

# Smart, Responsible, and Upper Caste Only: Measuring Caste Attitudes through Large-Scale Analysis of Matrimonial Profiles

Ashwin Rajadesingan,<sup>1</sup> Ramaswami Mahalingam,<sup>2</sup> David Jurgens<sup>1</sup>

<sup>1</sup>School of Information, <sup>2</sup>Department of Psychology  
{arajades,ramawasi,jurgens}@umich.edu  
University of Michigan, Ann Arbor

## Abstract

Discriminatory caste attitudes currently stigmatize millions of Indians, subjecting individuals to prejudice in all aspects of life. Governmental incentives and societal movements have attempted to counter these attitudes, yet accurate measurements of public opinions on caste are not yet available for understanding whether progress is being made. Here, we introduce a novel approach to measure public attitudes of caste through an indicator variable: openness to intercaste marriage. Using a massive dataset of over 313K profiles from a major Indian matrimonial site, we precisely quantify public attitudes, along with differences between generations and between Indian residents and diaspora. We show that younger generations are more open to intercaste marriage, yet attitudes are based on a complex function of social status beyond their own caste. In examining the desired qualities in a spouse, we find that individuals open to intercaste marriage are more individualistic in the qualities they desire, rather than favoring family-related qualities, which mirrors larger societal trends away from collectivism. Finally, we show that attitudes in diaspora are significantly less open, suggesting a bi-cultural model of integration. Our research provides the first empirical evidence identifying how various intersections of identity shape attitudes toward intercaste marriage in India and among the Indian diaspora in the US.

## 1 Introduction

Traditionally, marriages in India are considered as a union of two families. Typically, parents initiate and mediate the search for a spouse within kinship networks and ensure compatibility based on factors such as caste, education, affluence, horoscope and physical characteristics. These “arranged” marriages are usually endogamous with the bride and groom belonging to the same caste, reinforcing caste lines (Ambedkar 2004). However, recent studies show a decline in parental control in the matchmaking process, with more individuals reporting that they chose their spouse independently or at least jointly with their parents (Allendorf 2013; Allendorf and Pandian 2016). Coupled with a greater openness to marry outside their caste, this increase in agency

has the potential to alter the social fabric of the society, especially pertaining to caste relations. In this work, using data from over 313K profiles from a major Indian matrimonial site, we study social inclusion by examining partner preferences and factors that affect openness to intercaste marriage in India and the Indian diaspora in the US.

Previous research on intercaste marriage primarily relies on rich ethnographies (Fuller and Narasimhan 2008; Kannan 1963), qualitative interviews (Dhar 2013), surveys (Goli, Singh, and Sekher 2013; Ray, Chaudhuri, and Sahai 2017) and audit studies (Banerjee et al. 2013; Ahuja and Ostermann 2016). Unlike surveys which depend on self-reported openness which may have a social desirability bias (Krumpal 2013) or measuring incidence of intercaste marriages which may be different from actual openness, in our work, we use data from online matrimonial profiles where users specify their openness to intercaste marriage when they register. As profiles contain rich demographic information, we can identify how various intersections of identity shape attitudes toward inter-caste marriage.

In Indian matrimonial sites, either individuals or their family can create a matrimonial profile, allowing us to study generational differences in spousal preferences, and by proxy, caste attitudes. Further, given these websites’ popularity among Indian diaspora, through openness to intercaste marriage, we analyze how intercaste relations differ among diaspora in the US who live as a minority in a fundamentally different cultural context. These online profiles provide a unique real-time estimate of changing societal attitudes on caste and, because of their intensely personal nature, accurately reflect social preferences.

Our work provides the following three main contributions. First, we find younger generations are substantially more open to intercaste marriage when controlling for demographics and background (§4). However, the demographic factors associated with openness are highly similar across generations, with the exception that higher education in children shows increased openness when compared with their family. Moreover, we find clear geographic trends in acceptance to intercaste marriage, with South Indian states being less open. Second, we show that individuals who are open to intercaste marriage are more likely to desire a spouse

based on individual traits, whereas those not open to intercaste marriage emphasize family; these trends mirror the shift in Indian culture from collectivist to more individualist (§5). Third, Indian diaspora show a lower acceptance to intercaste marriage compared to Indians in India (§6). This result aligns with current theory (Mahalingam 2013) which contends that some diasporic individuals, because of their minority status, essentialize the notion of caste as a way maintaining cultural identity which results in resistance to intercaste marriage.

## 2 Caste System

Caste is a form of social hierarchy which assigns status to individuals based on birth. Originally based on the Hindu *varna* system which grouped people based on their occupation, caste has evolved into a form of social identity based on birth which decides hierarchy and social order in society (Srinivas 2003). The caste system divides people into groups and is organized based on perceived notions of “purity” and “pollution” (Dumont 1980). Brahmins hold the highest rank in this caste hierarchy, Other Forward Castes (OFC) are non-Brahmin upper castes, Other Backward Castes (OBC) are castes considered to be lower economically and socially, and Scheduled Castes (SC or Dalits) and Scheduled Tribes (ST) are historically disadvantaged groups. The prejudice propagated by caste through centuries denied individuals in lower caste groups dignity and self respect (Kumar 2010) as well as access to education (Chauhan 2008), job opportunities (Banerjee and Knight 1985), religious institutions (Tejani 2013) and public resources such as water (O’Reilly and Dhanju 2014). Despite increased awareness (Banerjee and Gupta 2015), government policies (Agrawal, Agrawal, and Aggarwal 1991) and work of social reformers (Ambedkar 2014), caste based discrimination is still prevalent today. This form of discrimination greatly reduces social and economic mobility among these groups and increases inequality in society (Deshpande 2011; Thorat and Neuman 2012). Caste discrimination is not limited to India, the Dalit diaspora in the US also face similar verbal, physical and workplace discrimination, with one in two Dalit respondents reporting that they fear being “outed” (Zwick-Maitreyi et al. 2017).

One of the primary institutions that reinforces caste lines is arranged marriages (Hopkins 2017), which are usually endogamous. Only 5% of marriages in India are intercaste marriages where the bride and groom belong to different caste groups (Desai and Vanneman 2015). Couples engaging in intercaste marriage face social seclusion (Parish 1996), loss of family support (Dhar 2013) and violence (Chowdhry 1997) even amounting to murder (Kumar 2012). Nevertheless, there is increasing support for intercaste marriages through NGO networks (Kumar 2012), activists (Rinker 2013) and government schemes that provide monetary support to encourage intercaste marriage. Therefore, precise measurement of changing attitudes on caste is readily needed to highlight where increased pressure and incentives can be used to ultimately bring about an equitable society.

	Self posted	Family posted	Total
Male profiles	175,611	50,245	225,856
Female profiles	39,039	48,413	87,452
<b>Total</b>	<b>214,650</b>	<b>98,658</b>	<b>313,308</b>

Table 1: The stratified matrimonial profile dataset.

Label	Caste Category	Pop. %	Data %
Brahmin	Brahmin	5.2	16.62
OFC	Other Forward Castes	22.9	27.43
OBC	Other Backward Castes	40.5	39.38
SC	Scheduled Castes	21.2	10.84
ST	Scheduled Tribes	8.6	1.74
Intercaste	Intercaste	-	0.17
Other	Other	1.3	3.82

Table 2: Groupings of caste by IHDS (Desai and Vanneman 2015) for individuals self-identifying as Hindu and their prevalence in society and on our dataset. Pop. % is the estimated percentage of population in India. Data % is the percentage of users in our dataset stratified by caste. “Others” refers to individuals who do not self-identify with any caste or for whom there is no established mapping to caste category. Data on intercaste populations are unavailable

## 3 Matrimonial Profile Data

The search for a spouse has given rise to multiple online matrimonial sites that host millions of profiles. Crucially, unlike *dating* websites, matrimonial sites are specifically tailored for marriage and include language prohibiting their use for casual relationships. Like most matchmaking websites, individuals provide detailed information including photos, a free-form self description section, family details, age, location, caste, education level, income and current job. However, unlike Western matchmaking platforms, these sites allow others (parents, siblings, other relatives and friends) to post on behalf of a person, thereby facilitating the traditional arranged marriage process in an online setting (Seth and Patnayakuni 2011). Critically, the platform directly asks individuals to indicate their openness to intercaste marriage, which we use as a proxy for their attitude toward caste.

Strong evidence indicates that the preferences expressed in these online matrimonial profiles accurately reflect offline personal attitudes in the Indian social context. Matrimonial sites in Indian communities facilitate—rather than replace—the traditional arranged marriage process, aiding the family or individual in selecting a spouse (Seth and Patnayakuni 2011; Mathur 2007). These matrimonial sites help make connections that would have been traditionally made by extended family members but are potentially more difficult to make in modernized India due to geographic mobility (Agrawal 2015). Indeed, user interviews indicate that these sites provide individuals with greater agency over the desired traits in a spouse, compared to the traditional parent-driven selection process, and therefore the expressed preferences are more closely mirror individual preferences (Titzmann 2015).

**Data Collection** Data was collected from Shaadi.com, a leading Indian matrimonial site, by creating a profile to access public profile information available on the website. This experimenter-created profile included minimal information using, whenever possible, default profile attributes and site-provided text content. Profiles included in the study were gathered by querying using the website’s profile search functionality with different combinations of age, gender, caste, and height.<sup>1</sup> Each such query returned a maximum of 600 profiles. While the data collected is not considered as human subjects research by our Institutional Review Board (IRB), efforts to remove personally-identifiable information were taken. All occurrences of individuals’ names in the free-form self description were removed. Unique identifiers associated with each profile such as names and usernames (after being used for de-duplication) were removed. We discuss in detail the ethics behind collecting this dataset and performing research on it in Section 7.

In total, through this method, we collected 411,292 profiles. Most profiles collected were in English (99.2%), and the rest was filtered out using `langdetect`. Further, to guard against survival bias, where older profiles are potentially systematically different in attributes compared to newer profiles, we restrict our analysis to the profiles created from 2017 onward. Also, as our topic of interest is openness to intercaste marriage in India and among the Indian diaspora in the US, we restrict the profiles to within the two countries. We filter out friend posted profiles (about 1% of profiles) and collectively refer to profiles posted by a parent (53.5%), sibling (43.3%), or other relatives (3.2%) as *family-posted profiles*. We note that although a sibling may have posted the profile, parents are typically heavily involved throughout the arrangement process on these sites (Seth and Patnayakuni 2011). In total, this filtering resulted in 313,308 profiles as shown in Table 1.

Caste is predominantly a Hindu phenomenon, with some scholars arguing that it is even intrinsic to Hinduism (for debate, see Rambachan (2008) and Nadkarni (2003)). Therefore, we focus our research on individuals self-identifying as Hindu, which is the social context in which the effects of caste identity are strongest and most salient. The profiles matched 352 castes, which are mapped to one of the groups shown in Table 2 using the caste classification in the nationally representative Indian Human Development Survey (IHDS) (Desai and Vanneman 2015). Estimating the proportion of individuals belonging to different caste categories in India is highly contentious (Vithayathil 2018), with the results on caste populations from the Socio Economic and Caste Census of 2011 (SECC) still yet to be released. Therefore, the population estimates for different caste categories in Table 2 are not definitive and must be treated accordingly. Broadly, relative to the available estimates of their national populations, SC and ST populations are underrepresented in our dataset, and Brahmin and OFC populations are over-represented, perhaps due to systemic economic disadvantages and differential levels of access to internet (Sreeku-

<sup>1</sup>No private data, e.g., communication, IP addresses, contact details or other user hidden data, was collected.

Variable	Description
Gender	Male, <b>Female</b>
Age	z-score of years-old
Account age	log(number of days on site)
First Marriage	Yes, <b>No</b>
In a Large City	Yes, <b>No</b>
Education	<b>High school</b> , College, Postgraduate
Income	Above Median, Below Median, <b>Not Answered</b>
Parental	Both, Father, Mother,
Employment	<b>Neither</b> , Not Answered
Affluence	Upper, Upper Middle Class, Lower-middle or Middle Class, <b>Not Answered</b>
Caste	Brahmin, OFC, OBC, SC, ST, Intercaste, <b>Other</b>

Table 3: Variables used in predicting openness, based on responses in online matrimonial profiles. The reference category for all categorical variables is bolded; ages are numeric.

mar 2007; Ahuja and Ostermann 2016). The results from our analysis are also more likely to reflect urban Indians’ attitudes on caste because of poor infrastructure and low Internet penetration in rural areas (West 2015).

## 4 Openness to Intercaste Marriage

Caste remains an critical identity variable in present-day India (Srinivas 1996; Madheswaran and Attewell 2007; Desai and Dubey 2012). The caste of a potential spouse matters for many individuals—to the point of ruling out potential mates on the basis of this variable alone (Banerjee et al. 2013). What personal features influence whether an individual is open to intercaste marriage? Prior studies have suggested that increased openness is associated with older age and increased education (Kannan 1963), being in an urban area (Reddy and Rajanna 1984), and being a member of a lower-status caste (Goli, Singh, and Sekher 2013). Yet, accurate measurements of these factors is significantly hindered by (i) cultural norms that vary substantially across states (Banerjee et al. 2013; Allendorf and Pandian 2016; Béteille 2012) and (ii) the social desirability response bias where individuals report the more socially-acceptable option (increased openness) while privately maintaining a different attitude (Edwards 1953; Paulhus 1984). Here, we mitigate both issues to analyze intercaste acceptability through large-scale analysis of the features of matrimonial profiles to ask two questions: (i) to what degree do caste attitudes differ between generations and (ii) what demographic factors are associated with increased openness to intercaste marriage.

### Data

The website allows direct measurement of openness to intercaste marriage through its registration process where new users are instructed to indicate whether caste is an important factor for their spousal choice: “Not particular about my partner’s Caste/Sect (Caste No Bar).” Notably, this option defaults to being *not open* to intercaste marriage, underscoring the societal bias against intercaste marriage. While survey responses can suffer from the social desirability response bias, the direct impact of this choice on the person (i.e., marrying that potential spouse) suggests that public

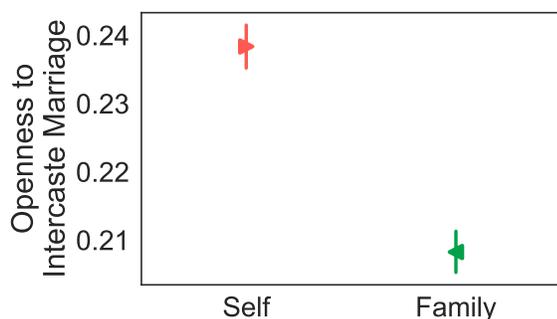


Figure 1: Under one-to-one almost exact matching conditions, profiles in which individuals posting for themselves are more open to intercaste marriages than those which were posted by family for individuals with identical demographics. Here and throughout the paper, bars show 95% confidence intervals.

and private sentiments are fully aligned. Indeed, an individual’s choice directly affects how they are categorized and to whom they are visible in the websites. Further, this choice is a blanket choice towards all members of other castes and not a response to one particular person, making it correspond more closely to a general attitude.

These matrimonial profiles include multiple demographic attributes that are selected from a fixed set of options, shown in Table 3. Affluence is a self-reported measure of socioeconomic status; income is reported in ranges or as “decline-to-report,” which we separate into three categories (i-ii) lower or higher than median and (iii) unknown. Although similar to affluence, income also serves as a measure of financial independence which is closely linked to intercaste marriage. To control for the possible effects of how long an account has been shown on the site, we include the log of the number of days, with accounts older than two years being excluded from the study altogether. To focus solely on the attitudes on Indian residents, in this analysis, we only consider the 298,400 India-based profiles of which 68.74% are self-posted.

## Methods

Inter-generational differences in acceptance of intercaste marriage are tested through two methods. First, we construct control and treatment cohorts of individuals to test for differences in openness using one-to-one almost exact matching (Rosenbaum 2010), which is analogous to techniques like propensity score matching except that it uses direct category alignment. Individuals who self post are randomly paired with an individual with identical demographics (based on Table 3) whose family had posted their profile. To increase the percentage of individuals with counterparts, we match individuals against others with a maximum age difference of 5 years and account age difference of 90 days. Owing to our large dataset size, 76.9% of family posted profiles are matched with a corresponding self-posted profile (versus 46.2% with same-age, though we note that results described

later are nearly identical). Due to almost exact matching, the difference in openness is expected to be due to the identity of profile’s author, i.e., the person or their family.

Second, we analyze the demographic factors contributing to increased openness by constructing identical logistic regression models for both self- and family-posted profiles using the variables in Table 3. Past work on caste has shown large regional differences (Allendorf and Pandian 2016), and therefore, we include a random effect for the state that the user resides in. To identify changes in the influence of the interest variables in self-arranged and family-arranged model of marriages, we compare log odds coefficients of the two models using the following Z test (Clogg, Petkova, and Haritou 1995):  $Z = \frac{\beta_S - \beta_F}{\sqrt{SE_{\beta_S}^2 + SE_{\beta_F}^2}}$  where  $\beta_S$  and  $\beta_F$  are coefficients of interest variables from both models.

## Results and Discussion

Individuals are far more open to intercaste marriage than family members who write profiles for their counterparts with near identical demographics (23.94% vs 20.82%), as shown in Figure 1. Given that parents are heavily involved when accounts are posted by a family member, this result strongly suggests a substantial inter-generational difference in caste attitudes. However, the current data cannot fully rule out selection effects to establish age as the only cause of this difference; individuals who opt to write their own profiles could also be more open than their counterparts having a profile written for them, thus introducing the potential for selection bias to affect the magnitude of the difference. Nevertheless, our result points to a substantial attitude difference in the two populations.

The demographic factors associated with openness to intercaste marriage, shown in Figure 2, suggest that a *lower* social status, as a function of income, education, and affluence, is associated with less openness to intercaste marriage. Also, the coefficients for affluence and parents’ employment are statistically indistinguishable across the two models, suggesting caste attitudes associated with these demographics are robust across generations. However, in the case of caste, the trend of higher status being more open deviates. Notably, the Brahmin caste is least open to intercaste marriage dwarfing other demographic attributes. Together, this result identifies how various intersections of identity influence attitudes toward caste, showing that demographic variables in addition to the caste hierarchy play an important role, and pointing to the need for holistic strategies for changing attitudes, not just those based on caste.

Comparing both models, we identify one important difference between self- and family-posted profiles related to education and openness. Higher education levels of the individual are positively correlated to openness in self-created profiles, whereas a person’s education level does not have a significant effect when the family creates their profiles. This result suggests that family members of more-educated individuals are less supportive of their relative being in an intercaste marriage. This infantilization of adults results in moral policing and consequently, less openness to intercaste marriage (Ganguly 2016). However, as marriages move closer to

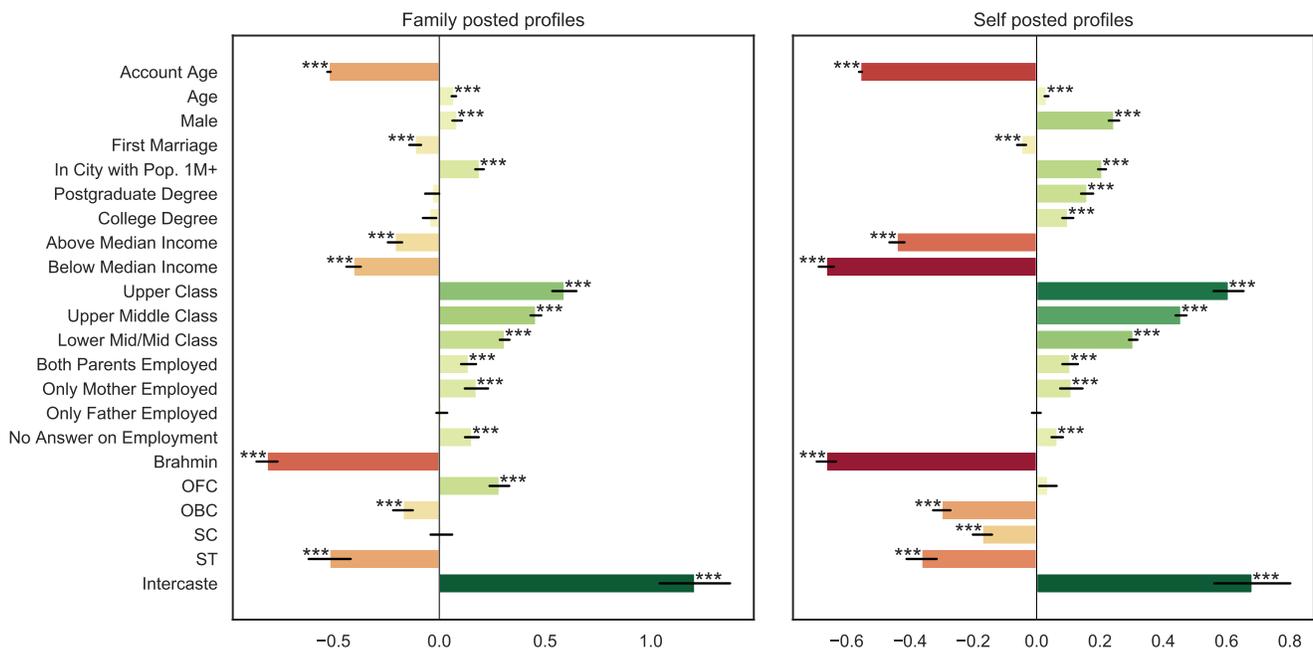


Figure 2: A comparison of demographic factors affecting openness to intercaste marriage in family-posted versus self-posted profiles shows that social status as a function of education, income, affluence, and to some degree caste, drive attitudes, where lower social status individuals are less open. Here and throughout the paper, significance levels for model coefficients are reported as ‘\*\*\*’ for  $p < 0.001$ , ‘\*\*’  $p < 0.01$ , and ‘\*’  $p < 0.05$ , and bars show standard errors.

Categories	Examples
family, profession, location	Looking for someone from a good family with a settled job in Delhi
caste, education	Seeking a suitable match of the same caste, ideally with a graduate degree
personality	Hoping for a like-minded partner who is warm, funny and supportive
education, physical attributes	Looking for a well-educated bride who is tall, fair and good-looking
caste, location	Looking for a match within the same community in the US or is willing to relocate

Table 4: Examples of preference statements found in profiles, lightly paraphrased to preserve privacy.

individual choice, higher levels of education may help break caste barriers.

Our result showing that the education of the self-poster plays a significant role in their openness directly contrasts with the result of Ray, Chaudhuri, and Sahai (2017), who found that only the education of the husband’s mother mattered but not the education of either spouse. We attribute this difference to the fact that our analysis is able to separate out *who* is initiating the search, though the family is likely to be heavily involved in the process for both (Seth and Patnayakuni 2011, p. 352). However, our result on the impact of the mother’s employment closely mirrors their result on the mother’s education (both associated with increases in openness), given that employment and education are associated (Thomas 1990; Doss 2013). Though our analyses are on slightly different data (preferences on matrimonial profiles vs. married couples), our finding presents an important insight for its implications on policy, suggesting that initiatives increasing education can potentially reduce the importance of caste in spouse selection.

Examining trends across caste identities, the relative differences in openness between castes was consistent across the models, with individuals identifying as intercaste being most open to intercaste marriage—in spite of likely being witness to and victim of their parents’ social ostracism by society (Parish 1996). Further, we find that in both models, openness to intercaste marriage follows the same order: Intercaste > OFC > SC > OBC > ST > Brahmin. Brahmins, who occupy the highest rung in the social hierarchy, are least open to intercaste marriage. Scheduled tribes are also resistant to intercaste marriage, which runs contrary to previous anthropological evidence (Sinha 1962). Considering the small sample size, the recent Hinduization of tribal identity (Gautam 2016), this phenomenon should be explored further with more qualitative research.

### Regional variations

Given known cultural differences across states (Dyson and Moore 1983), when controlling for all other demographic factors, to what degree do states vary in their openness? To

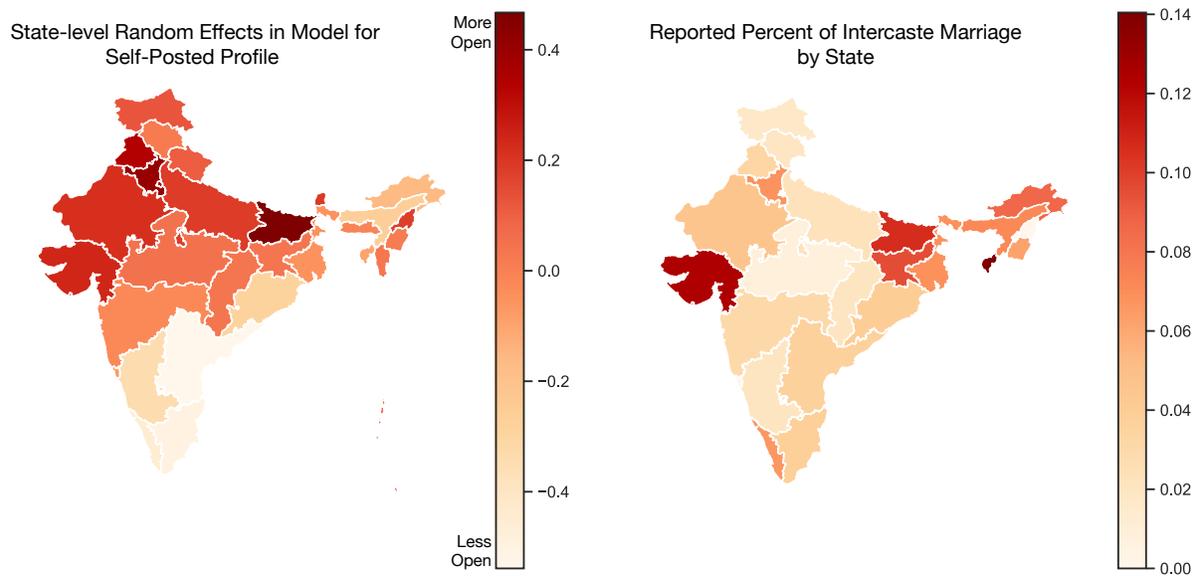


Figure 3: A comparison of the relative openness to intercaste marriage per state, controlling for demographics, (left) with the percentage of intercaste marriages (right) reveals a stark difference in attitudes versus lived reality. Here, openness is shown as states' random effect coefficients in the self-posted model, which controls for variation in demographic factors. We exclude two relatively less populous states, Sikkim and Meghalaya, from the right figure as their high rates of intermarriages skews visual interpretability.

answer this question, the states' random effect coefficients from the self-posted model are mapped and shown in Figure 3 (left),<sup>2</sup> which reveals strong regional trends: Northern states are substantially more open to intercaste marriage than Southern states. Further, considering that the Southern states consistently rank higher in gender equity (Rahman and Rao 2004) and socioeconomic levels (Ravallion and Datt 1999)—variables that are associated with higher openness in our model—we would expect the states to rank higher in openness to intercaste marriage as well. This trend is likely due to strong cultural norms around cross-cousin marriages (Trautmann 1981) and has also been observed in previous nationally representative studies (Goli, Singh, and Sekher 2013). Thus, our results suggest that while the *prima facie* expectations of openness in Southern states might be higher, when controlling for their demographic and social differences, our results show the opposite cultural norms. Compared to nationally representative percentages of actual intercaste marriages shown in Figure 3 (right), we also find substantial differences between the estimated national trends of openness to actual rates of intercaste marriage. Our result provides critical information on intercaste attitudes as relatively little information on openness to intercaste marriage is available; instead, most research relies on actual incidences of intercaste marriages, which we show differs substantially and could skew conclusions related to social inclusion.

<sup>2</sup>The map derived from family-posted accounts is highly similar and not shown here.

## 5 Attitude Shifts in Desirability

Questions about modernization and its relation to cultural norms have long been contentiously debated among scholars (Inglehart and Baker 2000): Do traditional norms persist or evolve in the face of modernization? Increased economic independence and changing family dynamics in a collectivist society have been argued to lead to a more individualistic society (Hamamura 2012). In India, the effect of globalization and urbanization has been associated with an evolution of marriage and family roles (Kashyap 2004). Given evidence of inter-generational differences in caste attitudes, does increased openness to intercaste marriage among self-posting individuals point to a shift in the attributes that are desired in a spouse, reflecting a broader shift away from caste as an important identity variable? Here, we examine profiles' text content to test for systematic shifts in spouse-seeking behavior.

### Data

Many profiles include at least one statement about their desired qualities in a spouse due to a prompt from the website to include a description of who they are looking for. We use a regular expression to extract statements about the desired attributes of the spouse as follows. Profiles are parsed into sentences and then filtered such that a sentence must contain a *partner* word and a *search* verb; here, search-related verbs are seek, need, search, look, looking, want, prefer, desire, expect and wish; and partner-related words include person, hubby, spouse, individual, soul mate, soulmate, partner, someone, some one, alliance, match, prospective and prospect. To streamline profile creating, the website pro-

Category	Example words
Age	years, young, younger
Attractiveness	beautiful, handsome, pretty
Caste	community, brahmin, iyer,
Education	well-educated, degree, IIT
Family	parents, mom, families
Finances	earning, financially, career
Health	smoking, non-smoker, teetotaler
Location	USA, mumbai, relocate
Personality	loyalty, honesty, funny
Physical attributes	slim, tall, fair
Profession	profession, work, job

Table 5: Categories of words for aspects of a desired spouse.

vides auto-generated descriptions that users can copy and use, e.g., “I am seeking someone who will be a great partner in my journey of life.” To remove possible effects from auto-generated text, we filter out any sentence repeated verbatim over 10 times from the analysis. Through this approach we identified sentences containing partner preferences from 55,520 profiles with a median length of 73 characters. Table 4 lists paraphrased examples of extracted sentences.

## Methods

To test for differences in spousal preference, we mirror prior work in testing for mentions of specific categories such as appearance (Perilloux, Fleischman, and Buss 2011). Past work in economics on dowry determinants suggested that education, height, and older men were preferred (Dalmia 2004). Other evolutionary psychology work suggests that parents look for good character, family background, similar social status, wealth, health and chastity (Apostolou 2010). Work on partner preferences in India highlighted the importance of kindness and understanding, health, mutual attraction, education and intelligence for males and females (Kamble et al. 2014). Men preferred young, physically attractive and house work oriented partners while women preferred attributes such as high social status, education, income and ambition (Kamble et al. 2014; Prakash and Singh 2014). Using these theory-based themes as reference, two annotators independently classified profile words into theory driven categories. Manual coding was done to account for emerging themes such as “location” which was not predicted by past research. After the first round of coding, discussions on word classifications and new themes were conducted before recoding, reaching high agreement at 0.91 Krippendorff’s  $\alpha$ . Categories and example words are shown in Table 5. These words are matched to the extracted preference statements in self-posted profiles.

## Results and Discussion

Both individuals who are open to intercaste marriage and those who are not have highly similar preferences in the attributes they look for in a spouse, as shown in Figure 4. This similarity implies that individuals open to intercaste marriage are no less choosy in the spousal choice. However, the relative mentions of Profession, Personality, Caste and

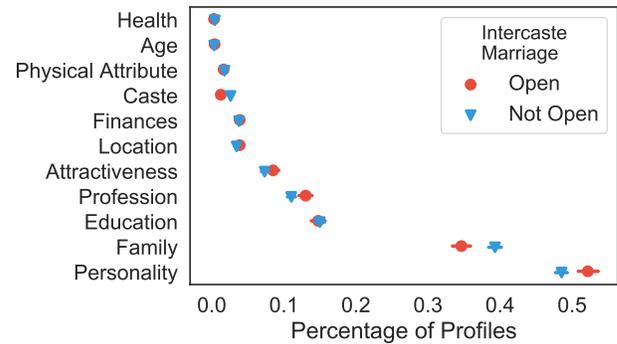


Figure 4: The percent of self-posted profiles mentioning categories of desired qualities shows individuals who are open to intercaste marriage are more likely to seek spouses on the basis their personality, rather than family background.

Family significantly differed between the groups ( $p < 0.01$ ), with the largest differences in the two most-frequently mentioned categories, Family and Personality. Specifically, we find that individuals open to intercaste marriage use a higher frequency of words related to the potential spouse’s personal qualities, whereas individuals who are not open use words more related to the social context of the spouse (caste and family). Individuals preferring to specify specific personal qualities as opposed to caste or family background in stating their preferences is a step forward in caste attitudes. A lower emphasis on family background of the prospective spouse suggests a turn towards a more individualistic concept of marriage, moving away from the traditional notion of marriage being a union of two families. Our large-scale data-driven results indicating a shift from family to personality provide further grounding to the qualitative findings from in-person interviews of individual’s preferences when seeking spouses on these sites (Titzmann 2013).

## 6 Openness to Caste in Diaspora

Economic and social mobility have led to substantial populations of Indians living abroad, both as immigrants and as citizens. Assimilation and modernization theories suggest that when immigrants move to more individualistic societies, they adopt individualism and the family plays a smaller role in individuals’ lives (Nee and Alba 2012; Tipps 1973). However, the bi-cultural integration model of acculturation (Maisuria 2004) contends that migrants retain cultural values in private settings while inculcating the host culture in professional settings. Using over 14,000 matrimonial profiles of the Indian diaspora living in the US, we test to assess whether attitudes and demographic factors associated with openness persist across cultural settings.

This study is rooted in a series of anthropological investigations of whether immigrants carry caste to their adopted lands. Studies on the Indian diaspora in the Caribbean islands report a decline and elimination of caste consciousness (Roopnarine 2006). However, similar research in the US demonstrates that caste and casteism continues to be preva-

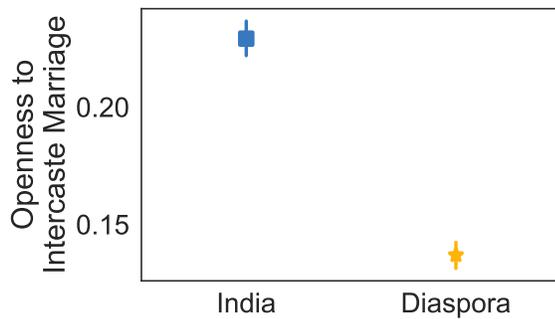


Figure 5: Under one-to-one almost exact matching conditions, individuals in the USA posting for themselves are less open to intercaste marriages than their counterparts with identical demographics in India.

lent among the Indian diaspora (Adur and Narayan 2017; Maira 2012). A crucial difference between the groups in the Caribbean and US is that the former were low-caste indentured workers with little cultural capital (Roopnarine 2003), while immigrants in the US are mostly high skilled professionals such as engineers and doctors. Thus, understanding factors that influence intercaste marriage in the US provides insight into how traditional caste boundaries are erased or reinforced in a relatively-privileged minority community within a larger western context.

### Data and Methods

We analyze data from 14,908 matrimonial profiles created by US-based users. First, we compare US based individuals to their Indian counterparts on openness to intercaste marriage. Then, we examine how demographic factors influence their openness to intercaste marriage. US and Indian profiles are compared using the one-to-one almost exact matching technique outlined in Section 4, matching users based on all demographic attributes in Table 3. 89.6% of US based individuals are matched with a random counterpart in India with identical demographics.

We repeat a similar setup as in Section 4, fitting a random effects logistic regression model on openness as the dependent variable. Here, we include the US state as a random effect and include an additional fixed effect for whether the person was raised in the US. Because of fewer profiles available for this analysis, we train a single model using both self- and family-posted profiles and include a fixed effect for whether the profile is self-posted. Based on results from Section 4 which showed factors affecting openness in self- and family-posted profiles are similar, we do not expect that using a single model will affect validity.

### Results and Discussion

Comparing the matched profiles in Figure 5, we find that individuals in the US are much less open to intercaste marriage than individuals with identical demographics in India (13.69% vs 22.91%). This result supports theory from social psychology regarding South Asian immigrants that suggests

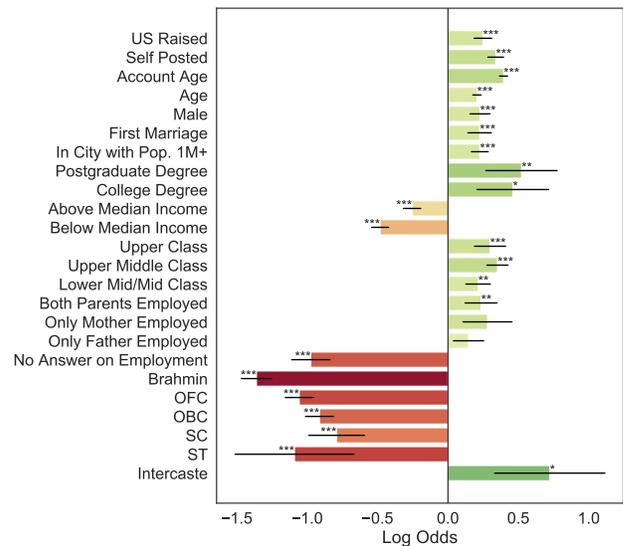


Figure 6: Demographic factors that influence openness to intercaste marriage in the diaspora, shown as their regression coefficients. These effects are largely similar to Indian residents (cf. Figure 2).

immigrants construct an essentialist notion of ethnic identity for self protection, which valorizes their culture and, in this circumstance, preserves caste salience (Mahalingam 2013).

The demographic aspects associated with increased openness to intercaste marriage in US diaspora, shown in Figure 6, largely mirror those of India-based profiles (cf. Figure 2). This result matches the expectations of the bi-cultural integration model of acculturation (Maisuria 2004) in which caste preferences are maintained in the more private setting of Indian-specific matrimonial sites, relative to the larger cultural (American) attitudes about caste. However, compared with Indian immigrants, US-raised Indians are *more* open to intercaste marriage, which supports the modernization theory that individuals will adopt aspects of the surrounding environment. Modernization theory notwithstanding, we find that the positive openness effect from being US raised is dwarfed by the resistance based on status as a function of income and the caste hierarchy—underscoring the importance of recognizing intersectional identity in understanding caste attitudes.

Similar to trends in India, the demographics of self-posting, age, residing in a large city, more education, and higher income—all signals of higher social status—are positively linked to openness. Contrary to trends in Section 4, individuals are more open to intercaste marriage during their first marriage, which should be more deeply explored. We speculate that individuals looking to get remarried may have married outside caste or ethnicity in the previous marriage, and are looking to seek in-group affinity within their own community. However, more data on their previous marriage is required to make definitive conclusions.

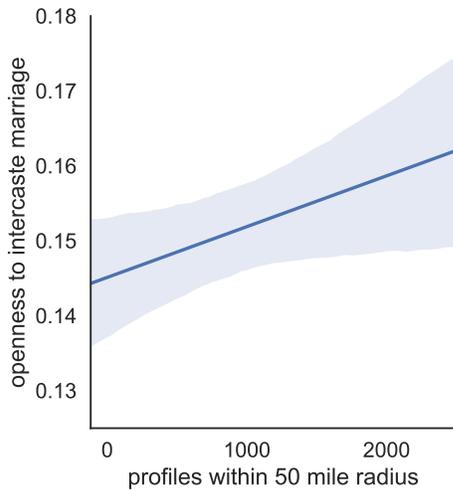


Figure 7: Openness to intercaste marriages is positively associated with the number of potential spouses nearby, shown here as the number of profiles for people within 50 miles. Shaded regions denote the 95% confidence interval for a logistic regression fit using the number of nearby profiles.

### Openness and Opportunity

Indian diaspora have settled throughout the US within varying sizes of communities. Prior studies on mate selection has found that with increased access to potential spouses, individuals become more selective (South 1991). However, prior sociological studies have also shown that urbanness is associated with increased tolerance for others of different races and background (Fischer, Merton, and Merton 1976). Given increased access to a larger pool of potential spouses who may also belong to the same caste, are individuals less open to intercaste marriage (i.e., more selective)?

To test the hypothesized relationship between openness and opportunity for marriage, we map the locations for all profiles and compute the number of profiles within a 50 mile radius as a proxy for potential matrimonial opportunity. The number of profiles correlates highly with the nearby urban population ( $r=0.90$ ) and therefore also serves as a proxy for urbanness. A logistic regression is then fit for predicting a profile's openness from the number of profiles. The results, shown in Figure 7, reveal an *increase* in openness as the number of potential spouses increase. This result suggests that increasingly urban settings are associated with less discriminatory attitudes towards caste. However, we note that this current evaluation is insufficient to establish causality; indeed, the effect could be due to a selection bias where more open individuals move towards urban areas. Nevertheless, the results do indicate that despite increased opportunities to marry within their own caste, individuals in larger cities increasingly choose to be open to intercaste marriages.

## 7 Ethical Considerations

The use of matrimonial profile data in research warrants a discussion of the ethical implications and decisions made.

The approach taken in this study is especially influenced by recent work around ethical research using social media data (Zimmer and Kinder-Kurlanda 2017; Vitak, Shilton, and Ashktorab 2016; Fiesler and Proferes 2018). In particular, precautions were taken to ensure the design, decision-making, data collection, and handling was done in an ethical manner according to recommended best practices and guidelines (Townsend and Wallace 2016; Mislove and Wilson 2018; Markham and Buchanan 2012). Here, we discuss three key aspects: (1) whether this data is public, (2) data and privacy protection practices, and (3) the overall risk-benefit trade-off in analyzing the data.

**Public or Private** Research using public information on the internet is typically not considered human subjects research (Vitak, Shilton, and Ashktorab 2016). The IRB at the authors' home institution of the University of Michigan also verbally confirmed that they do not consider work on this data to be human subjects research because all the information analyzed is public. However, the IRB is primarily concerned with legal and institutional regulations, and therefore, we consider the ethical implications of considering this data to be public. In their discussion of the process for deciding whether using particular data is ethical, Townsend and Wallace (2016) highlight a case study on data from a dating app that is a close analog to our study. Here and in the case study, individuals posting the data have the expectation that strangers will be viewing their profiles, and therefore we consider the data to be public in nature. However, as users do not expect their profiles to be publicly indexed and archived outside the site, releasing the data may violate their contextual expectations of privacy (Nissenbaum 2011). Also, previous releases of anonymized data have been de-anonymized, putting users at possible risk (Zimmer 2010). Therefore, we report only paraphrased examples and, though it limits reproducibility, opt not to release the data.

**Data Privacy and Protection** Although we consider the data to be public, substantial precautions were taken during the data collection process to ensure that no personally identifiable information was retained. We collect only the minimal data required to study the relationship of personal attributes associated with caste discrimination (e.g., education, caste identity). This process intentionally excludes data such as photos, contact details, their interactions with other users, or users' listed preference in other partner attributes, apart from openness to intercaste marriage. Once the data collection process was finished, all unique identifiers associated with each profile such as names and usernames (after being used for de-duplication) were removed. The fully-anonymized profiles are then stored on an encrypted hard disk accessible only to the researchers involved in the study. Although our data collection process required the creation of a new account, no communication or contact was made with any users on the site before, during, or after collection.

**Risk-Benefit Analysis** In critically examining this project, we aim to minimize potential harm while maximizing the

societal benefits of our research towards understanding casteism. Casteism is a serious and widespread form of social discrimination that results in major emotional, financial, and, at times, physical harm to those experiencing it, as detailed in Section 2. Our research provides valuable insights into where efforts to mitigate these discriminatory attitudes might be best placed; for example, geographic findings (Figure 3) show a stark difference in geographic trends in our data when compared to the limited data on rates of intercaste marriage, which point to where governmental and NGO efforts might be better addressed. As another example, while Ray, Chaudhuri, and Sahai (2017) found no association between discriminatory attitudes and education levels of the married couple, our regression analysis (Figure 2) shows that increased education is in fact associated with decreased discriminatory attitudes in self-posted profiles. Our research now identifies education as a potential long-term intervention strategy to change cultural attitudes in certain situations.

The largest risk from this research is loss of individuals' privacy. As a result, we have attempted to mitigate privacy concerns by anonymizing the collected data and by our decision to not share the data. We report only general trends—indeed, these general trends are what are most critical for directing efforts towards removing the scourge of casteism.

Finally, while the website's Terms of Service (TOS) prohibit crawling, our choice to disregard these terms was again motivated by a harm-benefit analysis. Our crawler issued limited queries that were appropriately spaced to have minimal impact on the site and our data privacy practices minimize the potential privacy risks to users. Our research from the resulting data presents an unprecedented view into caste attitudes that is otherwise difficult to measure and thus provides a substantial benefit towards understanding caste attitudes and their prevalence. This harm-benefit reasoning is analogous to those used by audit studies studying discrimination on platforms whose TOS prohibit such activities (Hannák et al. 2017; Chen et al. 2018).

## 8 Conclusion

Caste based discriminatory attitudes currently stigmatize millions of individuals, denying the equal access to employment, education, and even basic human rights. Governmental incentives and social movements have attempted to counter these attitudes, yet accurate measurements of public opinions on caste are unavailable despite being essential for understanding whether progress is being made. Here, we introduce a novel approach to measuring public attitudes of caste through an indicator variable: openness to intercaste marriage. Using a dataset of over 313K profiles from a major Indian matrimonial site, we precisely quantify attitudes on intercaste marriage, along with differences between generations and between Indian residents and diaspora.

Our work provides the following three main contributions towards understanding attitudes towards caste. First, we show attitudes are changing between generations, with younger individuals being more open to intercaste marriage and, in a holistic analysis of identity, show that lower social status as a function of multiple factors (e.g. income, education) is predictive of decreased openness to intercaste mar-

riage. Further, we provide the first large-scale measurement of openness to intercaste marriage, showing that attitudes are not well aligned with incidences of intercaste marriage. Second, we uncover signs of cultural shift towards individualism by examining the desired qualities in a spouse: as attitudes become more liberal towards intercaste marriage, less emphasis is on the familial aspects of a spouse and, instead, more emphasis is on their individual personality. Finally, we find that Indian diaspora are substantially less open to intercaste marriage. While some selection bias exists for diasporic individuals seeking a spouse on an Indian matrimonial site, even when controlling for where they grew up, our results support the bi-cultural theory of integration (Maisuria 2004) where individuals adopt the norms of the host country in public settings while maintaining their own culture's norms in private settings. Our research provides the first empirical evidence identifying how various intersections of identity shape attitudes toward intercaste marriage in India and among the Indian diaspora in the US, where caste hierarchy, transnational social location, social status, and other demographic variables all play an important role.

## Acknowledgements

The authors thank Ceren Budak, Paul Resnick, and the Computational Social Science seminar at UMSI for their helpful feedback on earlier drafts of this work.

## References

- Adur, S. M., and Narayan, A. 2017. Stories of dalit diaspora: Migration, life narratives, and caste in the us. *Biography* 40(1):244–264.
- Agrawal, S. P.; Agrawal, S.; and Aggarwal, J. 1991. *Educational and Social Uplift of Backward Classes: At what Cost and How?: Mandal Commission and After*, volume 26. Concept Publishing Company.
- Agrawal, A. 2015. Cyber-matchmaking among indians: Rearranging marriage and doing 'kin work'. *South Asian Popular Culture* 13(1):15–30.
- Ahuja, A., and Ostermann, S. L. 2016. Crossing caste boundaries in the modern indian marriage market. *Studies in Comparative International Development* 51(3):365–387.
- Allendorf, K., and Pandian, R. K. 2016. The decline of arranged marriage? marital change and continuity in india. *Population and development review* 42(3):435–464.
- Allendorf, K. 2013. Schemas of marital change: From arranged marriages to eloping for love. *Journal of Marriage and Family* 75(2):453–469.
- Ambedkar, B. R. 2004. Castes in india: Their mechanism, genesis and development. *Readings in Indian Government And Politics Class, Caste, Gender* 131–53.
- Ambedkar, B. R. 2014. *Annihilation of caste: The annotated critical edition*. Verso Books.
- Apostolou, M. 2010. Parental choice: What parents want in a son-in-law and a daughter-in-law across 67 pre-industrial societies. *British Journal of Psychology* 101(4):695–704.

- Banerjee, R., and Gupta, N. D. 2015. Awareness programs and change in taste-based caste prejudice. *PloS one* 10(4):e0118546.
- Banerjee, B., and Knight, J. B. 1985. Caste discrimination in the indian urban labour market. *Journal of development Economics* 17(3):277–307.
- Banerjee, A.; Duflo, E.; Ghatak, M.; and Lafortune, J. 2013. Marry for what? caste and mate selection in modern india. *American Economic Journal: Microeconomics* 5(2):33–72.
- Béteille, A. 2012. The peculiar tenacity of caste. *Economic and Political Weekly* 41–48.
- Chauhan, C. P. S. 2008. Education and caste in india. *Asia Pacific Journal of Education* 28(3):217–234.
- Chen, L.; Ma, R.; Hannák, A.; and Wilson, C. 2018. Investigating the impact of gender on rank in resume search engines. In *Proceedings of CHI*, 651. ACM.
- Chowdhry, P. 1997. Enforcing cultural codes: Gender and violence in northern india. *Economic and Political Weekly* 1019–1028.
- Clogg, C. C.; Petkova, E.; and Haritou, A. 1995. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology* 100(5):1261–1293.
- Dalmia, S. 2004. A hedonic analysis of marriage transactions in india: estimating determinants of dowries and demand for groom characteristics in marriage. *Research in Economics* 58(3):235–255.
- Desai, S., and Dubey, A. 2012. Caste in 21st century india: Competing narratives. *Economic and political weekly* 46(11):40.
- Desai, S., and Vanneman, R. 2015. India human development survey-ii (ihds-ii), 2011-12. icpsr36151-v2. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]* 31.
- Deshpande, A. 2011. *The grammar of caste: economic discrimination in contemporary India*. Oxford University Press.
- Dhar, R. L. 2013. Intercaste marriage: A study from the indian context. *Marriage & Family Review* 49(1):1–25.
- Doss, C. 2013. Intrahousehold bargaining and resource allocation in developing countries. *The World Bank Research Observer* 28(1):52–78.
- Dumont, L. 1980. *Homo hierarchicus: The caste system and its implications*. University of Chicago Press.
- Dyson, T., and Moore, M. 1983. On kinship structure, female autonomy, and demographic behavior in india. *Population and development review* 35–60.
- Edwards, A. L. 1953. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology* 37(2):90.
- Fiesler, C., and Proferes, N. 2018. “participant” perceptions of twitter research ethics. *Social Media+ Society* 4(1):2056305118763366.
- Fischer, C. S.; Merton, R. K.; and Merton, R. K. 1976. *The urban experience*. Harcourt Brace Jovanovich New York.
- Fuller, C. J., and Narasimhan, H. 2008. Companionate marriage in india: the changing marriage system in a middle-class brahman subcaste. *Journal of the Royal Anthropological Institute* 14(4):736–754.
- Ganguly, D. 2016. Privacy and consent: Public morality in india. *Social Sciences International Research Journal ISSN* 2395–0544.
- Gautam, A. 2016. The hinduization of tribals of Jharkhand: An outline since beginning. *Anthropology* 4(1).
- Goli, S.; Singh, D.; and Sekher, T. 2013. Exploring the myth of mixed marriages in india: evidence from a nation-wide survey. *Journal of comparative family studies* 44(2):193–206.
- Hamamura, T. 2012. Are cultures becoming individualistic? a cross-temporal comparison of individualism–collectivism in the united states and japan. *Personality and Social Psychology Review* 16(1):3–24.
- Hannák, A.; Wagner, C.; Garcia, D.; Mislove, A.; Strohmaier, M.; and Wilson, C. 2017. Bias in online freelance marketplaces: Evidence from taskrabbit and fiverr. In *Proceedings of CSCW, 1914–1933*. ACM.
- Hopkins, J. C. 2017. *Structure and change in Indian society*. Routledge.
- Inglehart, R., and Baker, W. E. 2000. Modernization, cultural change, and the persistence of traditional values. *American sociological review* 19–51.
- Kamble, S.; Shackelford, T. K.; Pham, M.; and Buss, D. M. 2014. Indian mate preferences: Continuity, sex differences, and cultural change across a quarter of a century. *Personality and Individual Differences* 70:150–155.
- Kannan, C. T. 1963. *Intercaste and inter-community marriages in India*. Allied Publishers.
- Kashyap, L. 2004. The impact of modernization on indian families: The counselling challenge. *International Journal for the Advancement of Counselling* 26(4):341–350.
- Krumpal, I. 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity* 47(4):2025–2047.
- Kumar, R. 2010. *Dalit Personal Narratives: Reading Caste, Nation, and Identity*. Orient Blackswan.
- Kumar, A. 2012. Public policy imperatives for curbing honour killings in india. *Journal of Politics & Governance* 1(1):33–37.
- Madheswaran, S., and Attewell, P. 2007. Caste discrimination in the indian urban labour market: Evidence from the national sample survey. *Economic and political Weekly* 4146–4153.
- Mahalingam, R. 2013. Cultural psychology of immigrants: An introduction. In *Cultural psychology of immigrants*. Psychology Press. 17–28.
- Maira, S. 2012. *Desis in the house: Indian American youth culture in NYC*. Temple University Press.
- Maisuria, C. N. 2004. *Growing up Bicultural: A qualitative study of acculturation and ethnic identity in second generation Indian Americans*. Ph.D. Dissertation, University of Hartford.
- Markham, A., and Buchanan, E. 2012. Ethical decision-making and internet research: Version 2.0. recommendations from the aoir ethics working committee. Available online: [aoir.org/reports/ethics2.pdf](http://aoir.org/reports/ethics2.pdf).

- Mathur, D. 2007. What's love got to do with it? Parental involvement and spouse choice in urban india. *Working Paper*. University of Chicago, Department of Economics.
- Mislove, A., and Wilson, C. 2018. A practitioner's guide to ethical web data collection. *The Oxford Handbook of Networked Communication*.
- Nadkarni, M. 2003. Is caste system intrinsic to hinduism? demolishing a myth. *Economic and Political Weekly* 4783–4793.
- Nee, V., and Alba, R. 2012. Rethinking assimilation theory for a new era of immigration. In *The new immigration*. Routledge. 49–80.
- Nissenbaum, H. 2011. A contextual approach to privacy online. *Daedalus* 140(4):32–48.
- O'Reilly, K., and Dhanju, R. 2014. Public taps and private connections: the production of caste distinction and common sense in a rajasthan drinking water supply project. *Transactions of the Institute of British Geographers* 39(3):373–386.
- Parish, S. M. 1996. *Hierarchy and its discontents: Culture and the politics of consciousness in caste society*. University of Pennsylvania Press.
- Paulhus, D. L. 1984. Two-component models of socially desirable responding. *Journal of personality and social psychology* 46(3):598.
- Perilloux, C.; Fleischman, D. S.; and Buss, D. M. 2011. Meet the parents: Parent-offspring convergence and divergence in mate preferences. *Personality and Individual Differences* 50(2):253–258.
- Prakash, R., and Singh, A. 2014. Who marries whom? changing mate selection preferences in urban india and emerging implications on social institutions. *Population Research and Policy Review* 33(2):205–227.
- Rahman, L., and Rao, V. 2004. The determinants of gender equity in india: examining dyson and moore's thesis with new data. *Population and Development Review* 30(2):239–268.
- Rambachan, A. 2008. Is caste intrinsic to hinduism? *Tikkun* 23(1):59–61.
- Ravallion, M., and Datt, G. 1999. *Why have some Indian states done better than others at reducing rural poverty?* The World Bank.
- Ray, T.; Chaudhuri, A.; and Sahai, K. 2017. Whose education matters? an analysis of intercaste marriage in india. Technical report, Indian Statistical Institute, Delhi Economics and Planning Unit.
- Reddy, G. S., and Rajanna, C. 1984. Intercaste marriage: A study. *Social Welfare* 31(1):10–12.
- Rinker, J. A. 2013. Why should we talk to people who do not want to talk to us? inter-caste dialogue as response to caste-based marginalization. *Peace & Change* 38(2):237–262.
- Roopnarine, L. 2003. East indian indentured emigration to the caribbean: Beyond the push and pull model. *Caribbean studies* 97–134.
- Roopnarine, L. 2006. Indo-caribbean social identity. *Caribbean quarterly* 52(1):1–11.
- Rosenbaum, P. R. 2010. *Design of observational studies*. Springer.
- Seth, N., and Patnayakuni, R. 2011. Online matrimonial sites and the transformation of arranged marriage in india. In *Virtual Communities: Concepts, Methodologies, Tools and Applications*. IGI Global. 951–974.
- Sinha, S. 1962. State formation and Rajput myth in tribal central India. *Man in India* 42(1):36–80.
- South, S. J. 1991. Sociodemographic differentials in mate selection preferences. *Journal of Marriage and the Family* 928–940.
- Sreekumar, T. 2007. Cyber kiosks and dilemmas of social inclusion in rural india. *Media, Culture & Society* 29(6):869–889.
- Srinivas, M. N. 1996. *Caste: its twentieth century avatar*. Viking.
- Srinivas, M. N. 2003. An obituary on caste as a system. *Economic and Political Weekly* 455–459.
- Tejani, S. 2013. Untouchable demands for justice or the problem of religious "non-interference": The case of temple entry movements in late-colonial india. *Journal of Colonialism and Colonial History* 14(3).
- Thomas, D. 1990. Intra-household resource allocation: An inferential approach. *Journal of Human Resources* 635–664.
- Thorat, S., and Neuman, K. S. 2012. *Blocked by caste: economic discrimination in modern India*. Oxford Univ. Press.
- Tipps, D. C. 1973. Modernization theory and the comparative study of national societies: A critical perspective. *Comparative studies in society and history* 15(2):199–226.
- Titzmann, F.-M. 2013. Changing patterns of matchmaking: The indian online matrimonial market. *Asian Journal of Women's Studies* 19(4):64–94.
- Titzmann, F.-M. 2015. Matchmaking 2.0. the representation of women and female agency in the indian online matrimonial market. *Internationales Asienforum* 42(3-4):239–256.
- Townsend, L., and Wallace, C. 2016. Social media research: A guide to ethics. *Aberdeen: University of Aberdeen*.
- Trautmann, T. R. 1981. *Dravidian kinship*. Sage Publications Pvt Limited.
- Vitak, J.; Shilton, K.; and Ashktorab, Z. 2016. Beyond the belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In *Proceedings of CSCW*, 941–953. ACM.
- Vithayathil, T. 2018. Counting caste: Censuses, politics, and castelessness in india. *Politics & Society* 46(4):455–484.
- West, D. M. 2015. Digital divide: Improving internet access in the developing world through affordable services and diverse content. *Brookings Institution*.
- Zimmer, M., and Kinder-Kurlanda, K. 2017. *Internet research ethics for the social age: New challenges, cases, and contexts*. Peter Lang International Academic Publishers.
- Zimmer, M. 2010. "but the data is already public": on the ethics of research in facebook. *Ethics and information technology* 12(4):313–325.
- Zwick-Maitreyi, M.; Soundararajan, T.; Dar, N.; Bhel, R.; and Balakrishnan, P. 2017. Caste in the United States. A Survey of Caste among South Asian Americans. Technical report, Equality Labs, USA.