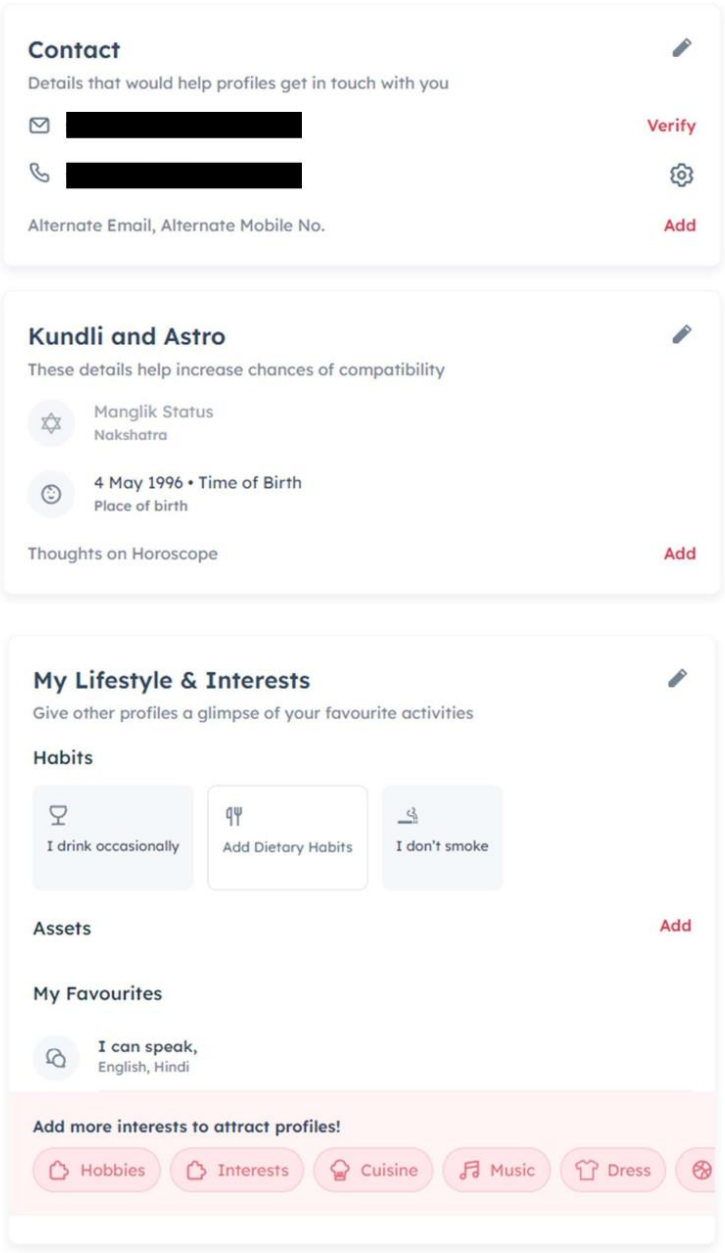
Career

Give a glimpse of your professional life

Banking Professional
Private Sector

Organisation Name, Thoughts on settling abroad Add

Write about your career...


Contact
Details that would help profiles get in touch with you
[redacted]
Verify
[redacted]
Alternate Email, Alternate Mobile No.
Add
Kundli and Astro
These details help increase chances of compatibility
Manglik Status
Nakshatra
4 May 1996 • Time of Birth
Place of birth
Thoughts on Horoscope
Add
My Lifestyle & Interests
Give other profiles a glimpse of your favourite activities
Habits
I drink occasionally
Add Dietary Habits
I don't smoke
Assets
Add
My Favourites
I can speak,
English, Hindi
Add more interests to attract profiles!
Hobbies
Interests
Cuisine
Music
Dress