Determining the Extent of Taste-Based and Statistical Discrimination: Evidence from a Field Experiment in India

Asad Islam[*]
Debayan Pakrashi[‡]
Liang Choon Wang[‡]
Yves Zenou[†]

February 2021

Abstract

We present a novel method that distinguishes choices that are unambiguously statistically discriminatory from choices consistent with both taste-based and statistical discrimination. By eliciting patients' preference rankings of physicians of different castes and experiences in a controlled field experiment and using additional information from a survey and a lab-in-the-field experiment, we estimate that taste-based discrimination explains at most 23 to 31 percent of caste discrimination in the demand for healthcare in Uttar Pradesh, India. Belief differences that are consistent with cohort variation in how Affirmative Action was implemented in higher education suggest that statistical discrimination explains caste-based discrimination in the demand for healthcare in India.

JEL codes: I15, J15, O12

Keywords: Field experiment, taste-based discrimination, statistical discrimination, health, affirmative action.

[*] Monash University, CDES and Department of Economics. Email: asadul.islam@monash.edu.
[‡] Indian Institute of Technology, Kanpur. Email: pakrashi@iitk.ac.in.
[‡] Monash University, Department of Economics. Email: liang.c.wang@monash.edu.
[†] Monash University, Department of Economics, and CEPR. Email: yves.zenou@monash.edu.

1. Introduction

The economics literature posits two major sources of discrimination: taste-based and statistical. The former is due to the fact that agents dislike some categories of the population while the latter occurs in an environment of imperfect information where agents form beliefs based on limited signals of quality. To empirically differentiate these two types of discrimination is challenging because differential treatments across groups are often consistent with the predictions of both theories. Understanding whether differences in beliefs are the main driver of agents' discriminatory behaviors is important in the context of Affirmative Action (AA), because the use of AA as a solution to the problem may unintentionally reinforce the differences in beliefs and help sustain discrimination (Coate & Loury, 1993).

We conduct a controlled field experiment across 40 localities in Uttar Pradesh, India, to examine whether taste-based discrimination or statistical discrimination is the main driver of patients' discriminatory behaviors against doctors of different castes. Uttar Pradesh is a particularly fitting setting to study this question because it is one of the Indian states with the largest concentration of lower-caste population, who are beneficiaries of a range of AA policies, including caste-based reservations for medical college seats. Through a simple theoretical framework that describes the experimental setting, we first show that taste-based discriminators, whether they have homophily or heterophily caste preference, exhibit preference rankings that completely overlap with a subset of statistical discriminators' preference rankings. The findings highlight the empirical challenge in distinguishing discriminatory behaviors that are driven by taste and those driven by beliefs. Because the predictions of taste-based discrimination overlap with a subset of predictions of statistical discrimination, our novel methodology allows us to bound the extent of taste-based discrimination and to tighten the bound under some reasonable assumptions.

To bound the extent of taste-based discrimination, we use a correspondence method to elicit 3,128 patients' rankings of four physicians of two different caste groups and two different years of experience. For each set of four physicians, a patient can have one out of 24 feasible preference rankings of doctors. Our theoretical framework shows that at most only eight out of these 24 rankings are possible under taste-based discrimination, while all 24 of them are possible under statistical discrimination. The data reveal that 36% of patients have rankings of doctors consistent with both homophily taste-based discrimination and statistical discrimination of doctors, 21% of patients have rankings consistent with both heterophily taste-based discrimination

and statistical discrimination of doctors, and 43% of patients have rankings that are unambiguously consistent with statistical discrimination of doctors. The results highlight that in the majority (57%) of cases, preference rankings of taste-based discriminators and those of statistical discriminators are indistinguishable.

We use two methods to further tighten the bound of taste-based discrimination by assuming that caste preferences are *deep* such that they similarly apply to all aspects of an individual's choices and are context invariant. In the first method, we rely on survey questions regarding attitudes towards different caste groups. In the second method, we rely on giving behaviors elicited in four (incentivized) dictator games. In both methods, we classify whether a patient's caste preference inferred in the field experiment is consistent with the patient's caste preference revealed in the survey or dictator games. For example, if a person's ranking of doctors is consistent with heterophily taste-based discrimination in the field experiment, while this person's survey attitude or giving behavior is more positive towards their own caste group, then we classify this person's ranking of doctors as consistent with statistical discrimination only (rather than both theories). We show that the percentage of taste-based discriminators decreases to 23% under the first method and 31% under the second method. Thus, by tightening the bound of taste-based discrimination, we find that the majority discriminators are likely to be statistical discriminators.

The finding that the majority of patients statistically discriminate doctors based on their castes raises the question about how caste-based quotas in higher education may shape patients' beliefs about the quality of doctors from different caste groups. In India, increasingly more medical school seats are allocated to younger cohorts of low-caste students under AA. Hence, younger cohorts of low-caste students face less competition in gaining medical school admissions relative to high-caste students as well as older cohorts of low-caste students. The opposite is true for younger cohorts of high-caste students. On the one hand, such trend may lead patients to downweigh the quality of older cohort of high-caste doctors relative to younger cohort of high-caste doctors, despite the general tendency for patients to associate better quality of doctor with more years of experience. On the other hand, cohort differences in competition for college seats and years of experience may both help to reinforce patients' beliefs that younger cohorts of low-caste doctors who benefitted the most in gaining access to medical schools are of poorer quality.

In our field experiment, we randomly assigned half of the patients to rank among doctors with four years and eight years of experience while the half of the patients to rank among doctors

with four years and 12 years of experience. We find that the return to experience from eight years to 12 years of experience is positive for low-caste doctors but negative for high-caste doctors. Furthermore, younger cohorts of low-caste doctors who benefitted the most from AA in college admissions are also the most discriminated against. These patterns of patients' preference rankings are consistent with the patterns of patients' beliefs under the specific form of caste-based AA policy implemented in India. Thus, statistical discrimination is likely the main driver of caste-based discrimination in the demand for healthcare in India.

There is a large literature focusing on detecting discrimination and examining the sources of discrimination using laboratory experiments, field experiments, natural experiments, and non-experimental approaches. In particular, the correspondence method has been the primary approach used in recent studies to investigate discrimination in a variety of settings, including employment (Bertrand & Mullainathan, 2004; Banerjee et al., 2009; Guiletti et al., 2019), housing (Ewens et al., 2014), product markets (Gneezy et al., 2012; Doleac & Stein, 2013; Zussman, 2013; Siddique et al., 2020), financial markets (Bayer et al., 2018), education (Hanna & Linden, 2012), and along different dimensions, including race, ethnicity, caste, gender, age, disability, sexual orientation, obesity, and religion.[1] However, as noted by Charles and Guryan (2011) and Neumark (2018), few studies were able to empirically differentiate behaviors consistent with taste-based discrimination and statistical discrimination, and, and when they do, it is implicitly done at the aggregate (group) level.

This paper helps advance the research on discrimination in several ways and provides evidence that caste-based statistical discrimination in the demand for healthcare is linked to caste-based AA policy in medical college admissions. First, because our approach focuses on eliciting *individual* discriminators' rankings of service providers of different types, we demonstrate that behaviors consistent with taste-based discrimination are indistinguishable from behaviors consistent with statistical discrimination, but not vice versa. The technique of eliciting individual preference rankings was previously implemented in the laboratory setting to test theories about self-regarding and other-regarding preferences (Levati et al., 2014) and correlated beliefs (Cason et al., 2020). To the best of our knowledge, this is the first time that this technique was implemented

[1] See Altonji and Blank (1999), Guryan and Charles (2013), Lang and Lehmann (2012), and Neumark (2018), for general overviews, Anderson et al. (2006) and Lane (2016), for tests of discrimination in the laboratory, Baert (2018), for correspondence experiments, and Riach and Rich (2002), Rich (2014), and Bertrand and Duflo (2017), for field experiments.

in the field setting and employed to examine the sources of discrimination in the *final stage of real transactions*. Because our approach circumvents a shortcoming of past studies on discrimination in which an *individual* discriminator's preference rankings over the choice set was unobserved, it enables us to identify individuals who are "pure" statistical discriminators and bound the extent of taste-based discrimination in various ways.

Second, we show that the overall pattern of discrimination is consistent with patients statistically discriminate against low-caste doctors and link the belief differences at the aggregate level to cohort differences in how caste-based medical college seats are implemented.[2] Our approach of linking differences in beliefs to patterns of discrimination complements the recent advancement in the literature by Bohren et al. (2019). Our focus, however, differs from Bohren et al.'s (2019), as we are agnostic whether the beliefs are biased or not. Instead, we exploit cohort differences in how caste-based AA was implemented in India to shed light on the source of belief differences at the aggregate level. We find evidence consistent with caste-based quotas in college seats fueling caste-based discrimination of college-educated professionals in India. This finding has important implications for the design and use of caste-based, race-based and gender-based AA policies in college admissions in various jurisdictions, including India, US, and Malaysia, where there have been ongoing debates about the designs and merits of AA in higher education.

2. The benchmark model

We outline the two major economic theories of discrimination in the context of our field experiment and present the situations in which the predictions of statistical discrimination completely overlap with the predictions of taste-based discrimination and the situations in which they do not. In order to match our field experiment, we assume only two types of castes for both the doctors and the patients: the low and the high caste, and two different levels (years) of experience for the doctors: low and high. We denote the caste of a doctor by $c = c_H, c_L$, where c_H corresponds to the high caste and c_L corresponds to the low caste, and the caste of a patient by $c^p = c_H^p, c_L^p$, where the superscript p refers to the patient. We denote the experience of a doctor by $e = e_H, e_L$, with $e_H > e_L$.

[2] For general overviews on the pros and cons of AA policies, see Holzer and Neumark (2000, 2006) and Arcidiacono and Lovenheim (2016). The literature examining the unintended consequences of AA policies under statistical discrimination dated back to the seminal work of Coate and Loury (1993).

2.1. Taste-based discrimination

According to Becker's (1957) theory of taste-based discrimination, prejudiced employers (or workers or consumers) dislike employing (or working with, or purchasing from) people from a certain group (e.g., race, gender, caste, etc.).

Here, we model taste-based as a consumer choice problem. The patient (consumer) chooses a doctor consumption bundle, x, which contains both the social closeness (Φ) to the doctor's caste group relative to the patient's own caste group and the perceived quality of health care (q) provided by the doctor, among the patient's consumption set X of all possible doctors. We denote by $\Phi(c^p, c)$, the social closeness of a patient from caste c^p to a doctor from caste c. Furthermore, the perceived quality of health care provided by a doctor is fully characterized by the years of experience e the doctor has been practicing medicine. On average, we expect the perceived quality to be increasing in years of experience, but there can be individual patients who perceive a doctor with fewer years of experience as one with better quality. [3] We thus allow for the possibility that $q'(e) > 0$ for some patients, but $q'(e) < 0$ for other patients. A consumption bundle $\mathrm{x} \in X$ is thus represented by a vector $\mathrm{x} \in \mathbb{R}^2_+$.

We assume that patients' preferences for doctors with various levels of social closeness and perceived health care quality satisfy five fundamental axioms of consumer choice: completeness, transitivity, continuity, strict monotonicity, and strict convexity. It follows that a patient's preference relation between two doctors with different levels of social closeness and perceived health care quality can be represented by a real-valued utility function: $\mathrm{U}: \mathbb{R}^2_+ \to \mathbb{R}, \forall \mathrm{x}^0, \mathrm{x}^1 \in \mathbb{R}^2_+$, such that $\mathrm{U}(\mathrm{x}^0) \geq U(\mathrm{x}^1) \Leftrightarrow \mathrm{x}^0 \gtrsim \mathrm{x}^1$. We denote this utility function for a patient c^p by $U(\Phi(c^p, c), q)$ and assume that it is strictly increasing in both arguments.

2.1.1. Homophily versus heterophily taste-based discrimination

[3] What matters to patients is likely not to be the experience of a doctor per se, but the quality of health care, such as the accuracy of diagnosis, the efficacy of the prescribed medicine, etc., provided by the doctor. We assume that the quality of health care can be fully described by the doctor's experience as it is the dimension that we manipulate in the experiment. In reality, patients may use experience as a noisy signal of quality $\mathbb{E}(q|e)$ to predict the quality of health care to be provided by the doctor. We can accommodate this possibility in the model by having the patient to first use the experience of the doctor to predict the quality of health care that the doctor will provide (with some errors) before choosing doctors with different levels of predicted quality and caste backgrounds. In order to put aside the possibility of statistical discrimination, we must however assume patients do not utilize the information related to the caste of a doctor when predicting the quality of health care that the doctor will provide.

We consider two possible manifestations of social closeness. The first is patients with a preference bias for own type or homophily (McPherson et al., 2001; Currarini et al., 2009), so that, for a patient of a given caste, there is a cost of interacting with a doctor from a different caste. This is a standard assumption that the discrimination literature typically makes (at least for the average individual). Thus, for a low-caste patient with homophily preference, the social closeness to a low-caste doctor is higher than to that of a high-caste doctor, $\Phi(c_L^p, c_L) > \Phi(c_L^p, c_H)$. Likewise, for a high-caste patient with homophily preference, the social closeness to a low-caste doctor is lower than to that of a high-caste doctor, $\Phi(c_H^p, c_L) < \Phi(c_H^p, c_H)$.

The second type of preference for social closeness is patients with out-group preference bias or heterophily. For this type of patient, there is a cost of interacting with a doctor from the same caste. We do not expect the majority of individuals to be this type, but it is possible that there are some of them and our theoretical framework allows for such a possibility. For a low-caste patient with heterophily preference, we have $\Phi(c_L^p, c_L) < \Phi(c_L^p, c_H)$. Likewise, for a high-caste patient with heterophily preference, we have $\Phi(c_H^p, c_L) > \Phi(c_H^p, c_H)$.

2.1.2. Rankings from the perspective of a low-caste patient

On the basis of the axioms and assumptions specified above for taste-based discriminators, we now list the possible preference rankings of doctors for a low-caste patient c_L^p who has homophily preference, given their choice set of doctors with two levels of experience e_L and e_H and two caste backgrounds c_L and c_H. There are in total 24 possible rankings of the four different doctors, and we list those that are consistent with taste-based discrimination.

For a homophily low-caste patient who views more years of experience as an indication of better health care quality, the best possibility is a doctor $c_L e_H$, i.e., a doctor from a low-caste group c_L and with a high experience level e_H, because the social closeness and quality of this doctor are both the highest. For this patient, the worst possibility is a doctor $c_H e_L$, as the social closeness and quality of the doctor are both the lowest. For a homophily low-caste patient who perceives more years of experience as better quality, there are two possible rankings of doctors that are compatible with taste-based discrimination:

$$c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$$

$$c_L e_H \succ c_H e_H \succ c_L e_L \succ c_H e_L$$

Figures 1A and 1B illustrate these two possible rankings for a homophily low-caste patient given an arbitrary utility function $U\left(\Phi\left(c_L^p, c\right), q\right)$ that satisfies the axioms and assumptions of consumer choice specified above. Note that, in the first ranking (Figure 1A), the patient exhibits stronger preference for social closeness than for quality, while in the second ranking (Figure 1B), the patient exhibits stronger preference for quality than for social closeness. We show these results an additively separable utility function that generates all rankings consistent with taste-based discrimination in Online Appendix A.

[Figure 1 here]

For a homophily low-caste patient who views fewer years of experience as an indication of better health care quality, the best possibility is a doctor $c_L e_L$ and the worst possibility is a doctor $c_H e_H$. For such a homophily low-caste patient, there are two possible rankings of doctors that are compatible with taste-based discrimination:

$$c_L e_L \succ c_L e_H \succ c_H e_L \succ c_H e_H$$
$$c_L e_L \succ c_H e_L \succ c_L e_H \succ c_H e_H$$

Figures 1C and 1D illustrate these two possible rankings for a homophily low-caste patient given an arbitrary utility function $U\left(\Phi\left(c_L^p, c\right), q\right)$ that satisfies the axioms and assumptions of consumer choice specified above.

For a heterophily low-caste patient who views more years of experience as an indication of better health care quality, the best possibility is a doctor $c_H e_H$, as both the social closeness and quality of this doctor are the highest. The worst possibility for this patient is a doctor $c_L e_L$ as both the social closeness and quality of this doctor are the lowest. For a heterophily low-caste patient, there are two possible rankings of doctors that are compatible with taste-based discrimination:

$$c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$$

For a heterophily low-caste patient who views fewer years of experience as an indication of better health care quality, there are two possible rankings of doctors that are compatible with taste-based discrimination:

$$c_H e_L \succ c_H e_H \succ c_L e_L \succ c_L e_H$$
$$c_H e_L \succ c_L e_L \succ c_H e_H \succ c_L e_H$$

Figure 2 illustrates all four possible rankings for a heterophily low-caste patient given an arbitrary utility function $U\left(\Phi\left(c_L^p, c\right), q\right)$ that satisfies the axioms and assumptions of consumer choice specified above.

[Figure 2 here]

In sum, there are 8 out of 24 possible rankings for a low-caste patient who taste-based discriminates doctors by caste.

2.1.3. Rankings from the perspective of a high-caste patient

We now provide the possible preference rankings of doctors for a high-caste patient whose preferences satisfy the axioms and assumptions specified above. We list the possible rankings of doctors for different types of high-caste patients in the following order: (1) a homophily high-caste patient who views more years of experience as an indication of health care quality; (2) a homophily high-caste patient who views fewer years of experience as an indication of health care quality; (3) a heterophily high-caste patient who views more years of experience as an indication of health care quality; and (4) a heterophily high-caste patient who views fewer years of experience as an indication of health care quality.

For a homophily high-caste patient who views more years of experience as an indication of health care quality, the two possible rankings are:

$$c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$$

Note that the above two rankings are identical to those for a heterophily low-caste patient who views more years of experience as an indication of health care quality since both patients

view a high-caste doctor being socially closer than a low-caste doctor and more years of experience as an indication of better health care quality.

For a homophily high-caste patient who views fewer years of experience as an indication of health care quality, the two possible rankings are:

$$c_H e_L \succ c_H e_H \succ c_L e_L \succ c_L e_H$$
$$c_H e_L \succ c_L e_L \succ c_H e_H \succ c_L e_H$$

Again, note that the above two rankings are identical to those for a heterophily low-caste patient who views fewer years of experience as an indication of health care quality.

For a heterophily high-caste patient who views more years of experience as an indication of health care quality, the two possible rankings are:

$$c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$$
$$c_L e_H \succ c_H e_H \succ c_L e_L \succ c_H e_L$$

The above two rankings are identical to those for a homophily low-caste patient who views more years of experience as an indication of health care quality.

For a heterophily high-caste patient who views fewer years of experience as an indication of health care quality, the two possible rankings are:

$$c_L e_L \succ c_L e_H \succ c_H e_L \succ c_H e_H$$
$$c_L e_L \succ c_H e_L \succ c_L e_H \succ c_H e_H$$

The above two rankings are identical to those for a homophily low-caste patient who views fewer years of experience as an indication of health care quality.

In sum, there are 8 out of 24 possible rankings for a high-caste patient who taste-based discriminates doctors by caste. These 8 possible rankings are identical to those for a low-caste patient who taste-based discriminates doctors by caste. Therefore, there are 16 out of 24 possible rankings that are inconsistent with taste-based discrimination.

2.2. Statistical discrimination

Phelps (1972) and Arrow (1973) pioneered statistical discrimination theory. The theory posits that, in the absence of direction information about quality, a decision maker would use group averages

(beliefs) to make inferences. For instance, labor market discrimination may exist because employers do not know with certainty workers' productivity and, therefore, may base their employment decisions on the workers' visible features, such as group identity or race, as long as these features are correlated with the unobserved productivity. This type of discrimination can result in self-fulfilling behavior from the disadvantaged groups. For example, Verdier and Zenou (2004) show that, if all agents, including blacks themselves, believe with no reason that blacks are more criminal than whites, blacks can become more criminal than whites because, based on wrong beliefs, employers pay them less, which forces them to reside far away from job centers, which leads blacks to rationally commit more crime than whites.[4]

As in the case of taste-based discrimination, we assume that a statistically discriminating patient's preference relations satisfy five fundamental axioms consumer choice: completeness, transitivity, continuity, strict monotonicity, and strict convexity. This statistically discriminating patient in our experiment, however, does not have any preference for social closeness on the basis of caste. The preference is only for the quality of health care provided. Because the actual quality of health care, q, to be delivered by a doctor is unobserved to the patient before the transaction takes place, she uses the information about the doctor's caste, c, and experience, e, to predict this quality, $\mathbb{E}(q|e,c)$. It is important to note that what matters to this patient is not the experience of a doctor per se but the quality of health care or treatment provided by the doctor. Her consumption bundle, x, contains only the expected quality of health care, $\mathbb{E}(q|e,c)$. The patient's preference relation between two doctors is *now* represented by a real-valued utility function: $\mathrm{U}: \mathbb{R}_+ \rightarrow \mathbb{R}, \forall \mathrm{x}^0, \mathrm{x}^1 \in \mathbb{R}_+$, such that $\mathrm{U}(\mathrm{x}^0) \geq U(\mathrm{x}^1) \Leftrightarrow \mathrm{x}^0 \succsim \mathrm{x}^1$.

We follow Phelps (1972) and Aigner and Cain (1977) to model statistical discrimination for our case. The experience e of a doctor from caste group c now provides a signal of the doctor's quality q with an error (noise) ε so that:

$$e = q + \varepsilon \tag{1}$$

where $\varepsilon \sim N\left(0, \sigma_{\varepsilon,c}^2\right)$ and $q \sim N\left(\beta_c, \sigma_{q,c}^2\right)$. It is assumed that $cov(q, \varepsilon) = 0$. Thus, $\mathbb{E}(e_c) = \beta_c$ and $Var(e_c) = \sigma_{q,c}^2 + \sigma_{\varepsilon,c}^2$. Each patient infers the expected value of the doctor quality q from the noisy signal e (experience) using the available information, including the caste of the doctor c. In

[4] For a substantive survey on the theory of statistical discrimination, see Fang and Moro (2011).

order to choose (rank) a doctor, each patient forms $\mathbb{E}(q|e,c)$. Since q and e are jointly normally distributed, for each caste of doctor $c = c_L, c_H$, we have:

$$\hat{q}_c \equiv \mathbb{E}(q|e,c) = (1-\gamma_c)\beta_c + \gamma_c e_c \quad (2)$$

where $0 < \gamma_c < 1$ is given by:

$$\gamma_c = \frac{\sigma_{q,c}^2}{\sigma_{q,c}^2+\sigma_{\varepsilon,c}^2} = \frac{Cov(q_c,e_c)}{Var(e_c)} \quad (3)$$

where $Cov(q_c, e_c) > 0$ as experience is a positive signal of quality according to equation (1). In other words, for a given caste of doctor, c, a doctor with higher experience is perceived to be providing a higher quality health care. On the other hand, if experience is a negative signal of quality, then $Cov(q_c, e_c) < 0$ and $-1 < \gamma_c < 0$.

Equation (2) says that $\hat{q}_c \equiv \mathbb{E}(q|e,c)$, the conditional distribution of q given e and c, follows a normal distribution with mean equal to a weighted average of the signal e_c and the unconditional group mean β_c. If the signal e_c is very noisy, i.e., the variance of ε, $\sigma_{\varepsilon,c}^2$, is large, the expected conditional value of doctor's quality is close to β_c, the population average of caste group c, regardless of the signal's value. In other words, when experience is not informative of quality, the patient uses the average quality of health care provided by the doctor's caste group to make inferences about a particular doctor's quality. On the other hand, if the signal is very precise, i.e., $\sigma_{\varepsilon,c}^2$ close to zero, then the signal e_c provides an accurate estimate of the doctor's quality. γ_c is often interpreted as the "reliability" of the signal since the higher is γ_c, the less noisy and thus more precise is the signal e_c.

2.2.1. Belief differences across caste groups

The choice of a doctor from a patient of caste c^p will depend on $\mathbb{E}(e_c) = \beta_c$, the signal e_c, and γ_c, the "reliability" of the signal. Different cases may arise depending on what the beliefs of patients are.

We are agnostic about how beliefs about doctors from different castes may arise for an *individual* patient. One can imagine that everyone starts with a common prior about the quality of doctor from various caste groups, but the idiosyncratic experiences with different types of doctors over a patient's lifetime lead to the patient having different posterior beliefs about different types of doctors. In other words, their beliefs are shaped by the draws of doctors (from different caste groups and with different years of experience) they have encountered. These draws are unlikely to

be independently and identically distributed, and the posterior beliefs that they lead to are unlikely to map to the posterior beliefs of the average patient.[5]

Consider an assumption typically made in the labor market racial discrimination literature, where the signal of labor productivity is assumed to be noisier for minority workers to explain why minority individuals with strong test performance may be discriminated on average (e.g., Aigner & Cain, 1977). In the case of an individual patient discriminates doctors based on caste, the equivalence would be to assume that the patient have the beliefs that $\beta_{c_H} = \beta_{c_L}$, $\sigma^2_{q,c_H} = \sigma^2_{q,c_L}$, $Cov(q_c, e_c) > 0$, and $\sigma^2_{\varepsilon,c_L} > \sigma^2_{\varepsilon,c_H}$. These beliefs imply that $0 < \gamma_{c_L} < \gamma_{c_H} < 1$, so that the signal about a doctor's experience is less informative the quality of health care for the lower caste c_L doctor than the higher caste c_H doctor. In this standard case, three possible rankings of doctors may arise:

$$c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_L e_L \succ c_H e_L$$

In the above three possible preference rankings of the four doctors, the first two of them are identical to the preference rankings that a taste-based discriminator (either a homophily high-caste or a heterophily low-caste patient) who views greater experience as an indicator of better health care quality may have. Only the last of the three rankings is unique for a statistical discriminator, as it violates the transitivity axiom. Panels A to B in Figure 3 illustrate these three possible rankings. We show the relationship between $\hat{q}_c$ and e_c for simplicity given that $\mathrm{U}: \mathbb{R}_+ \to \mathbb{R}, \forall \mathrm{x}^0, \mathrm{x}^1 \in \mathbb{R}_+$ such that $\mathrm{U}(\mathrm{x}^0) \geq U(\mathrm{x}^1) \Leftrightarrow \mathrm{x}^0 \gtrsim \mathrm{x}^1$ for a caste-based statistical discriminator.

[Figure 3 here]

[5] Ewens et al. (2014) shows that for racial discrimination in the US rental apartment market, when the draws are correlated within neighborhoods in which landlords are leasing apartments, the differences in variances across groups may arise since the variance includes all the pairwise correlations of observations. Bohren et al. (2019) examine whether belief differences across groups get updated as new information arrives in the context of gender discrimination in online evaluations of user-generated mathematics discussions. They find that without prior evaluations, women face significant discrimination, but the direction of discrimination reverses following a sequence of positive evaluations. Their findings imply that the beliefs are biased.

It is highly unlikely for all patients to share the beliefs that $\beta_{c_H} = \beta_{c_L}$, $\sigma^2_{q,c_H} = \sigma^2_{q,c_L}$, $Cov(q_c, e_c) > 0$, and $\sigma^2_{\varepsilon,c_L} > \sigma^2_{\varepsilon,c_H}$, even though it might be the case that the *average* patient has these beliefs. More generally, the idiosyncratic experiences of individual patients may give rise to all kinds of beliefs, including $\beta_{c_H} > \beta_{c_L}, \beta_{c_H} < \beta_{c_L}, \sigma^2_{q,c_H} > \sigma^2_{q,c_L}, \sigma^2_{q,c_H} < \sigma^2_{q,c_L}, Cov(q_c, e_c) < 0$, $\sigma^2_{\varepsilon,c_L} < \sigma^2_{\varepsilon,c_H}$, and $\sigma^2_{\varepsilon,c_L} > \sigma^2_{\varepsilon,c_H}$. Indeed, it is even possible for a patient to have the belief that $Cov(q_{c_H}, e_{c_H}) > 0$ and the belief that $Cov(q_{c_L}, e_{c_L}) < 0$, for example. Without data on the history of encounters that each patient has had, it is impossible to estimate these beliefs at the *individual* patient level. Nonetheless, we can *infer* preference rankings that are consistent with the beliefs of statistical discriminators and identify those preference rankings that are distinguishable from the preference rankings of taste-based discriminators.

2.2.2. Rankings from the perspective of any patient

We now list the patient's preference rankings of doctors that are consistent with statistical discrimination and those that overlap with taste-based discriminator's preference rankings of doctors according to the beliefs of statistical discriminators.

[Table 1 here]

Table 1 provides the list of all possible rankings of doctors under statistical discrimination by $Cov(q_c, e_c)$. In panel A, six rankings of doctors are possible when $Cov(q_c, e_c) > 0$ for both low-caste and high-caste doctors. Among these six rankings, the first four (i.e., r_1 to r_4) are identical to those under taste-based discrimination, while the last two (i.e., r_5 and r_6) are unique to statistical discrimination. Panel B shows that there are six rankings of doctors that are possible when $Cov(q_c, e_c) < 0$ for both low-caste and high-caste doctors. Similar to when $Cov(q_c, e_c) > 0$ for both low-caste and high-caste doctors, the first four rankings (i.e., r_7 to r_{10}) are identical to those under taste-based discrimination while the last two (i.e., r_{11} and r_{12}) are unique to statistical discrimination. Panels C and D show that there are 12 rankings of doctors that are possible when $Cov(q_c, e_c) > 0$ for doctors from one caste group but $Cov(q_c, e_c) < 0$ for doctors from the other caste group. Thus, all 24 possible preference rankings are consistent with the beliefs of statistical discriminators, but only eight of them (i.e., r_1 to r_4 and r_7 to r_{10}) completely overlap with the

preference rankings of taste-based discriminators. In our field experiment, it is impossible to identify whether a patient taste-based discriminates or statistically discriminates when the patient reports any one of these eight preference rankings that are consistent with both theories of discrimination.[6]

3. Context and experimental design

In this section, we first provide some background on the caste system in India and then discuss caste-based affirmative action, particularly in the higher education, that has been in place to improve the outcomes of low-caste individuals. Then, we explain the way we implemented the field experiment.

3.1. Caste system in India

The caste system in India played an important role in ancient Hindu tradition. The term caste originated from the Iberian word Casta, which means "lineage" or "breed." Most Indians use the terms "Varna" and "Jat" when referring to caste. The caste system is hierarchical. At the top of the hierarchy are the Brahmins, who were priests, teachers, and intellectuals. The second category is the Kshatriyas, who were rulers and aristocrats of the society. The next category is Vaishyas, which includes traders, landlords, farmers, and businessmen. The bottom category is Shudras, who were peasants and working class of the society. Below these castes are the outcasts who are untouchable to these four castes, who worked in degrading jobs and call themselves the Dalits.

The first three castes had social and economic rights, which the Shudras and the Dalits did not have, and were classified as high caste. The lower castes were listed in three categories – (1) Scheduled Castes (SCs) who were untouchables; (2) Scheduled Tribes (STs), who did not accept the caste system and preferred to reside in the forests and mountains; and (3) Other Backward Classes (OBCs), which include Shudras, untouchables who converted from Hinduism to other religions, nomads and tribes.

For centuries, caste dictated almost every aspect of Hindu religious and social life. Traditionally, the system bestowed many privileges on the upper castes while sanctioning repression of the lower castes by privileged groups. In recent decades, the influence of caste has

[6] Our taste-based discrimination model assumes only two goods, while patients may have preferences over more than two dimensions in reality. We assess the robustness of our results to such possibility in Online Appendix D.1.

somewhat declined, especially in cities where different castes live side-by-side and interact economically and socially. Despite the changes, caste identities remain strong, and surnames are almost always indications to which caste a person belongs. There is also evidence indicating that individuals from low caste groups continue to face discrimination, stigmatization, exclusion and rejection (Madheswaran & Attewell, 2007; Thorat & Attewell, 2007; Banerjee et al., 2009; Siddique, 2011).

3.2. Caste-based affirmative action policy in education

After independence of India, discriminating a person based on caste was legally forbidden. In 1950, the Indian government launched Affirmative Action (AA), which is known as reservation policy in India, to promote equal opportunities for SCs and STs in areas of government jobs, government-funded education, and politics (Deshpande, 2012). Following the Mandal Commission's recommendations, quotas for government jobs were extended to OBCs in the early 1990s. In 1992, the Supreme Court of India put a cap on reservation and ruled that reservations should not exceed 50%. In 2006, educational quotas for OBCs were established through the 93rd educational amendment. Government-funded colleges and universities allot seats according to caste-based quotas, which assign 7.5% to STs, 15% to SCs, and 27% to OBCs (Deshpande, 2012).

Candidates from SCs, STs, and OBCs have to take entrance examinations for higher educational institutions, but they compete only among themselves to fill the allocated number of reserved seats for their caste groups. Depending on the extent of competition within each caste group, admission requirements adjust to fill the allocated number of reserved seats. By lowering the admission requirements, high education institutions are able to increase the share of low-caste students enrolled in various medical science programs. Figure 4 shows that the share of low-caste students enrolled in various medical science programs in Indian higher education institutions has been steadily rising from a little under 40% in 2012 to near the 50% cap in 2018.

[Figure 4 here]

Several studies have examined how caste-based quotas in college admissions affect the quality of admitted students, focusing on their academic performance and subsequent earnings. For example, Bertrand et al. (2010) find the qualifying scores for admission were roughly 480 out

of a possible 900 for upper caste applicants, 419 for OBC applicants, and 182 for SC applicants among all students applying to an engineering college. Bagde et al. (2016) also find similar gradation of entrance scores by caste for a sample of 214 engineering colleges. Despite the lower qualifying scores among the admitted lower caste engineering students, Bertrand et al. (2010) find that their future earnings improve as a result of college attendance. On the other hand, Frisancho and Krishna (2016) find that in highly technical programs, SC students start with lower grades and graduate with lower grades. Furthermore, they find that SC students who enroll in more selective majors due to quotas end up earning less than what they would have earned if they had enrolled in less selective majors.

3.3. The field experiment

We conducted a field experiment to test for the presence and source of caste discrimination in the demand for health care in the Kanpur Nagar district of Uttar Pradesh (UP), India. With a population of about 225 million, UP has the most population and also the largest concentration of lower caste people among all Indian states. Caste-based issues and policies have historically dominated the state's politics.

The field experiment took place in 40 localities across the Kanpur Nagar district between August and October 2017. Figure B2 in Online Appendix B shows these 40 localities. We selected these locations because their demographic and social characteristics are representative of the overall demographic and social characteristics of the UP state. A total of 3,128 adults participated in the field experiment. Table 2 shows that the average demographic and social economic characteristics of the participants are similar to the demographic and social economic characteristics of individuals in UP state.

[Table 2 here]

We implemented the field experiment in four stages. In the first stage, participants registered interests and expressed preferences for different types of doctors presented to them. In the second stage, participants answered a short survey questionnaire. In the third stage, participants were assigned to doctors and appointments. In the fourth stage, participants received the health services.

In the first stage, we randomly approached households in each locality to advertise for an upcoming, free-of-charge health check service offered by a mobile clinic. Due to safety and ethical concerns, individuals with potential urgent and life-threatening diseases or injuries were advised to seek immediate medical attention at the local hospital, instead of waiting for the upcoming health check. At the point of registration, we requested participants to express their preferences over four potential doctors listed on a sign-up sheet. In India, mobile medical units are a common practice in places where medical facilities are inadequate, or in areas populated by low-income households. Thus, the main advantage of our field experiment is that it occurred in a "natural" environment since people in these areas have used such services. It is also common for patients to register their interest for an upcoming service and express their preference. We therefore believe that the participants did not know that they were taking part in a caste discrimination study and they acted the way they normally would.

The sign-up sheet on which participants expressed their preferences showed a two-by-two matrix containing information for four different doctors: (i) a doctor with a high-caste surname and a high number of years of experience ($c_H e_H$); (ii) a doctor with a low-caste surname and a high number of years of experience ($c_L e_H$); (iii) a doctor with a high-caste surname and a low number of years of experience ($c_H e_L$); and (iv) a doctor with a low-caste surname and a low number of years of experience ($c_L e_L$).[7] The high-caste surnames that appear on the sign-up sheet belong to the general-category (GC) caste. The low-caste surnames that appear on the sign-up sheet belong to one of the three low-caste groups (i.e., either SC, ST, or OBC). Similarly, the high number of years of experience is either 12 years or eight years but never both.[8] The low number of years of experience is always four years. We randomized the order in which each type of doctor appeared in the matrix. We did not disclose the first name of the doctor but only the initial. Participants were randomly assigned to either a female-doctor group or a male-doctor group and they were informed about the gender of the doctors. This design is to ensure that there are only two dimensions (caste group and experience level) that the four doctors differ.

Participants were instructed to rank the four doctors from their most desired (rank 1) to their least desired (rank 4), without the possibility of an equal rank. They were also explained that

[7] The sign-up sheet also indicated that, in case the participant was not assigned to any of the listed doctors, an alternative doctor would be provided.

[8] The high-caste surnames used are: Bajpai, Dixit, Mishra, and Pandey. The low-caste surnames used are: Katiyar, Pal, Rajput, Yadav, Kanaujiya, Kureel, Sonkar, and Valmiki.

they had a higher chance of getting the more preferred doctor than the less preferred doctor. There are several reasons why we did not allow participants to rank doctors equally. First, it is *impossible* to elicit true indifference. The fact that a participant chooses one doctor over another is completely consistent with the theoretical notion of indifference between two doctors, because when a patient is indifferent between two doctors the patient chooses one at random. Second, by forcing participants to give strict ranking, we prevent the situations in which participants with weak preferences give equal ranking out of social desirability concern. Third, as long as we empirically detect ranking differences by doctor type, any measurement errors due to indifferences are differenced out on average.[9]

In stage two, the participants filled out a short demographic and social economic survey. The survey collects information about their age, gender, caste identity, caste identity, educational attainment, religious affiliation, etc. The short survey also includes questions about their attitudes toward individuals of different castes. By surveying the after the elicitation exercise, we minimized any potential priming effect. The correspondence study effectively concluded by the end of stage two.

In stage three, we informed the participants about the doctor they were assigned to and the location and time of their upcoming health-check appointment. In stage four, the mobile clinic arrived in the locality to deliver service. The service was delivered within one week of registration.

4. Results

4.1. Evidence of caste discrimination

We examine whether on average there is any evidence of caste discrimination first by pooling the responses of all patients and ignore the preference rankings of doctors at the individual patient level. Specifically, we focus on the overall ranking of the four doctors by looking at the mean reverse rank (five minus the rank) that each caste-experience doctor-type receives in the full sample.

[Figure 5 here]

[9] We also show in Online Appendix D that our results are robust to the possibilities that patients may be indifferent between some doctors, do not truthfully rank some doctors, and non-discriminatory.

Figure 5 reports the aggregate ranking of the four doctors based on the pooled sample of patients from all caste groups. On average, high-caste, high-experience ($c_H e_H$) doctors receive the highest rank (mean reverse rank = 3.23), followed by low-caste, high-experience ($c_L e_H$) doctors (mean reverse rank = 3.15), high-caste, low-experience ($c_H e_L$) doctors (mean = 1.89), and low-caste, low-experience ($c_L e_L$) doctors (mean = 1.73).

If there were no caste discrimination, for the same level of experience there should not be any difference between high-caste and low-caste doctors. The difference in average ranks is statistically significant between high-caste and low-caste doctors at each level of experience ($P < 0.01$). For two doctors from different castes but with the same years of experience, the high-caste doctor is better ranked than the low-caste doctor, especially for low-experienced doctors. Such significant differences also mean that by not offering patients the possibility to rank doctors equally is not a major concern.[10] Thus, there is clear evidence of discrimination against low-caste doctors, on average.

Figure 5 indicates that the average ranking of doctors is $c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$, which is consistent with both theories of discrimination. If the average patient is a taste-based discriminator, Figure 5 implies that this average patient prefers high-caste doctors to low-caste doctors and also prefers more years of experience. Given that the majority of patients in the sample are low-caste individuals, the ranking implies that the average patient has an out-group preference (heterophily caste preference), which contradicts the assumption of in-group preference that the discrimination research typically makes about taste at the aggregate level. In contrast, if the average patient is a statistical discriminator, the ranking is consistent with the belief that $\sigma^2_{\varepsilon,c_L} > \sigma^2_{\varepsilon,c_H}$, which is the standard assumption of noisier signal for the discriminated group that most statistical discrimination models make at the aggregate level. Thus, if one follows the standard assumptions in the literature, Figure 5 suggests evidence of statistical discrimination. However, if one is not willing to impose strong assumptions about the forms of caste preference and beliefs at the aggregate level, it remains unclear whether the primary source of caste discrimination is taste-based or statistical (belief) based given that the ranking $c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$ is consistent with both theories of discrimination.

[10] If patients randomize due to indifference, we would not have detected the statistical differences.

4.2. Bounding taste-based discrimination using individual preference rankings

We now turn to the preference rankings of doctors at the individual patient level given that we cannot credibly identify the primary source of discrimination at the aggregate level without imposing strong assumptions. By examining the share of rankings that are consistent with both theories of discrimination, we now provide an upper bound of taste-based discrimination. If this upper bound is less than 50% of rankings, then we can conclude that statistical discrimination is likely the main source of caste discrimination.

Table 3 reports the share of patients who report each of the 24 possible rankings of doctors. The majority of participants have preference rankings of doctors consistent with the (expected) quality of health care is increasing with the years of experience of the doctor (i.e., the shares in column 1 sum up to approximately 80%).

[Table 3 here]

In Table 3, we quantify these differences by the type of discrimination and by caste of patients. Column (1) in Table 3 shows that, for low-caste patients, 33% of them have preference rankings of doctors consistent with homophily taste-based discrimination and statistical discrimination, 23% of them have preference rankings of doctors consistent with both heterophily taste-based discrimination and statistical discrimination, and a little over 43% of them have preference rankings of doctors consistent with only statistical discrimination. Column (2) in Table 3 shows that for high-caste patients, 44% of them have preference rankings of doctors consistent with both homophily taste-based discrimination and statistical discrimination, 15% of them have preference rankings of doctors consistent with both heterophily taste-based discrimination and statistical discrimination, and a little over 41% of them have preference rankings of doctors consistent only with statistical discrimination. Overall, low-caste patients are slightly more likely to have preference rankings consistent with only statistical discrimination, while high-caste are more likely to have preference rankings consistent with both homophily taste-based discrimination and statistical discrimination.

Finally, column (3) in Table 3 provides a more general picture of preference rankings consistent with different theories of discrimination. We see that 36% of patients have preference rankings consistent with both homophily taste-based discrimination and statistical discrimination,

21% of patients have preference rankings consistent with both heterophily taste-based discrimination and statistical discrimination, and 43% of patients have preference rankings consistent with only statistical discrimination. Thus, the upper bound of taste-based discrimination can be as high as 57% and we are unable to identify the primary source of caste discrimination.

4.3. Tightening the bound of taste-based discrimination

Given that we cannot clearly identify the primary source of discrimination by looking at the rankings of doctors at the aggregate level and at the individual patient level, we now demonstrate two methods to tighten the upper bound of taste-discrimination. In particular, these methods rely on the assumption that caste preferences are deep in the sense that they similarly permeate all aspects of an individual's choices. Given this assumption, caste preferences are context invariant. We can then further separate individual patients whose rankings are consistent with both types of discrimination in the previous section into those whose rankings are consistent with only statistical discrimination. The first method relies on survey instruments, while the second method relies on four dictator games conducted in a lab-in-the-field experiment.

By exploiting the way participants contradict themselves between the field experiment when health treatment matters, and the self-reported or the lab-in-the-field experiment about interactions with people from different castes, we are able determine individuals who are "pure" statistical discriminators among those who have preference rankings consistent with both types of discriminators. In both methods, we first identify a participant's caste preference revealed in the survey or lab-in-the-field experiment. We then classify whether this revealed caste preference contradicts the participant's caste preference inferred from their preference ranking of doctors in the field experiment. If there is a contradiction, we reclassify their preference ranking of doctors as consistent only with statistical discrimination, since all preference rankings that are consistent with taste-based discrimination are also consistent with statistical discrimination.

4.3.1. Survey method

In stage two of the field experiment, we asked participants whether they "strongly disagreed," "disagreed," "neither disagreed nor agreed," "agreed," or "strongly agreed," with a set of statements about different caste groups in order to gauge their attitudes towards individuals from different caste backgrounds. Four of these attitudinal questions are useful for us to infer their

relative taste for high-caste and low-caste persons. These four questions are as follows: (i) "It really upsets me if anyone says anything negative about people from Backward Caste"; (ii) "I have very positive attitudes towards people from Backward Caste"; (iii) "It really upsets me if anyone says anything negative about people from General Caste"; (iv) "I have very positive attitudes towards people from General Caste". We code each response on a 1-5 Likert scale, with a higher scale indicates stronger agreement. To measure a person's positive attitudes towards low-caste individuals, we compute the total points for the first two questions and denote this score as $S^p_{c_L}$. Similarly, to measure a person's positive attitudes towards high-caste individuals, we compute the total points for the last two questions and denote this score as $S^p_{c_H}$.

If a participant has relatively positive attitudes towards their own caste in the survey, then the participant exhibits homophily caste preference in general. If the participant's ranking of doctors in the field experiment is one of the four consistent with both statistical discrimination and heterophily taste-based discrimination, then we will instead classify their preference ranking of doctors as consistent with statistical discrimination. Similarly, if a participant has relatively negative attitudes towards their own caste in the survey, then the participant exhibits heterophily caste preference in general. If the participant's ranking of doctors in the field experiment is one of the four consistent with both statistical discrimination and homophily taste-based discrimination, then we will instead classify his/her preference ranking of doctors as consistent with statistical discrimination. For example, if the ranking of doctors is $c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$ for a high-caste patient, but the patient reveals $S^p_{c_H} \geq S^p_{c_L}$ in the survey, then the patient cannot taste-based discriminate against a high-caste doctor in the field experiment under the assumption that caste preference is context invariant. Similarly, if the ranking of doctors is $c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$ for a high-caste patient, but the patient reveals $S^p_{c_L} \geq S^p_{c_H}$ in the survey, then the patient cannot taste-based discriminate against a low-caste doctor in the field experiment under the assumption that caste preference is context invariant.

In the data, approximately 44% of low-caste patients and 46% of high-caste patients who have preference rankings of doctors that are consistent with homophily taste-based discrimination do not exhibit homophily preference in the survey. The patients with preference rankings of doctors that are consistent with heterophily preference but do not exhibit heterophily preference in the survey are approximately 86% for low-caste patients and 79% for high-caste patients. Thus, the majority of patients' caste preferences according to the rankings of doctors in the field

experiment actually contradict their caste preferences revealed in the survey. Panel A of Table 4 shows that we are able to tighten the upper bound of taste-discrimination to 23% after reclassifying caste preferences based on the survey method. Thus, the majority of patients statistically discriminate doctors based on their castes.

[Table 4 here]

4.3.2. Lab-in-the-field experiment

In late October 2017, we invited a random subset of the initial field-experimental participants in 30 randomly selected localities to participate in a lab-in-the-field experiment. The purpose of the lab-in-the-field experiment was to allow tightening the bound of taste-based discrimination with an incentivized method, instead of the survey method. Similar to the survey method, we reclassify patients' preference rankings of doctors by checking whether their revealed caste preferences in the lab-in-the-field experiment are consistent with their inferred caste preferences in the field experiment. In total, 482 subjects participated in the lab-in-the-field experiment.

The lab-in-the-field experiment includes four dictator games. In each dictator game, each participant, who had an endowment of 100 Rupees, decided how much to keep from this endowment (a number between 0 and 100 (inclusive) Rupees), given that what was not kept went to a randomly drawn "partner" from a particular group of participants. The four games correspond to four different groups of partners: high-caste, low-caste, above the poverty line (APL), and below the poverty line (BPL). If the participant is a low-caste person, we reminded the participant that the low-caste partner was from the same low-caste group. The anonymous partner was randomly drawn from our field experiment and the allocation was later given to them.

At the beginning of each game, an envelope with the group identity of an anonymous partner written on the envelope was drawn from a set of four envelopes. The participant was then given the group identity of the anonymous partner and the envelope with 100 Rupees (10 x Rs10 notes). The experimenter then instructed the participants to go to quiet corner to allocate whatever amount they wished for themselves and put the remaining amount in the envelope they wanted to give to the anonymous partner. They were also informed that, once they finished the task, they would drop the envelope in a bag full of similar-looking envelopes that the experimenter placed in

a different corner.[11] By letting them allocate the money in a quiet corner and drop each of the envelopes in a bag full of similar-looking envelopes away from the scrutiny of the experimenter, we ensured minimal experimenter demand effect.

Table C1 in Online Appendix C shows that the characteristics of these 482 participants in the lab-in-the-field experiment are similar to the 3,128 participants in the initial field experiment. The main findings in Table 4 are replicated when we restrict the sample to these 482 participants (Table C2 in Online Appendix C). Thus, these 482 participants behave, on average, exactly as the 3,128 participants in the field experiment. Figure C1 in Online Appendix C displays the mean amounts of giving in this lab-in-the-field experiment. There is a lot of variation in the giving behavior of individuals. Both low-caste individuals and high-caste individuals tend to give more to low-caste individuals and individuals living below the poverty line than to high-caste individuals and individuals living above the poverty line. Thus, it seems that high-caste individuals tend to exhibit heterophily preference in their giving behavior, whereas both high-caste individuals and low-caste individuals exhibit preference to help those who are financially less-privileged.

To tighten the bound of taste-based discrimination using the incentivized method, we focus on a variant of List's (2004) method to classify whether a participant's caste preference in the field experiment is consistent with the participant's caste preference in the lab-in-the-field experiment. In List's (2004) method, a person's caste preference is inferred by the amount given to one group, say the low-caste group, relative to the amount given to the other group, say the high-caste group. In our lab-in-the-field experiment, both the high-caste and low-caste participants tend to give more to individuals from a low-caste background and to individuals living below the poverty line. Since low-caste individuals are more likely to live below the poverty line, it is plausible that the tendency for all individuals to give more to low-caste individuals reflects preferences to help people in need of financial assistance. In order to have a more accurate classification of the caste preference revealed in giving behavior that is independent of the tendency to help the poor, we refine List's (2004) method to estimate the following equation by caste of patient and area to obtain the residuals $\hat{\epsilon}_i^p$:

$$G^p_{c_H,i} - G^p_{c_L,i} = \alpha_0 + \alpha_1 G^p_{APL,i} + \alpha_2 G^p_{BPL,i} + \beta' Z^p_i + \epsilon^p_i$$

[11] In the inside of each envelope, each participant's unique ID is written, so the amount can be linked to their responses in the field experiment.

where $G^p_{c_H,i}$ and $G^p_{c_L,i}$ are the amounts given to the high-caste partner and the low-caste partner by participant i, $G^p_{APL,i}$ and $G^p_{BPL,i}$ are the amounts given to the APL partner and the BPL partner by participant i, and Z^p_i is a set of characteristics of participant i, which include gender, religion, poverty status, and education.[12]

We infer a person's caste preference in the lab-in-the-field experiment using the residuals $\hat{\epsilon}^p_i$. If the residual is below zero, i.e., $\hat{\epsilon}^p_i < 0$, then the participant exhibits a preference for low caste. This is because the negative residual informs us that this participant gives relatively more to a low-caste person than a high-caste person even after considering the participant's tendency to give more to the poor (who are more likely to be low-caste). The excess giving is consistent with this patient preferring low-caste. If the residual is greater than zero, i.e., $\hat{\epsilon}^p_i > 0$, then the participant exhibits a preference for high caste.

In the data, approximately 53% of low-caste patients with preference rankings of doctors that are consistent with homophily taste-based discrimination do not exhibit homophily preference in their giving behavior. This figure is 36% for high-caste patients. On the other hand, approximately 52% of low-caste patients with preference rankings of doctors that are consistent with heterophily taste-based discrimination do not exhibit heterophily preference in their giving behavior, while approximately 39% of high-caste patients with preference rankings of doctors that are consistent with heterophily taste-based discrimination do not exhibit heterophily preference in their giving behavior. Panel B of Table 4 shows that we are able to tighten the upper bound of taste-discrimination to 31% after reclassifying caste preferences based on the incentivized method. Thus, the majority of patients statistically discriminate doctors based on their castes.

4.4. Affirmative Action and belief differences

By bounding taste-based discrimination using individual level data, we have shown that statistical discrimination is likely the major source of caste-based discrimination in the demand for health care in India. Given that statistical discrimination is a belief-based explanation, the results imply

[12] We have also performed the classification of caste preferences based on List's (2004) method. Because the estimates fall between those obtained under the survey method and this refinement of List's (2004) method, while our purpose in this section is tighten the bound of taste-based discrimination, we report the estimates Table C3 in Online Appendix C.

that the variation in the implementation of AA policy in higher education could potentially shape beliefs about different types of doctors at the aggregate level and explain the overall patterns of caste-based discriminatory behavior.

The caste-based reservation policy in higher education has resulted in an increasing share of college students enrolled in various medical science programs to come from a low-caste background. The way in which higher education institutions increase the share of low-caste students at the expense of high-caste students is achieved by setting qualifying scores for admission to be lower among low-caste applicants than among high-caste applicants (Bertrand et al., 2010; Bagde et al., 2016; Deshpande, 2012; Frisancho & Krishna, 2016). Thus, the admission requirements are relatively lower for younger cohorts of low-caste medical school students than older cohorts of low-caste medical school students, but relatively higher for younger cohorts of high-caste medical school students than older cohorts of high-caste medical school students.

If the average patients believe that the qualifying score for admission into the medical school and subsequent labor market performance of a doctor are positively related, then the relatively lower admission requirements and the truncation of qualifying scores at a lower level among older cohorts of high-caste medical school students would translate into the following beliefs: (1) lower mean and higher variance of health care quality for older cohorts of high-caste doctors than for younger cohorts of high-caste doctors; (2) higher mean and lower variance of health quality for older cohorts of low-caste doctors than for younger cohorts of low-caste doctors.

By differentiating $\hat{q}_c$, which is defined in equation (2) in Section 2.2., we obtain:

$$\frac{\partial \hat{q}_c}{\partial \beta_c} = 1 - \gamma_c \tag{4}$$

$$\frac{\partial \hat{q}_c}{\partial \sigma^2_{q,c}} = (e_c - \beta_c)\left[\frac{-\sigma^2_{q,c}}{(\sigma^2_{q,c}+\sigma^2_{\varepsilon,c})^3}\right] \tag{5}$$

Expression (4) is non-negative given that $-1 < \gamma_c < 1$. The sign of expression (5) is ambiguous, but for a sufficiently large e_c, it is likely to be negative.

With the beliefs of lower β_{c_H} and larger σ^2_{q,c_H} for older cohorts of high-caste doctors relative to younger cohorts of high-caste doctors, at a high level of experience, the expected quality of health care is likely to decrease as experience increases. In contrast, with the beliefs of higher β_{c_L} and smaller σ^2_{q,c_L} for older cohorts of low-caste doctors relative to younger cohorts of low-

caste doctors, the expected quality of health care is likely to increase as experience increases. Thus, the return to experience is likely to be larger for low-caste doctors than for high-caste doctors.

In the field experiment, we randomly assigned half of the participants to choose among doctors with four years or eight years of experience, while the other half to choose among doctors with four years or 12 years of experience. By comparing the mean difference in the reverse ranks between doctors with eight years of experience and doctors with 12 years of experience receive within the same caste group of doctors, we now examine whether there is any evidence to suggest that AA policy in higher education shapes the differences in beliefs at the aggregate level.

Figure 6 shows that the return to experience is *negative* for *high-caste doctors* between 8 years and 12 years of experience, but (weakly) *positive* for *low-caste doctors* between 8 years and 12 years of experience. The negative return to experience is significant for high-caste doctors ($P<0.01$). Thus, the return to experience is not only lower for high-caste doctors at high levels of experience, but it actually becomes negative. These results suggest that differences in beliefs at the aggregate level are consistent with how AA policy in higher education has been implemented in India. Furthermore, younger cohorts of low-caste doctors are more discriminated against when AA policy improved their access to higher educational institutions. In sum, the aggregate patterns of patients' discriminatory behavior are also consistent with the predictions of statistical discrimination under caste-based AA.

[Figure 6 here]

5. Robustness

We perform a number of robustness checks to examine whether the results are sensitive to a number of assumptions made. The results are reported and discussed in Online Appendix D. Here, we briefly summarize the key points. First, we examine whether our results are sensitive to allowing taste-based discriminatory patients to have preferences for the gender of a doctor, which is an additional unobserved good or characteristic of a doctor that is correlated with experience but not theoretically modelled under taste-based discrimination (see Online Appendix D.1). Second, we examine whether the estimated upper bound of taste-based discrimination is sensitive to the possibility that patients misreported their preference rankings strategically as in the concerns raised in the matching literature (see Online Appendix D.2). Third, we examine whether the estimated

upper bound of taste-based discrimination is robust to the possibility that patients randomized the two bottom-ranked doctors due to their lack of interest in them (see Online Appendix D.3). Fourth, we examine whether the estimated upper bound of taste-based discrimination is robust to the presence of non-discriminators by simulating various shares of non-discriminators in the sample (see Online Appendix D.4). In all four cases, we find that the estimated upper bound of taste-based discrimination remains similar to that reported in Table 4.

6. Conclusion

This paper highlights the challenge in separating the predictions of taste-based discrimination from the predictions of statistical discrimination using a simple theoretical framework that models discrimination of patients against doctors based on their castes and years of experience in a field experiment conducted in Uttar Pradesh, India. The framework yields a novel methodology to bound the extent of taste-based discrimination. Specifically, by eliciting individual patients' preference rankings of doctors, we show that patients discriminate against low-caste doctors on the average. In 57% of patients, we are unable to distinguish whether they taste-based discriminate or statistically discriminate, while we identify the remaining 43% of patients as statistical discriminators. If we assume that caste preferences are deep in the sense that they similarly permeate all aspects of an individual's choices, we can tighten the upper bound of taste-based discriminators to 23% using additional attitudinal questions collected from a survey and to 31% using incentivized giving behavior in four dictator games. Thus, by bounding taste-based discrimination at the individual level on the basis of individuals' preference rankings, we find that statistical discrimination is likely the primary source of caste discrimination in the demand for health care in India.

Because statistical discrimination is a belief-based explanation, we further examine whether there is any evidence that aggregate differences in patients' beliefs about different types of doctors are consistent with the variation in how caste-based AA has been implemented in higher educational institutions. We find that the negative return to experience for high-caste doctors with high levels of experience is consistent with this belief-based explanation under AA. Moreover, low-caste doctors suffer greater discrimination when AA policy improved their access to higher educational institutions.

The findings here have implications for the use of AA policy in improving the access to higher educational institutions and professional training programs among low-caste individuals. On the one hand, the finding that patients discriminate against low-caste doctors provides justification for the use of caste-based AA policies to address discrimination. On the other hand, we find suggestive evidence that the discrimination that low-caste doctors suffer is linked to the way in which caste-based reservation in higher educational institutions has been implemented in India. Hence, policy makers need to be cautious about the design of AA. In particular, because admission scores adjust to help fill caste quotas in college seats, the lesser competition for low-caste quotas and the stronger competition for high-caste quotas result in relative higher admission requirements for high-caste students than for low-caste students. The resulted differences in the quality of admitted students across caste groups reinforce the beliefs that low-caste graduates are less qualified when they work. Therefore, there are two countervailing effects of AA: it increases the representation of low-caste individuals in higher education and certain occupational groups but may reinforce the negative stereotype people have about the quality and productivity of low-caste individuals. It is worthwhile to consider alternative forms of AA that are less likely to reinforce negative stereotype people have against low-caste individuals. For instance, instead of using caste-based quotas in college seats, caste-based educational resources could be devoted in primary and secondary education to improve the preparation of low-caste students for college admissions.

More generally, our methodology can be applied in many other settings to understand the primary source of discrimination along different dimensions, such as race, ethnicity, gender, religion, age, etc. Given the recent major events around the world, such as the Black Lives Matter movement and Covid-19 related racism, identifying the primary source of discrimination can be a first step towards implementing appropriate policy responses to address the underlying problems.

References

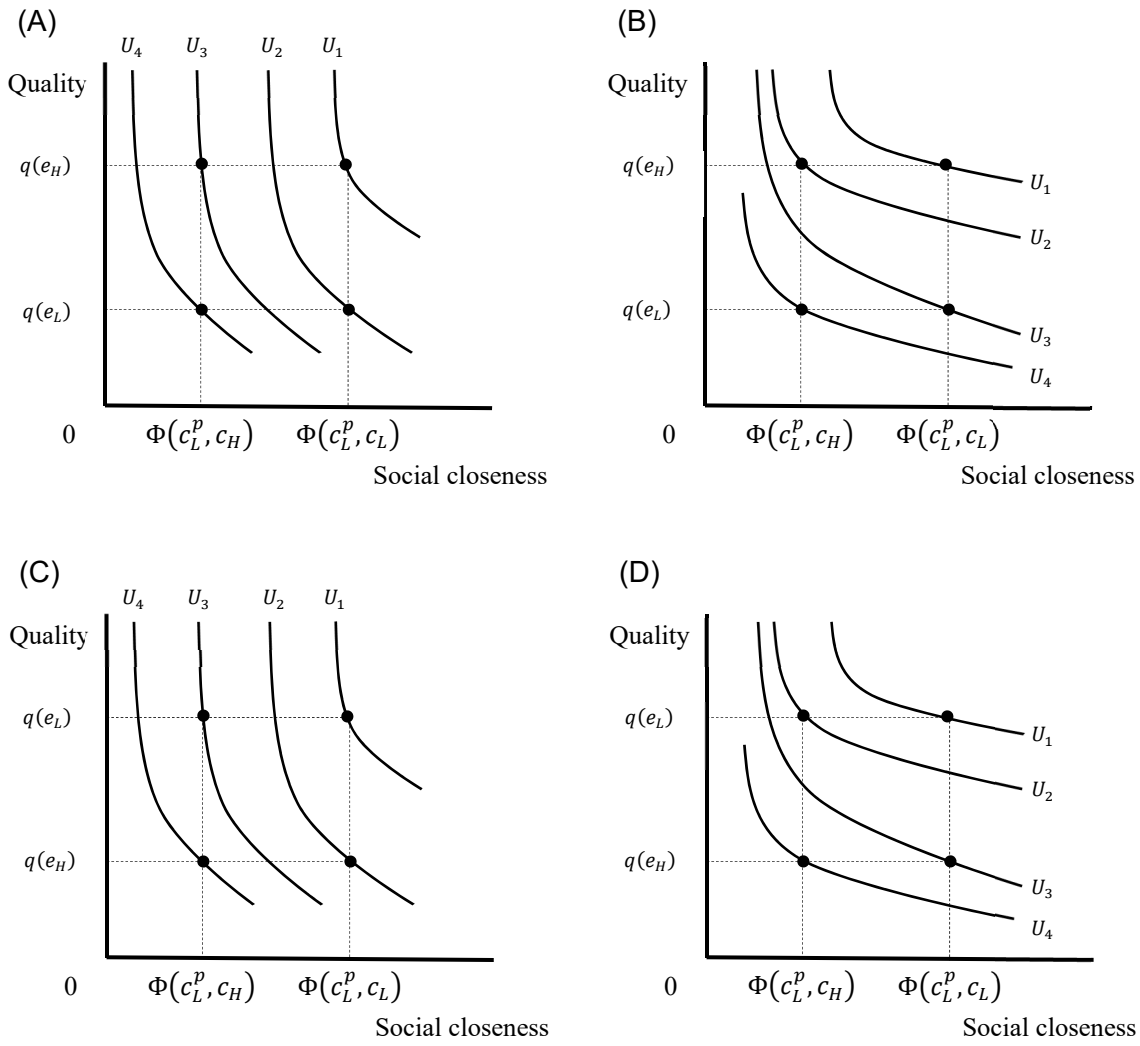
1. Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Ilr Review*, *30*(2), 175-187.
2. Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. *Handbook of labor economics*, *3*, 3143-3259.
3. Anderson, L., Fryer, R., & Holt, C. (2006). Discrimination: experimental evidence from psychology and economics. *Handbook on the Economics of Discrimination*, 97-118.
4. Arcidiacono, P., & Lovenheim, M. (2016). Affirmative action and the quality-fit trade-off. *Journal of Economic Literature*, *54*(1), 3-51.
5. Arrow, K. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Ed.), *Discrimination in Labor Markets* (pp.3-33). Princeton, NJ: Princeton University Press.
6. Baert, S. (2018). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. *Audit studies: Behind the scenes with theory, method, and nuance*, 63-77.
7. Bagde, S., Epple, D., & Taylor, L. (2016). Does affirmative action work? Caste, gender, college quality, and academic success in India. *American Economic Review*, *106*(6), 1495-1521.
8. 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.
9. Bayer, P., Ferreira, F., & Ross, S. L. (2018). What drives racial and ethnic differences in high-cost mortgages? The role of high-risk lenders. *The Review of Financial Studies*, *31*(1), 175-205.
10. Bertrand, M., Hanna, R., & Mullainathan, S. (2010). Affirmative action in education: Evidence from engineering college admissions in India. *Journal of Public Economics*, *94*(1-2), 16-29.
11. Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. *Handbook of economic field experiments*, *1*, 309-393.
12. Bertrand, M., & 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), 991-1013.
13. Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American economic review*, *109*(10), 3395-3436.
14. Cason, T. N., Sharma, T., & Vadovič, R. (2020). Correlated beliefs: Predicting outcomes in 2× 2 games. *Games and Economic Behavior*, *122*, 256-276.
15. Charles, K. K., & Guryan, J. (2011). Studying discrimination: Fundamental challenges and recent progress. *Annu. Rev. Econ.*, *3*(1), 479-511.
16. Coate, S., & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes?. *The American Economic Review*, 1220-1240.
17. Currarini, S., Jackson, M. O., & Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, *77*(4), 1003-1045.
18. Deshpande, A. (2012). Social Justice through affirmative action in India: An assessment. In *Capitalism on Trial: Explorations in the Tradition of Thomas Weisskopf*, Northampton, MA: Edward Elgar Publishing.
19. Doleac, J. L., & Stein, L. C. (2013). The visible hand: Race and online market outcomes. *The Economic Journal*, *123*(572), F469-F492.
20. Ewens, M., Tomlin, B., & Wang, L. C. (2014). Statistical discrimination or prejudice? A large sample field experiment. *Review of Economics and Statistics*, *96*(1), 119-134.

21. Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: A survey. *Handbook of social economics*, *1*, 133-200.
22. Frisancho, V., & Krishna, K. (2016). Affirmative action in higher education in India: Targeting, catch up and mismatch. *Higher Education*, 71(5), 611-649.
23. Gneezy, U., List, J., & Price, M. K. (2012). *Toward an understanding of why people discriminate: Evidence from a series of natural field experiments* (No. w17855). National Bureau of Economic Research.
24. Giulietti, C., Tonin, M., & Vlassopoulos, M. (2019). Racial discrimination in local public services: A field experiment in the United States. *Journal of the European Economic Association*, *17*(1), 165-204.
25. Guryan, J., & Charles, K. K. (2013). Taste-based or statistical discrimination: the economics of discrimination returns to its roots. *The Economic Journal*, *123*(572), F417-F432.
26. Hanna, R. N., & Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy*, *4*(4), 146-68.
27. Holzer, H., & Neumark, D. (2000). Assessing affirmative action. *Journal of Economic literature*, *38*(3), 483-568.
28. Holzer, H. J., & Neumark, D. (2006). Affirmative action: What do we know?. *Journal of policy analysis and management*, *25*(2), 463-490.
29. Lane, T. (2016). Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review*, *90*, 375-402.
30. Lang, K., & Lehmann, J. Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, *50*(4), 959-1006.
31. Levati, M. V., Nicholas, A., & Rai, B. (2014). Testing the single-peakedness of other-regarding preferences. *European Economic Review*, *67*, 197-209.
32. List, J. A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics*, *119*(1), 49-89.
33. Madheswaran, S., & Attewell, P. (2007). Caste discrimination in the Indian urban labour market: Evidence from the National Sample Survey. *Economic and Political Weekly*, 42(41), 4146-4153.
34. Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, *56*(3), 799-866.
35. McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415-444.
36. Phelps, E. S. (1972). The statistical theory of racism and sexism. *American economic review*, *62*(4), 659-661.
37. Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *The economic journal*, *112*(483), F480-F518.
38. Rich, J. (2014). *What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted since 2000* (No. 8584). IZA.
39. Siddique, Z. (2011). Evidence on caste-based discrimination. *Labour Economics*, 18, S146-S159.
40. Siddique, A., Vlassopoulos, M., & Y. Zenou (2020), "Market competition and discrimination," CEPR Discussion Paper No. 14535.
41. Thorat, S., & Attewell, P. (2007). Legacy of social exclusion: A correspondence study of job discrimination in India. *Economic and Political Weekly*, 42(41), 4141-4145.

42. Verdier, T., & Zenou, Y. (2004). Racial beliefs, location, and the causes of crime. *International Economic Review*, *45*(3), 731-760.
43. Zussman, A. (2013). Ethnic discrimination: Lessons from the Israeli online market for used cars. *The Economic Journal*, *123*(572), F433-F468.

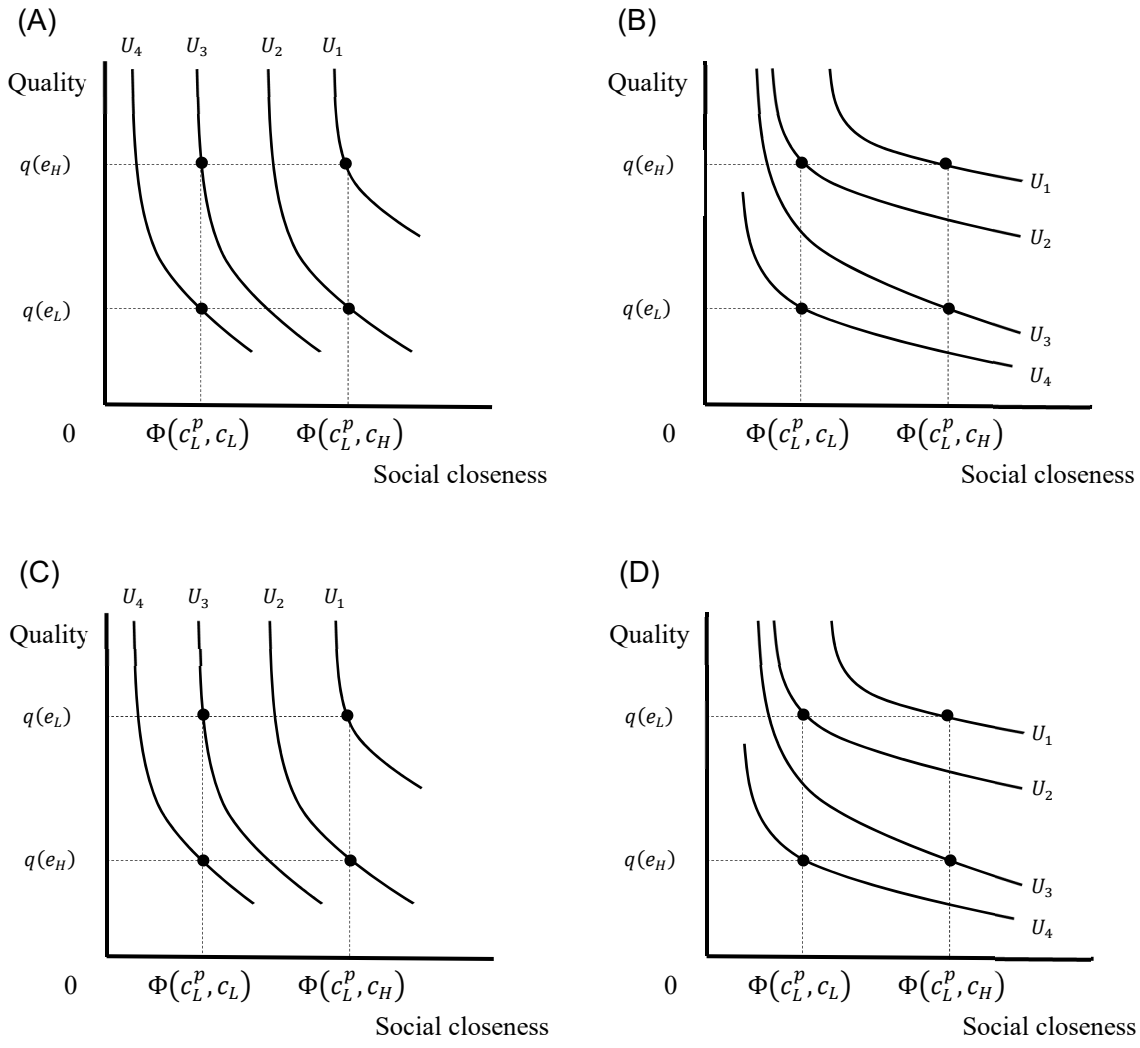
Figures

Figure 1: Low-caste homophily patients' preferences for and rakings of doctors with different levels of social closeness and quality of health care



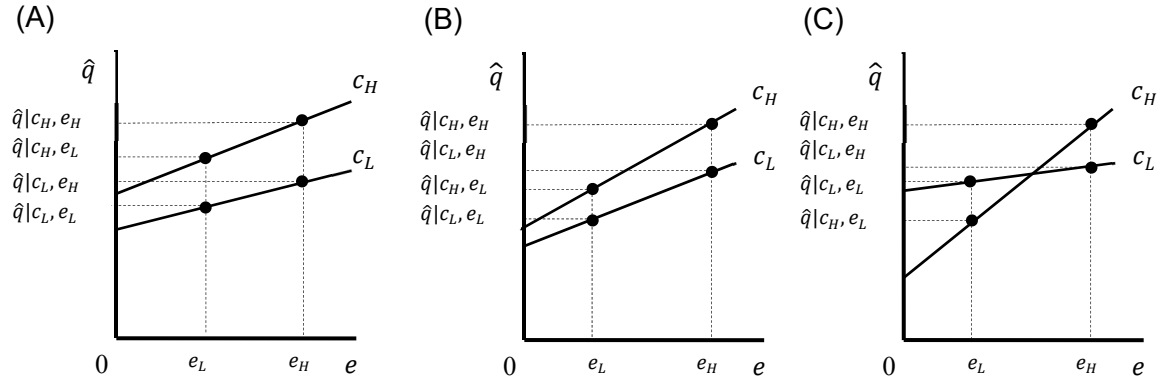
Notes: In panels A and B, the low-caste homophily patient prefers more years of experience to fewer years of experience. In panels C and D, the low-caste homophily patient prefers fewer years of experience to more years of experience. All four panels are identical for high-caste heterophily patients.

Figure 2: Low-caste heterophily patients' preferences for and rakings of doctors with different levels of social closeness and quality of health care



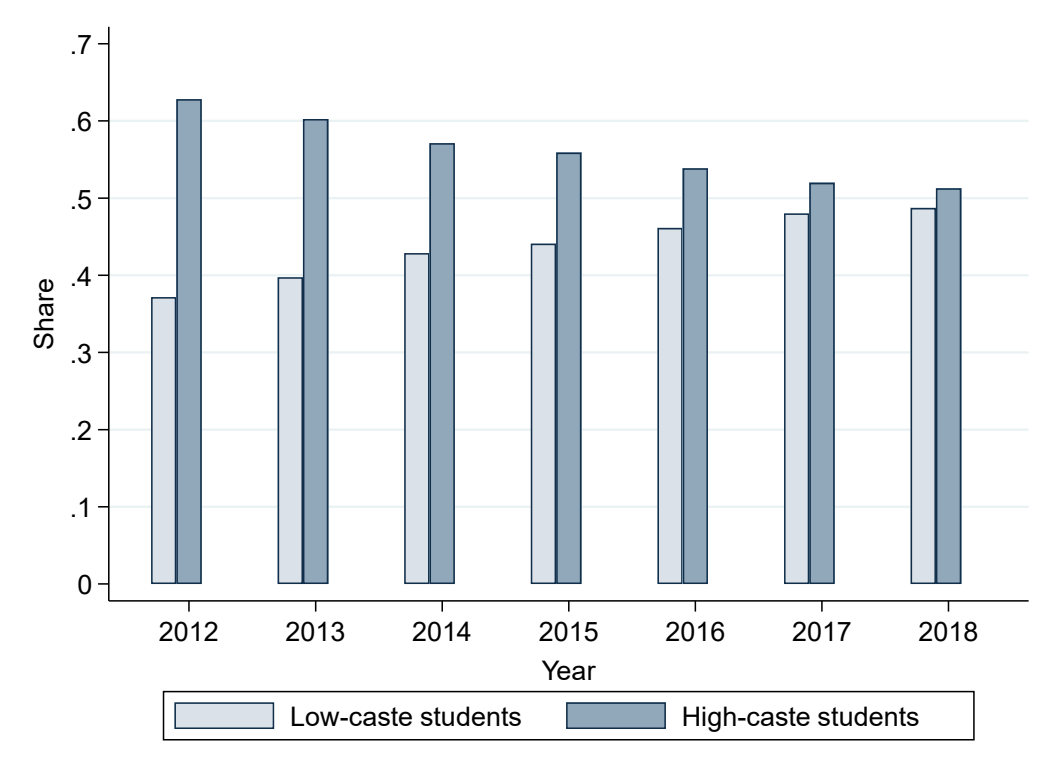
Notes: In panels A and B, the low-caste heterophily patient prefers more years of experience to fewer years of experience. In panels C and D, the low-caste heterophily patient prefers fewer years of experience to more years of experience. All four panels are identical for high-caste homophily patients.

Figure 3: The relationship between expected quality of health care and years of experience when patients statistically discriminate



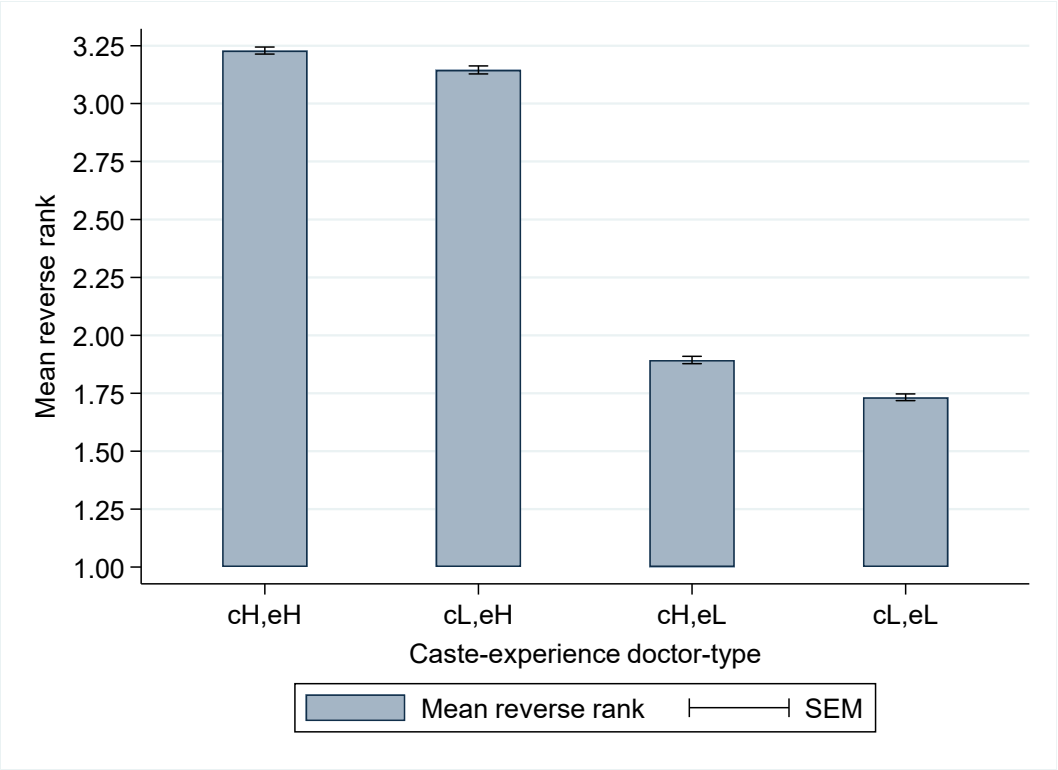
Notes: In all cases, the beliefs are assumed to be $\beta_{c_H} = \beta_{c_L}$, $\sigma^2_{q,c_H} = \sigma^2_{q,c_L}$, $Cov(q_c, e_c) > 0$, and $\sigma^2_{\varepsilon,c_L} > \sigma^2_{\varepsilon,c_H}$.

Figure 4: Share of students enrolled in medical science programs by caste and year



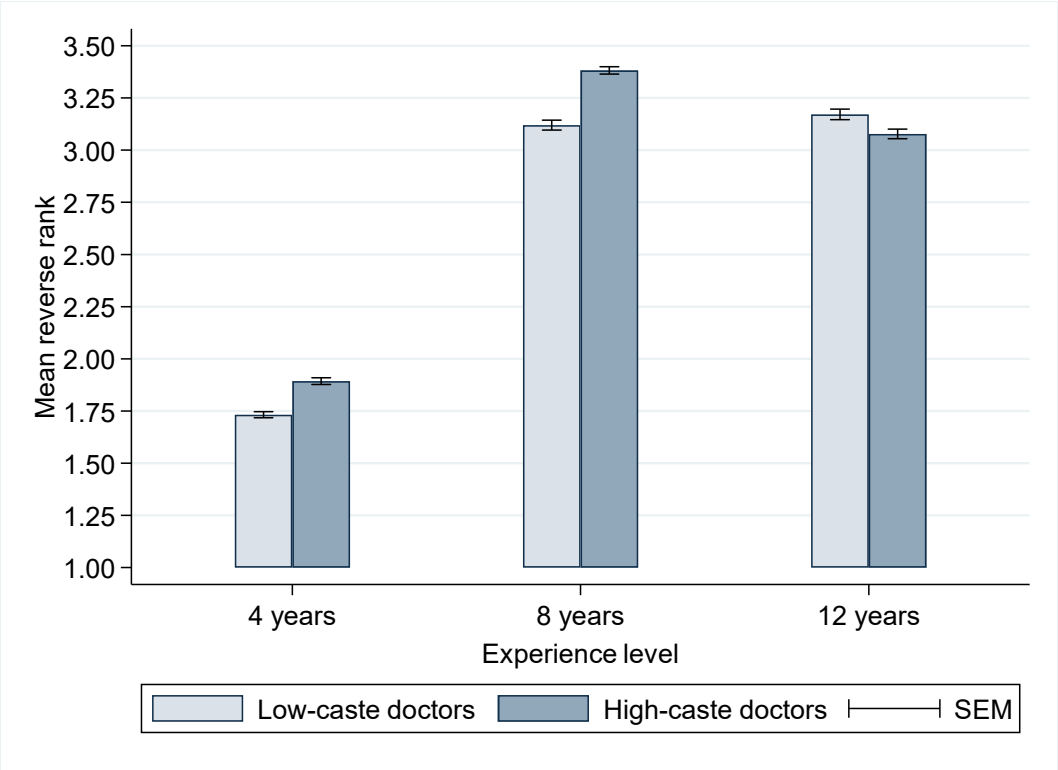
Notes: Authors' calculation based on data drawn from the All India Survey of Higher Education (AISHE). The data include tertiary students enrolled in all medical science programs (e.g., MBBS, pharmacy and pharmacology, nursing, etc.).

Figure 5: Mean reverse rank by doctor type



Notes: Reverse rank is five minus the rank of a doctor given by a patient. A doctor who receives a higher reverse rank is more preferred. cH,eH = high-caste high-experience dctor; cL,eH = low-caste high-experience doctor; cH,cL = high-caste low-experience doctor; and cL,eL = low-caste low-experience doctor. SEM denotes the standard error of mean.

Figure 6: Return to experience of doctor by caste of doctor



Notes: Reverse rank is five minus the rank of a doctor given by a patient. A doctor who receives a higher reverse rank is more preferred. SEM denotes the standard error of mean.

Tables

Table 1: All 24 possible preference rankings and the corresponding shares in the data

A. $Cov(q_c, e_c) > 0$ **for** $c = c_L, c_H$	**%**	**B.** $Cov(q_c, e_c) < 0$ **for** $c = c_L, c_H$	**%**
$r_1: c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$	7.99	$r_7: c_H e_L \succ c_H e_H \succ c_L e_L \succ c_L e_H$	1.92
$r_2: c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$	18.16	$r_8: c_H e_L \succ c_L e_L \succ c_H e_H \succ c_L e_H$	0.67
$r_3: c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$	5.47	$r_9: c_L e_L \succ c_L e_H \succ c_H e_L \succ c_H e_H$	1.18
$r_4: c_L e_H \succ c_H e_H \succ c_L e_L \succ c_H e_L$	20.97	$r_{10}: c_L e_L \succ c_H e_L \succ c_L e_H \succ c_H e_H$	0.99
$r_5: c_H e_H \succ c_L e_H \succ c_L e_L \succ c_H e_L$	11.19	$r_{11}: c_H e_L \succ c_L e_L \succ c_L e_H \succ c_H e_H$	0.51
$r_6: c_L e_H \succ c_H e_H \succ c_H e_L \succ c_L e_L$	16.08	$r_{12}: c_L e_L \succ c_H e_L \succ c_H e_H \succ c_L e_H$	0.77
Subtotal	79.86	Subtotal	6.04
C. $Cov(q_{c_H}, e_{c_H}) > 0$ **and** $Cov(q_{c_L}, e_{c_L}) < 0$	**%**	**D.** $Cov(q_{c_H}, e_{c_H}) < 0$ **and** $Cov(q_{c_L}, e_{c_L}) > 0$	**%**
$r_{13}: c_H e_H \succ c_L e_L \succ c_L e_H \succ c_H e_L$	0.32	$r_{19}: c_H e_L \succ c_L e_H \succ c_L e_L \succ c_H e_H$	0.26
$r_{14}: c_H e_H \succ c_L e_L \succ c_H e_L \succ c_L e_H$	1.05	$r_{20}: c_H e_L \succ c_L e_H \succ c_H e_H \succ c_L e_L$	0.61
$r_{15}: c_H e_H \succ c_H e_L \succ c_L e_L \succ c_L e_H$	4.48	$r_{21}: c_H e_L \succ c_H e_H \succ c_L e_H \succ c_L e_L$	2.69
$r_{16}: c_L e_L \succ c_L e_H \succ c_H e_H \succ c_H e_L$	0.80	$r_{22}: c_L e_H \succ c_L e_L \succ c_H e_L \succ c_H e_H$	2.40
$r_{17}: c_L e_L \succ c_H e_H \succ c_H e_L \succ c_L e_H$	0.32	$r_{23}: c_L e_H \succ c_H e_L \succ c_H e_H \succ c_L e_L$	0.42
$r_{18}: c_L e_L \succ c_H e_H \succ c_L e_H \succ c_H e_L$	0.32	$r_{24}: c_L e_H \succ c_H e_L \succ c_L e_L \succ c_H e_H$	0.45
Subtotal	7.29	Subtotal	6.81

Notes: All 24 preference rankings are possible under statistical discrimination. Rankings r_1 to r_4 and r_7 to r_{10} are identical to those under taste-based discrimination. The shares are based on what patients actually reported in the field experiment.

Table 2: Descriptive statistics

	Uttar Pradesh	Experimental Sample	
	Mean	Mean	Std. Dev.
Male	0.51	0.51	0.50
Age	38.0	37.8	14.3
High caste	0.27	0.27	0.44
Hindu	0.80	0.80	0.40
College educated	0.08	0.11	0.32
Below poverty line	0.29	0.34	0.47
Urban resident	0.34	0.34	0.48

Notes: The field experiment sample include 3,128 participants. All statistics for Uttar Pradesh were sourced from NSS 68th Round, 2011-2012, except the below poverty line figure which came from World Bank (2016).

Table 3: Distribution of preference rankings consistent with different theories by caste of patient

	(1) Low-caste Patients	(2) High-caste Patients	(3) All Patients
Homophily taste-based or statistical discrimination	0.334	0.440	0.363
	(0.010)	(0.017)	(0.009)
Heterophily taste-based or statistical discrimination	0.232	0.154	0.211
	(0.009)	(0.013)	(0.007)
Uniquely statistical discrimination	0.434	0.406	0.426
	(0.010)	(0.017)	(0.009)

Notes: Taste-based discrimination is indistinguishable from statistical discrimination. High experience may signal better or worse quality of health care service when classifying with which theory a particular ranking is consistent. Standard errors reported in parentheses.

Table 4: Distribution of preference rankings consistent with different theories by caste of patient after reclassification of taste-based discrimination

	(1) Low-caste Patients	(2) High-caste Patients	(3) All Patients
A. Survey method of reclassification			
Homophily taste-based or statistical discrimination	0.188	0.236	0.201
	(0.008)	(0.015)	(0.007)
Heterophily taste-based or statistical discrimination	0.031	0.032	0.032
	(0.004)	(0.006)	(0.003)
Uniquely statistical discrimination	0.780	0.732	0.767
	(0.009)	(0.015)	(0.008)
B. Refinement of List's (2004) method of reclassification			
Homophily taste-based or statistical discrimination	0.165	0.306	0.201
	(0.020)	(0.042)	(0.018)
Heterophily taste-based or statistical discrimination	0.117	0.089	0.110
	(0.017)	(0.026)	(0.014)
Uniquely statistical discrimination	0.718	0.605	0.689
	(0.024)	(0.044)	(0.021)

Notes: Taste-based discrimination is indistinguishable from statistical discrimination. High experience may signal better or worse quality of health care service when classifying with which theory a particular ranking is consistent. Standard errors reported in parentheses.

Online Appendix

A. Rankings of doctors under taste-based discrimination

A.1. Homophily taste-based discrimination

We present a simple additively separable utility function for patients of caste $c = c_L, c_H$ to illustrate the eight possible rankings under taste-based discrimination when quality of health care is permitted to be positively or negatively correlated with experience. We note that these rankings are true for any utility function that satisfies the five axioms of consumer choice specified in Section 2.1.

The utility function for a low-caste patient c_L^p choosing a doctor of caste $c = c_L, c_H$ is given by:

$$U\big(\Phi\big(c_L^p, c\big), q\big) = q(e) - 1_{\theta_H} \tag{A.1}$$

where $1_{\theta_H} = \theta_H > 0$ if the doctor is from a high caste background and zero otherwise. The indicator 1_{θ_H} indicates whether there is any social distance between the doctor's caste group and the patient's caste group.

If $q'(e) > 0$ (i.e., doctors with more years of experience provide better quality health service), the possible rankings for a low-caste patient compatible with (A.1) are:

$$c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$$
$$c_L e_H \succ c_H e_H \succ c_L e_L \succ c_H e_L$$

In the first ranking, we need to give the condition for which $c_L e_L \succ c_H e_H$ (the other inequalities are always true by definition since $e_H > e_L$). The condition is:

$$\theta_H > q(e_H) - q(e_L) \tag{A.2}$$

In the second ranking, we need to give condition for which $c_H e_H \succ c_L e_L$ (the other inequalities are always true by definition since $e_H > e_L$). The condition is:

$$\theta_H < q(e_H) - q(e_L) \tag{A.3}$$

If $q'(e) < 0$ (i.e., doctors with fewer years of experience provide better quality health service), the possible rankings for a low-caste patient compatible with (A.1) are:

$$c_L e_L \succ c_L e_H \succ c_H e_L \succ c_H e_H$$
$$c_L e_L \succ c_H e_L \succ c_L e_H \succ c_H e_H$$

In the first ranking, we need to give the condition for which $c_L e_H \succ c_H e_L$ (the other inequalities are always true by definition since $e_L < e_H$). The condition is:

$$\theta_H > q(e_L) - q(e_H) \tag{A.4}$$

In the second ranking, we need to give condition for which $c_H e_H \succ c_L e_L$ (the other inequalities are always true by definition since $e_L < e_H$). The condition is:

$$\theta_H < q(e_L) - q(e_H) \tag{A.5}$$

The utility function for a high-caste patient c_H^p choosing a doctor of caste $c = c_L, c_H$ is given by:

$$U\big(\Phi\big(c_H^p, c\big), q\big) = q(e) - 1_{\theta_L} \tag{A.6}$$

where $1_{\theta_L} = \theta_L > 0$ if the doctor is from a high caste background and zero otherwise.

If $q'(e) > 0$ (i.e., doctors with more years of experience provide better quality health service), the possible rankings for a high-caste patient compatible with (A.6) are:

$$c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$$

In the first ranking, we need to give the condition for which $c_H e_L \succ c_L e_H$ (the other inequalities are always true by definition since $e_H > e_L$). The condition is:

$$\theta_L > q(e_H) - q(e_L) \tag{A.7}$$

In the second ranking, we need to give condition for which $c_L e_H \succ c_H e_L$ (the other inequalities are always true by definition since $e_H > e_L$). The condition is:

$$\theta_L < q(e_H) - q(e_L) \tag{A.8}$$

If $q'(e) < 0$ (i.e., doctors with fewer years of experience provide better quality health service), the possible rankings for a high-caste patient compatible with (A.6) are:

$$c_H e_L \succ c_H e_H \succ c_L e_L \succ c_L e_H$$
$$c_H e_L \succ c_L e_L \succ c_H e_H \succ c_L e_H$$

In the first ranking, we need to give the condition for which $c_H e_H \succ c_L e_L$ (the other inequalities are always true by definition since $e_L < e_H$). The condition is:

$$\theta_L > q(e_L) - q(e_H) \tag{A.9}$$

In the second ranking, we need to give condition for which $c_L e_L \succ c_H e_H$ (the other inequalities are always true by definition since $e_L < e_H$). The condition is:

$$\theta_L < q(e_L) - q(e_H) \tag{A.10}$$

To summarize, with taste-based discrimination, there are eight possible rankings in total. The four possible rankings for a low-caste homophily patient that are compatible with (A.1) are:

$$c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$$
$$c_L e_H \succ c_H e_H \succ c_L e_L \succ c_H e_L$$
$$c_L e_L \succ c_L e_H \succ c_H e_L \succ c_H e_H$$
$$c_L e_L \succ c_H e_L \succ c_L e_H \succ c_H e_H$$

The four possible rankings for a high-caste homophily patient that are compatible with (A.6) are:

$$c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$$
$$c_H e_L \succ c_H e_H \succ c_L e_L \succ c_L e_H$$
$$c_H e_L \succ c_L e_L \succ c_H e_H \succ c_L e_H$$

A.2. Heterophily taste-based discrimination

The utility function for a low-caste patient c_L^p choosing a doctor of caste $c = c_L, c_H$ is given by:

$$U\big(\Phi\big(c_L^p, c\big), q\big) = q(e) - 1_{\theta_L} \tag{A.11}$$

where $1_{\theta_L} = \theta_L > 0$ if the doctor is from a low caste background and zero otherwise.

If $q'(e) > 0$, the possible rankings for a low-caste patient compatible with (A.11) are:

$$c_H e_H \succ c_H e_L \succ c_L e_H \succ c_L e_L$$
$$c_H e_H \succ c_L e_H \succ c_H e_L \succ c_L e_L$$

If $q'(e) < 0$, the possible rankings for a low-caste patient compatible with (A.11) are:

$$c_H e_L \succ c_H e_H \succ c_L e_L \succ c_L e_H$$
$$c_H e_L \succ c_L e_L \succ c_H e_H \succ c_L e_H$$

The utility function for a high-caste patient c_H^p choosing a doctor of caste $c = c_L, c_H$ is given by:

$$U\left(\Phi\left(c_H^p, c\right), q\right) = q(e) - 1_{\theta_H} \tag{A.12}$$

where $1_{\theta_H} = \theta_H > 0$ if the doctor is from a high caste background and zero otherwise.

If $q'(e) > 0$, the possible rankings for a high-caste patient compatible with (A.12) are:

$$c_L e_H \succ c_L e_L \succ c_H e_H \succ c_H e_L$$
$$c_L e_H \succ c_H e_H \succ c_L e_L \succ c_H e_L$$

If $q'(e) < 0$, the possible rankings for a high-caste patient compatible with (A.12) are:

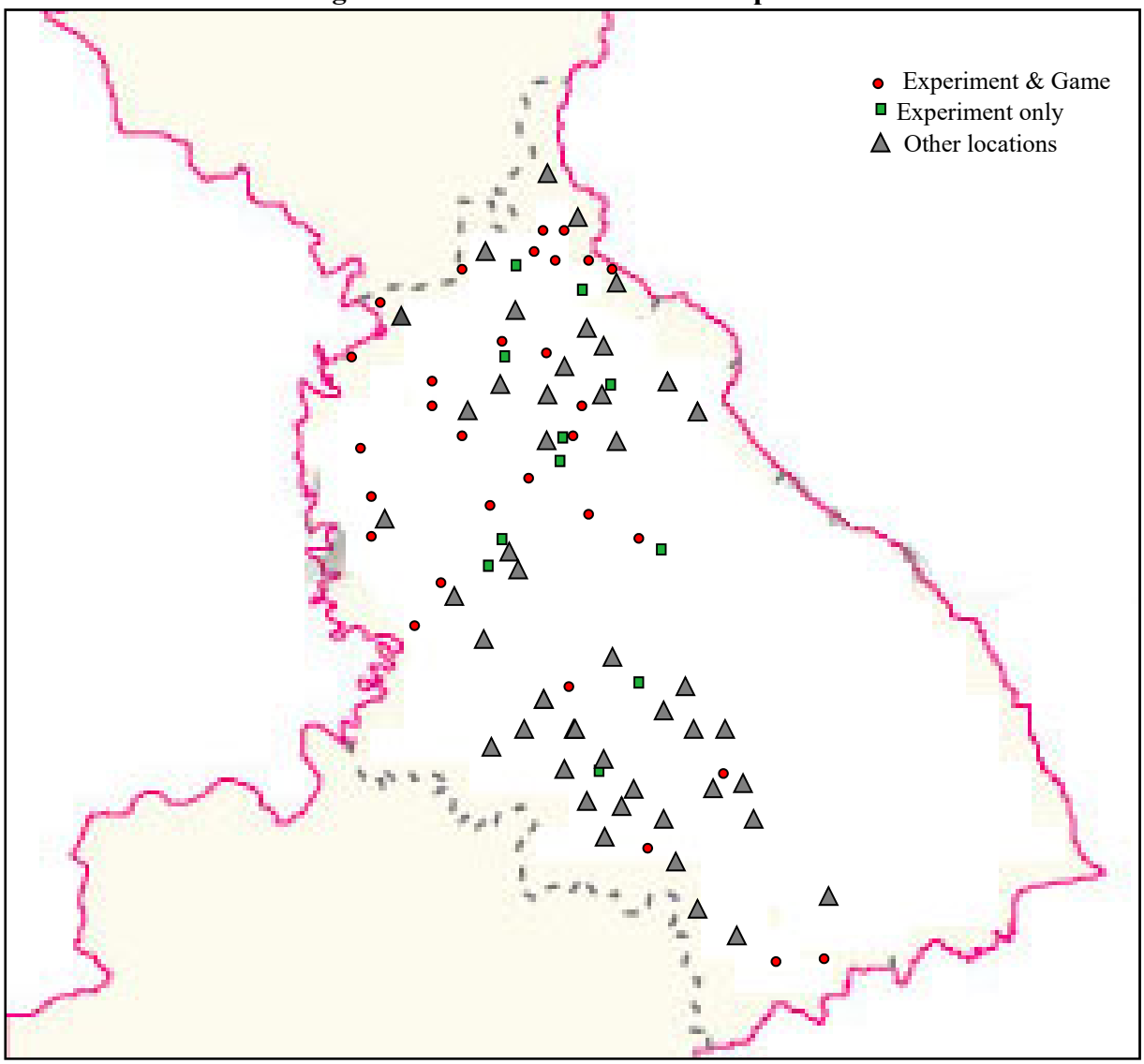
$$c_L e_L \succ c_L e_H \succ c_H e_L \succ c_H e_H$$
$$c_L e_L \succ c_H e_L \succ c_L e_H \succ c_H e_H$$

B. Locations of field experiment

Figure B1 shows the location of Kanpur Nagar district in the state of Uttar Pradesh while Figure B2 displays all the localities where the experiment was conducted. The areas covered in our study include Ratanpur, Lodhar, Kursauli, Maksudabad, Tikra, Singhpur, Hora, Paigupur, Pachor, Mandhana, Kukradev, Tikkanpurwa, Bairy, Mharajpur, Loharkheda, Pargahi, Guraha, Sandeela, Shadipur, Naurangabad, Baikunthpur, Sakshupurwa, Iswaringanj, Hradaypur, Parapratappur, Chandula, Pokharpurwa, Naramau, Karsaitpur, Madarpur, Indra Nagar, Kalyanpur Khud, Devi Shai Nagar, Sahab Nagar, Jai Prakash Nagar, Loharanbhatta, Fazalganj, Barasirohi, Mirjapur, and Maswanpur.

Figure B1: Location of Kanpur Nagar, Uttar Pradesh

Figure B2: Localities of the field experiment



C. Lab-in-the-field experiment

Table C1: Descriptive statistics

	Field experiment (n = 3,128)		Lab-in-the-field experiment (n = 482)	
	Mean	Std. Dev.	Mean	Std. Dev.
Male	0.51	0.50	0.53	0.50
Age	37.8	14.3	38.5	15.0
High caste	0.27	0.44	0.26	0.44
Hindu	0.80	0.40	0.76	0.43
College educated	0.11	0.32	0.09	0.28
Below poverty line	0.34	0.47	0.34	0.47
Urban resident	0.34	0.48	0.33	0.47

Table C2: Distribution of preference rankings consistent with different theories by caste of patient

	(1) Low-caste Patients	(2) High-caste Patients	(3) All Patients
Homophily taste-based or statistical discrimination	0.349	0.476	0.382
	(0.025)	(0.045)	(0.022)
Heterophily taste-based or statistical discrimination	0.246	0.145	0.220
	(0.023)	(0.032)	(0.019)
Uniquely statistical discrimination	0.405	0.379	0.398
	(0.026)	(0.044)	(0.022)

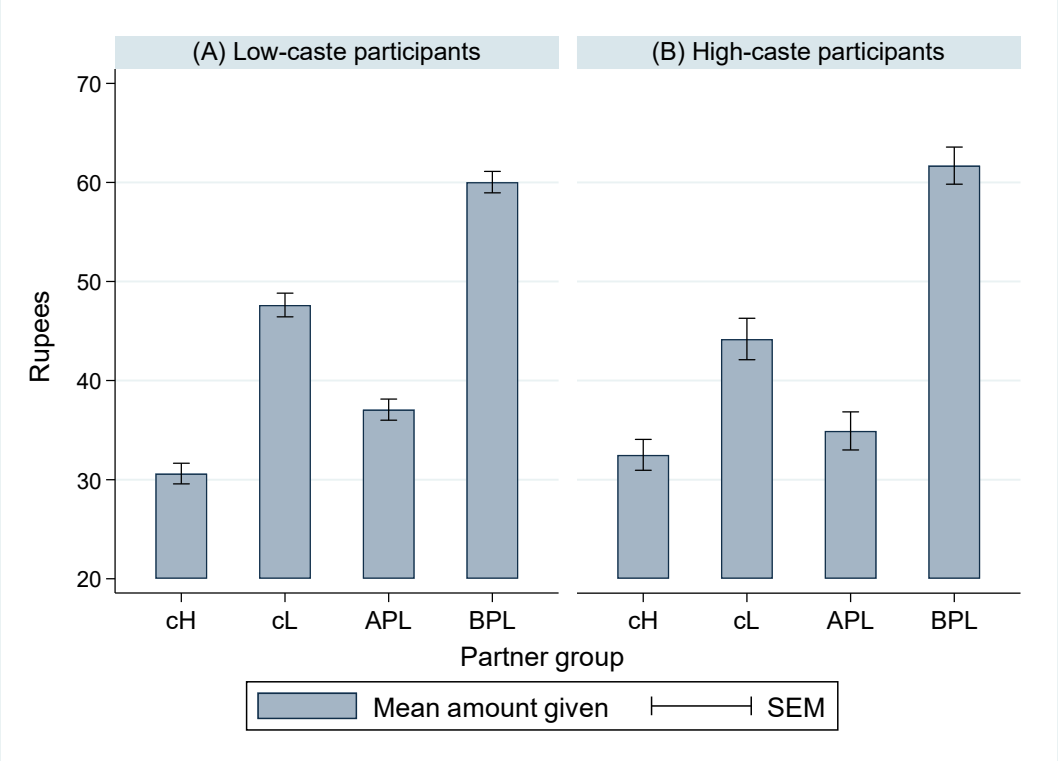
Notes: The sample size is 482. Taste-based discrimination is indistinguishable from statistical discrimination. High experience may signal better or worse quality of healthcare service when classifying with which theory a particular ranking is consistent. Standard errors reported in parentheses.

Table C3: Distribution of preference rankings consistent with different theories by caste of patient after reclassification of taste-based discrimination

	(1) Low-caste Patients	(2) High-caste Patients	(3) All Patients
Homophily taste-based or statistical discrimination	0.223	0.137	0.201
	(0.022)	(0.031)	(0.018)
Heterophily taste-based or statistical discrimination	0.020	0.121	0.046
	(0.007)	(0.029)	(0.010)
Uniquely statistical discrimination	0.757	0.742	0.753
	(0.023)	(0.039)	(0.020)

Notes: The sample size is 482. Reclassification of taste-based discrimination using List's (2004) method to define caste preference. A participant is defined as having a caste preference for one caste group when the amount given to that group is larger than the other caste group. When the caste preference defined in this way contradicts the caste preference inferred from the preference ranking of doctors in the field experiment, preference ranking consistent with taste-based discrimination is reclassified as consistent with statistical discrimination. Standard errors reported in parentheses.

Figure C1: Mean amount given to different groups by caste of patient



D. Robustness

D.1. Preferences for attributes other than caste and experience

In the taste-based discrimination model, we assume that a patient has only preferences for the social closeness of a doctor in terms of caste and the quality of health care, but nothing else. In reality, patients are likely to have preferences for other attributes of a doctor. A good example is the gender of a doctor. Now imagine that, in our experiment, patients are shown only two of the three attributes in their consumption bundle, x. Given that they also have preferences for the gender of a doctor, it is plausible to think that they may use the doctor's experience to infer the doctor's gender because female doctors are increasingly more represented in the medical profession in India (Bhadra, 2011). This possibility is similar to the Heckman's (1998) critique about using experiments to detect taste-based discrimination. Note that, to put aside the possibility of caste-based statistical discrimination, we still assume that these patients do not use the doctor's caste to help make inferences about gender. In this case, differences in the predicted gender of a doctor across patients will influence how they rank the four doctors. The possible rankings under taste-based discrimination for the case when patients try to infer the unobserved attribute from an observed attribute in $x \in \mathbb{R}^3_+$ may thus include more than the eight possible rankings we have highlighted for the case when $x \in \mathbb{R}^2_+$.

Table D1: Robustness of results to genders of patients and doctors

	(1) Male Patients Male Doctors	(2) Male Patients Female Doctors	(3) Female Patients Male Doctors	(4) Female Patients Female Doctors
A. Survey method				
Homophily taste-based or statistical discrimination	0.184	0.186	0.213	0.222
	(0.014)	(0.014)	(0.015)	(0.015)
Heterophily taste-based or statistical discrimination	0.041	0.030	0.023	0.033
	(0.007)	(0.006)	(0.005)	(0.006)
Uniquely statistical discrimination	0.775	0.784	0.764	0.745
	(0.015)	(0.015)	(0.015)	(0.016)
B. Refinement of List's (2004) method				
Homophily taste-based or statistical discrimination	0.169	0.237	0.224	0.182
	(0.032)	(0.039)	(0.041)	(0.035)
Heterophily taste-based or statistical discrimination	0.140	0.093	0.075	0.124
	(0.030)	(0.027)	(0.026)	(0.030)
Uniquely statistical discrimination	0.691	0.669	0.701	0.694
	(0.040)	(0.043)	(0.044)	(0.042)

Notes: Taste-based discrimination is indistinguishable from statistical discrimination. High experience may signal better or worse quality of health care service when classifying with which theory a particular ranking is consistent. Standard errors reported in parentheses.

If the share of (uniquely) statistical discriminators does not vary significantly when additional correlated attributes are taken into consideration, then our methods of bounding taste-based discrimination are unlikely to yield biased estimates due to taste-based discriminators using observable attributes of a doctor to infer other unobservable attributes of the doctor for which they have preferences.

We thus examine if the findings for male and female patients differ between the case when all the four doctors are females and the case when all the four doctors are males. The results are reported in Table D1. Column (1) reports the results for male patients and male doctors, column (2) reports the results for male patients and female doctors, column (3) reports the results for female patients and male doctors, and column (4) reports the results for female patients and female doctors. The share of discriminators who have preference rankings consistent with only statistical discrimination is similar regardless of whether they are presented with female or male doctors. Therefore, our results are robust to the possibility that taste-based discriminators use a doctor's experience to make inferences about other unobserved attributes of the doctor.

D.2. Patients may strategically manipulate the rankings of doctors

It is plausible that patients may strategically manipulate the rankings of doctors as in the matching literature, such as for the choice of schools.

Table D2: Robustness of results to changing the order of the two top-ranked doctors among rankings only consistent with statistical discrimination

	All patients
A. Survey method	
Homophily taste-based or statistical discrimination	0.304
	(0.016)
Heterophily taste-based or statistical discrimination	0.067
	(0.009)
Uniquely statistical discrimination	0.628
	(0.017)
B. Refinement of List's (2004) method	
Homophily taste-based or statistical discrimination	0.293
	(0.021)
Heterophily taste-based or statistical discrimination	0.205
	(0.018)
Uniquely statistical discrimination	0.502
	(0.023)

Notes: Taste-based discrimination is indistinguishable from statistical discrimination. High experience may signal better or worse quality of healthcare service when classifying with which theory a particular ranking is consistent. Standard errors reported in parentheses.

We believe that such an issue is not as relevant in our specific context. Our subjects were not making decisions in the setting with which the matching literature is typically concerned. First, our subjects were given a short amount of time to choose from a set of doctors that they never heard or visited before (doctors with fictitious names). The information given about these doctors was also very limited (surnames, experience and gender). Second, our subjects were making decisions for a one-off experience that was expected to last no more than 15 minutes. Third, subjects were not informed about the exact probability of each doctor being chosen. These features contrast the situations of school choice that the matching literature studies, where parents have rich information the schools available and the number of slots available relative to the number of applicants, have substantial amount of time to make decisions, and an understanding as to where their child will eventually end up attending the school for many years to come.

However, even if it is very unlikely, we investigate the sensitivity of our results to the possibility that patients strategically manipulate the top-two ranked doctors by putting their most preferred doctors as the second most-preferred doctors. For that, we swap the order of the two top-ranked doctors among patients who show rankings only consistent with statistical discrimination.

In column (1) in Table D2, we report the results based on the two methods of reclassifying taste-based discrimination. Quite naturally, the maximum share of taste-based discrimination increases and is now between 37% and 50%. The conclusion that the majority of patients are unambiguously statistical discriminators remains unchanged.

D.3. Patients may not care much about the lower-ranked doctors

Table D3: Robustness of results to changing the order of the two bottom-ranked doctors among rankings only consistent with statistical discrimination

	All patients
A. Survey method	
Homophily taste-based or statistical discrimination	0.324
	(0.008)
Heterophily taste-based or statistical discrimination	0.050
	(0.004)
Uniquely statistical discrimination	0.626
	(0.009)
B. Refinement of List's (2004) method	
Homophily taste-based or statistical discrimination	0.301
	(0.021)
Heterophily taste-based or statistical discrimination	0.195
	(0.018)
Uniquely statistical discrimination	0.504
	(0.023)

Notes: Taste-based discrimination is indistinguishable from statistical discrimination. High experience may signal better or worse quality of healthcare service when classifying with which theory a particular ranking is consistent. Standard errors reported in parentheses.

It is also plausible that some patients may only care about their most-preferred choices (the first-two doctors that they rank) and do not care much about their less-preferred choices (the last two doctors that they rank). In other words, they may be indifferent between lower-ranked doctors and, as a result, may randomly rank these less-preferred doctors. Because the ranking of the last two doctors affects our inference of whether the patient's ranking is consistent with statistical discrimination only or with both theories, we need to verify how these labels are affected when we change the ranking of the last two doctors. For that, we swap the two bottom-ranked doctors for rankings only consistent with statistical discrimination and turn them into rankings consistent with both theories to assess the sensitivity of our results to this type of randomization.

Table D3 reports the results based on the two methods of reclassifying taste-based discrimination. After we swap the two bottom-ranked doctors, the upper bound of taste-based discrimination varies between 37% and 50%. The results give us confidence that, at least half of the patients are unambiguously statistical discriminators.

D.4. Accounting for non-discriminators

Given the evidence of discrimination in Figure 5, we have assumed that all patients were discriminating to identify the major source of discrimination. To assess the sensitivity of our findings to the assumption that all patients discriminate, we now perform some numerical simulations where we randomly choose the share of non-discriminators and check how the other patients discriminate against doctors.

Table D4: The share of different types of discriminators

	Assume 95% discriminators	Assume 75% discriminators	Assume 50% discriminators
A. Survey method			
Homophily taste-based or statistical discrimination	0.201	0.201	0.201
	[0.002]	[0.004]	[0.007]
Heterophily taste-based or statistical discrimination	0.032	0.032	0.032
	[0.001]	[0.002]	[0.003]
Uniquely statistical discrimination	0.767	0.767	0.767
	[0.002]	[0.005]	[0.008]
B. Refinement of List's (2004) method			
Homophily taste-based or statistical discrimination	0.203	0.204	0.205
	[0.004]	[0.011]	[0.019]
Heterophily taste-based or statistical discrimination	0.109	0.110	0.110
	[0.004]	[0.008]	[0.015]
Uniquely statistical discrimination	0.688	0.686	0.686
	[0.005]	[0.013]	[0.021]

Notes: Mean share is computed on the basis of 1000 repetitions of simulation. Standard deviation of the share of rankings reported in the bracket.

We assume that a fraction of individuals in our field experiment do not discriminate doctors on the basis of caste and then bound taste-based discrimination using the two methods of reclassifying taste-based discrimination. Specifically, we select a random subset of participants according to the assumed share of discriminators in each round of simulation and perform 1,000 rounds of simulation for each assumed share of discriminators. The results are reported in Table D4. The share of individuals who are not discriminators is first assumed to be 5% in column (1), then 25% in column (2), and finally 50% in column (3). Whether we consider 50% or 95% discriminators, the upper bound of taste-based discrimination stays between 23% and 32%. Thus, our conclusion that the majority of patients are statistical discriminators is robust to allowing a fraction of patients to be non-discriminators.