Are Lawyers' Case Selection Decisions Biased? A Field Experiment on Access to Justice

Jens Frankenreiter[*] *Michael A. Livermore*[†]

Accepted for publication by the
Journal of Legal Studies on September 27, 2022[‡]

Abstract

The lawyer-client relationship is pivotal in providing access to courts. This paper presents results from a large-scale field experiment exploring how demographic information (encoded in potential clients' names) affects how attorneys respond to initial inquiries in private injury cases. Based on prior literature, we hypothesized that race would be a significant factor, but we also explore race and gender interactions. We find that ostensibly Black or Hispanic senders receive fewer responses than ostensibly White senders, a result largely driven by a preferential treatment of White female senders. The racial disparities are larger than those previously documented in some contexts, such as public services, but smaller than in others, such as employment. We also find suggestive evidence that White attorneys are more likely than others to treat White senders preferentially, implying that the differences in response rates are not merely a reaction to jurisdiction-level factors affecting the expected payoff of lawsuits.

[*]Associate Professor of Law, Washington University in St. Louis School of Law; fjens@wustl.edu.

[†]Edward F. Howrey Professor of Law, UVA School of Law, Charlottesville, Virginia; mlivermore@law.virginia.edu.

[‡]We are grateful to two anonymous referees as well as Daniel Chen, Adam Chilton, Charles Crabtree, Adrienne Davis, Christoph Engel, Jim Greiner, Chris Griffin, Daniel Harawa, Eric Helland, William Hubbard, Adi Leibovitch, Chang-Ching Lin, Julian Nyarko, and Sonja Starr for useful comments. The paper has furthermore profited from feedback by participants in the 2020 International Junior Scholars Forum in Law and Social Science, the summer workshop at University of Virginia School of Law, the 2019 Conference on Empirical Legal Studies (CELS), the 2019 Canadian Law and Economics (CLEA) Conference, the 2019 Empirical Legal Studies Conference at Academia Sinica in Taipei, Taiwan, and the internal seminar series at Max Planck Bonn. Martin Bolger, Sarah New, and Andrew Teal provided outstanding research support. The authors remain solely responsible for any omissions or errors.

1 Introduction

In the United States, access to justice often requires access to lawyers. Although, in principle, litigants can represent themselves in court, as a practical matter *pro se* litigants face a variety of challenges that make it difficult to vindicate their rights. In many criminal proceedings, the U.S. Supreme Court has found that access to counsel is so basic that the Constitution requires representation be provided to indigent defendants. The existence of a robust market for legal services also provides evidence of the belief among consumers that competent legal counsel is valuable. Where it has been studied, legal representation is oftentimes associated with better outcomes (Eagly and Shafer 2015; Greiner, Pattanayak, and Hennessy 2013), although this effect is not universal (Greiner and Pattanayak 2012).

The role and importance of counsel varies by litigation context, but access to lawyers may be particularly consequential in legal areas where fees are paid on a contingent basis. For these matters, lawyers provide not only counsel and advocacy, but also financing for litigation costs. In many civil cases, whether or not a plaintiff is able to pursue litigation at all will often depend on an attorney's decision of whether to take up a case. This paper presents results from a large-scale field experiment exploring how demographic information (encoded in potential clients' names) affects elements of attorneys intake decisions in private injury claims. In line with studies in other fields such as employment and housing, we find evidence suggesting that attorneys might disfavor members of racial/ethnic minorities in this context.

There are a variety of factors that may influence attorneys' intake decisions. A utility maximizing attorney will seek out the cases that provide the largest expected benefit at the lowest expected cost. Typically, "benefit" can be understood in terms of monetary remuneration—the likelihood of a settlement or judgment times the anticipated amount, discounted by the contigency fee percentage. There may also be reputational benefits associated with taking on certain high-profile cases, as well as non-pecuniary benefits akin to the psychological "warm glow" that accompanies charitable giving (Andreoni 1990). Costs are often, similarly, pecuniary in nature, but may also have reputational or psychological dimensions.

The decision concerning whether to take a case is sequential in nature, as prospective attorneys and clients gather and reveal information during their preliminary interactions. There are important information asymmetries on both sides, as clients have access to privately held information about their case, and attorneys have experience in pursuing similar matters. Once a matter is taken on, an attorney-client relationship attaches, with all of the associated responsibilities and liabilities on the part of the attorney. Most important, the attorney-client relationship makes it difficult to abandon a client "mid-stream" in the course of litigation, even if information is revealed that indicates that a case has a lower expected payout than initially anticipated. This lock-in effect means that information collected prior to creating an attorney-client relationship has particularly high value.

Even though intake decisions can be consequential, they nevertheless must be taken on the basis of quite limited information. For valuable claims, attorneys are in potential competition with each other, and delaying the decision to take up a case with the goal of gathering more information may result in losing the client altogether. Protracted initial deliberations, especially with unsophisticated clients, also raise the possibility of creating an informal attorney-client relationship, with the attendant duties and liabilities. In this initial stage, attorneys must carefully balance the costs and benefits of delay. Because attorneys often have to make their intake decision based on a preliminary understanding of the facts of a case, they may be influenced by guesses and hunches as much as reasoned analysis.

This study investigates the very beginning of attorney intake decisions—an initial request for information on the part of a prospective client. Such a request represents the first gate-keeping opportunity for a lawyer. By failing to respond or responding negatively to such a request, an attorney can terminate any dialogue before it begins. This initial contact also represents the lowest information point for the attorney, as a prospective client is reaching out "cold," and the attorney has access to only the information that the requestee has chosen to reveal.

We are interested in how basic demographic factors that are revealed in a prospective client's name (most importantly, a sender's membership in a racial/ethnic group) might affect attorneys' choices concerning whether to pursue additional information at this initial stage. To investigate this question, we conduct a large-scale, email-based field experiment. Our work builds on a robust social and behavioral science literature that examines how various economic and political interactions are affected by race and gender, utilizing a name-based experimental design. Leveraging the same basic methodology that has been developed in this literature (Crabtree 2018; Gaddis 2018), we sent initial information requests to more than 24,000 attorneys in jurisdictions across the United States, randomly assigning requestee names that signal different demographic characteristics. Our research question is whether the rates of responses that we receive to these requests are sensitive to these changes and what geographic and attorney characteristics predict response rates.

Our primary finding is that perceived race/ethnicity is a factor in attorney responses to initial inquiries concerning personal injury cases, with ostensibly Black or Hispanic senders less likely to receive a response than ostensibly White senders.[1] The existence of this effect

1. We note that identity markers are contested, and the discussion surrounding them is constantly evolving.

is validated by a split sample approach we employed in designing this study. We constructed the models used to test our hypotheses on the basis of an analysis of roughly half the available data. After preregistering the resulting research plan, we used the second half of the data to confirm that our results replicate.

In apparent contrast to similar studies in other fields, the preferential treatment of White senders is mostly driven by higher response rates for senders with names suggestive of persons who are White and female. While this effect can also be observed in both rounds of the study, it is markedly more pronounced in the first round. In particular, the results from the first round provided little indication that White male inquirers received more responses than male inquirers in other racial/ethnic categories. By contrast, in the second round, attorneys were more likely to respond to White male inquirers than to other male inquirers, although this difference is not statistically significant.

We furthermore find that the treatment effect described above may vary depending on the attorney's location and personal characteristics. Most importantly, we find some evidence that the preferential treatment of White senders is driven by certain subgroups of attorneys, among them attorneys with names that are more common among the White population in the United States. This result persists when we apply a matching procedure that seeks to eliminate potential confounding effects of the environments in which attorneys operate. We note that we were unable to replicate this result using the statistical test included in our preregistration statement. However, an analysis of the combined data still suggests the existence of such "in-group, out-group" effects, and we leave it to future research to confirm that these results can be replicated. If supported in future studies, this finding would suggest that the differential treatment of different groups cannot be explained solely as a response on the part of attorneys to jurisdiction-level factors that affect the expected payoff of lawsuits.

The remainder of this paper is structured as follows. Section 2 provides more details on the theory as well as an overview of the existing literature. Section 3 describes the design and execution of the field experiment, Section 4 discusses the the data and provides descriptive statistics. Section 5 presents the results of the study. Section 6 discusses the implications of these results, Section 7 concludes.

As we explain in more detail in Section 3.3 below, we use the term Black to refer to senders whose names are indicative of membership in the census category "Black or African-American" and the term Hispanic to refer to senders whose names are indicative of membership in the census category "Hispanic or Latino." We use this label, as opposed to other candidate labels such as Latino/Latina/Latinx, for members of the census category "Hispanic or Latino" because it tracks the usage in the papers from which we draw the information for the sender names in our study (Gaddis 2017a, 2017b; Crabtree and Chykina 2018).

2 Theory and Previous Literature

The primary aim of this study is to document whether lawyers in the United States respond differently to initial inquiries based on perceived demographic characteristics of prospective clients. We are particularly interested in documenting whether lawyers are less likely to respond to inquiries by members of racial or ethnic minorities. We also document differences in response rates to inquiries by senders with names suggesting different gender identities and the interaction between race/ethnicity and gender. Randomization in the experimental design allows us to investigate causal claims concerning the randomized variables, and the extent and diversity of the data allows us to test these claims for various subgroups of attorneys and in a variety of geographic settings. In other words, we can investigate whether changing a client's name *causes* differential attorney response rates, and we can estimate how those causal effects vary based on attorney characteristics and attorney geography.

By contrast, this study is not intended to untangle the precise mechanisms that lead to differences in attorney response rates. Differential access to legal remedies has important social and economic consequences irrespective of its precise sources. At the same time, a better understanding of the mechanism at play has implications for potential policy designs to counter these effects. With this reality in mind, we attempt to shed at least some light on the underlying mechanisms.

There are many behavioral processes that could affect how attorneys respond to initial inquiries. First, differences in the treatment of sender groups could be caused by hostility towards or an unwillingness to work with members of certain racial/ethnic groups (Becker 1971). For example, an attorney might enjoy the feeling of seeing her client win, but might enjoy this feeling less for clients from other racial/ethnic groups. This type of discriminatory behavior is sometimes described as "taste-based" (Moser 2012) or "animus-based" (Ayres 1991) discrimination in the economics literature. In the psychological literature, the term "prejudice" is used to describe these types of negative attitudes towards members of certain groups (Stangor 2009). From an economic perspective, an attorney engaging in such behavior sacrifices profits to satisfy her prejudices.

Second, a differential treatment of members of different racial/ethnic groups can also be the results of attorneys' attempts to maximize profits. This is the case when attorneys hold beliefs about members of different groups that influence the expected payoff of lawsuits in which they act as plaintiff. There are various types of belief that could cause the expected payoff of lawsuits to differ. Some of these beliefs are related to characteristics of the prospective client that are hidden from the attorney's view. For example, an attorney might assume that the income of a white inquirer is higher than the income of a member of a minority group, which can influence the size of the judgment. Other beliefs are related to the impact that the a client's group membership will have on the decisions of others. In the context of a lawsuit, an attorney might for example anticipate that judges and juries will treat members of different groups differently.

This second type of explanation is usually described as "statistical" discrimination by economists (Phelps 1972). Psychologists use the term "stereotyping" to describe the related idea that actors attribute to group members the traits that are typically associated with members of this group. Compared to taste-based discrimination and prejudice, the concepts of statistical discrimination and stereotyping are relatively vague. They comprise a number of different cognitive, evaluative, and psychological influences on decision making that can have different normative implications. In particular, most authors assume that statistical discrimination can be based on honestly held beliefs, whether or not these beliefs are well-informed (e.g., Ayres 1991). These authors tend to describe statistical discrimination as normatively problematic (e.g., Phelps 1972). This position also appears to prevail in the psychological literature on stereotyping (Stangor 2009). On the other hand, statistical discrimination is described as efficiency-enhancing by some, especially in the economics literature, partly based on a narrow description of statistical discrimination that excludes misinformed beliefs (e.g., Posner 1989).

In this paper, we do not attempt to distinguish between taste-based and statistical discrimination. Differentiating between the two forms of discrimination is usually elusive in the context of a field experiment like ours. Furthermore, because of the lack of clarity surrounding the normative consequences of statistical discrimination, a finding that the disparate treatment of members of different groups was caused by statistical discrimination on its own would be of limited value.

Instead, we explore patterns suggesting that attorneys are more likely to respond to inquiries by members of their own racial/ethnic group (see also Agan and Starr 2020; Block et al. 2021; Libgober 2020). While the presence of these patterns is compatible with a number of different explanations including animus against and ill-informed beliefs about members of other groups, it seems incompatible with certain other versions of statistical discrimination. In particular, such patterns appear incompatible with an explanation that attributes the differential response rates to differences in the value of representing clients from different racial/ethnic groups that are unrelated to the attorneys' personal characteristics. One example of such a factor is the expected downstream behavior by other actors (e.g., judges or juries) that influence the chances of a plaintiff to succeed in court. This question is not only interesting from an academic perspective. The types of intervention that will or will not be likely to succeed in facilitating more equal access to justice or addressing failures in the market for legal services are likely to depend, in part, on the underlying mechanisms at work.

This study relates to various previous literatures. First, it builds on the literature on so-called "audit studies" (Gaddis 2018), a stream of research that investigates whether the identity of a person influences other persons' decisions in various settings, including hiring and responses to inquiries by public officials. Our study is most similar to those in which researchers contact decision makers by mail or email sent under different names, some of which suggest membership in a minority group (e.g., Agan and Starr 2018; Bertrand and Mullainathan 2004). These studies, which are also known as "correspondence studies," appear to have been most widely applied on labor discrimination (Gaddis 2018, 9–21). Researchers have used this methodology to study decisions in various fields, including short-term market transactions (Ayres 1991; Ayres and Siegelman 1995; Cotropia, Masur, and Schwartz 2019) and interactions between public officials and citizens (Broockman 2013; Butler and Broockman 2011; Einstein and Glick 2017; White, Nathan, and Faller 2015).

The results of these studies vary depending on treatment and setting. In most instances, researchers were able to document pervasive differences in the response rates for applications or inquiries sent by members of different racial/ethnic groups across various fields in the United States (Bertrand and Duflo 2017; Block et al. 2021). This effect is strongest and least ambiguous in the hiring context. Black inquirers also receive substantially fewer responses than White inquirers in the rental housing context, while discrimination against Hispanic senders appears to be less severe in this context. To a lesser extent, both Black and Hispanic inquirers also face discrimination in accessing public services and other areas (Gaddis et al. 2021).

With regard to gender, the available evidence is more mixed. In hiring decisions, for example, gender discrimination seems to vary by occupational characteristics (Baert 2018). Gender effects seem not to have been broadly studied in the context of access to local politicians or street-level bureaucrats (e.g., Butler and Broockman 2011; White, Nathan, and Faller 2015).

Our study joins the nascent stream of literature that employs this methodology to investigate the decisions of lawyers and their role in the legal system. An early example of work in this field is Braucher, Cohen, and Lawless (2012), who use a study design that shares many features of an audit study to document race-dependent differences in the recommendations bankruptcy lawyers make to clients. Importantly, however, they do not study actual attorney behavior but merely responses to a vignette.

The closest prior work to our study is Libgober (2020), who also uses an email audit study to test for discrimination in intake behavior in various legal areas. In line with the findings in other fields, this study finds evidence that lawyers do indeed discriminate on the basis of race and that lawyers' racial preferences might be driving this effect. Our study builds on this work in several ways. First, our study is much larger in size, which allows us not only to study the effects in question with more precision, but also to analyze the heterogeneity in attorneys' responses in ways that are unavailable to Libgober. Second, Libgober includes only attorneys from two regions (California/Florida), and the differences in the results obtained in both settings raises difficult questions about the interpretation of his results. By contrast, our study is nationwide in scope. Finally, we expand the racial/ethnic classes that are studied and use a much larger combination of names, which allows us to mitigate the influence of potential effects that happen to be associated with particular names.

Our study is also related to work that documents the existence of discrimination in the legal system using methods other than audit studies (e.g., Abrams, Bertrand, and Mullainathan 2012; Alesina and La Ferrara 2014; Arnold, Dobbie, and Yang 2018; Miles 2012; Shayo and Zussman 2011; Underhill 2019; Yang 2015). Finally, our paper shares commonalities with other work that investigates the decision-making of legal actors using (field) experiments (Baradaran et al. 2013; Baradaran et al. 2014; Spamann and Klöhn 2016).

3 Research Design

3.1 Basic approach

We set up this study as a field experiment with a between-subject design.[2] We approach lawyers by email and record whether and how they respond to these inquiries. We contact each lawyer exactly once, randomly varying the name and email address of the purported inquirer in a way that suggest membership in a particular racial/ethnic group and a specific gender identity. Our main outcome variable is an indicator for whether attorneys responded to an inquiry or not.

This setup allows us to investigate how the assignment of emails and sender names (the treatment) affects the probability of a response. Drawing on the findings of most audit studies in the fields of hiring and access to local politicians and bureaucrats, we expect emails sent under names suggestive of a sender belonging to a racial/ethnic minority to receive fewer responses as compared to inquiries sent by White senders. By contrast, the available literature does not allow us to form any strong priors about the existence of a gender effect or interaction effects between race/ethnicity and gender.

2. This research was approved by the University of Virginia Institutional Review Board for Social and Behavioral Research.

Of course, sender name and email are not the only variables that influence the probability that an attorney will respond to an inquiry. It seems reasonable to assume that attorney characteristics will play a major role as well. This does however not threaten the validity of our results, for two related reasons. First, we are not interested in documenting absolute differences in response rates for different types of attorneys (e.g., attorneys located in different areas). Instead, we focus on whether attorneys differ in their likelihood to respond to an inquiry depending on the identity of the purported inquirer. Second, and more importantly, the random assignment of sender names and emails to attorneys ensures that there cannot be any systematic relationship between the purported identity of the sender and any attorney characteristic that might influence the response rate as well.

3.2 Attorney Data

As a basis for this study, we obtained lawyers' contact information from a commercial legal database. We focused on those lawyers for whom the database indicates that their practice encompasses personal injury law. Overall, we obtained information for more than 89,000 lawyers from all 50 states and the District of Columbia.[3] For slightly more than half of the attorneys in the database (over 52,000), we obtained their email addresses.[4]

Naturally, we have limited means to verify whether the information provided in the commercial database is accurate and complete, and whether the quality of the information is the same for different geographical areas. However, because of our research design, such factors (just like any other factors pertaining to the identity of an attorney) do not impair the validity of our results.

3.3 Treatment and Randomization

As discussed above, the most important feature of this study is that the treatment (i.e., the purported identity of the potential client) is randomly assigned to the attorneys we contact. In order to achieve this, we took the following steps.

First, we selected the sample lawyers that would be contacted from the population of lawyers for whom we obtained email addresses. We contacted each lawyer one time, at maximum. Attorneys were chosen randomly from our population of attorneys for which email addresses were available, subject to the constraint that only one lawyer per firm was included in the sample. This decision implies that lawyers working in bigger firms with

3. In the remainder of this paper, references to "states" include the District of Columbia.

4. Table OA3 in the Online Appendix reports information on the number of attorneys in the database whose practice includes personal injury law by state.

more than one attorney appearing in our dataset tend to be undersampled in our study. We sampled lawyers separately for each state, ensuring that the number of lawyers from each state included in our study is proportional to the number of lawyers from this state for whom we obtained email addresses.[5] Overall, we obtained a sample of 25,680 attorneys who were contacted during the course of the study.

Second, we settled on the names and email addresses of purported inquirers. U.S. government guidance describes five categories for data on race—American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White—and two categories for data on ethnicity, "Hispanic or Latino" and "Not Hispanic or Latino."[6] We generate six name types, corresponding to all possible combinations of two gender groups (male and female) and three race/ethnicity groups: Black or African-American; Hispanic or Latino; and White.

To construct the names, we relied on information in Gaddis (2017a, 2017b) which empirically validated racial/ethnic perceptions based on first names, and U.S. census data that correlated last names with racial/ethnic categories (Crabtree and Chykina 2018). We used six last names indicative of membership in the census categories Black or African American and Not Hispanic or Latino, six last names indicative of membership in the (ethnic) census category Hispanic or Latino, and six indicative of membership in the census categories White and Not Hispanic or Latino.[7] For each racial/ethnic group, we used 10 first names (5 indicative of a female sender, 5 indicative of a male sender) which we combine with the last names in the same group to form a set of 60 different combinations of first and last names per group, or 180 first-name-last-name combinations in total. Emails were sent from email addresses that contain the last name of the purported inquirer, but not the first name. We used one email address per last name, for a total of 18 email addresses.[8]

Third, we randomly assigned inquirer names (and, as a consequence, email addresses) to individual attorneys. As a result, each attorney we contacted had a 16.67% chance of

5. The only state whose lawyers are underrepresented in our study is Arkansas. We did not contact any lawyers from Arkansas during the first round of the study, because we had run a pilot study in this state which did not leave us with enough lawyers to ensure the inclusion of a proportional number of lawyers in both rounds of the study. In the second round of the study, we included a proportional number of attorneys from Arkansas.

6. Office of Management and Budget, Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity (Fed. Reg. Oct. 20, 1997)

7. Gaddis (2017a), Gaddis (2017b) and Crabtree and Chykina (2018) do not explicitly distinguish between all racial/ethnic category combinations (e.g. Black Hispanic versus Black non-Hispanic). Nonetheless, theses papers provided sufficient information to create three categories: Hispanic; White alone (non-Hispanic); Black alone (non-Hispanic). These categories are in keeping with Census practice in measuring racial and ethic diversity (Jensen et al. 2021).

8. As described in note 13 below, we had to replace 4 last names and the corresponding email addresses during the roll-out of the first round of the study due to technical issues.

receiving an email from an inquirer belonging to one of the following groups: (1) Black female, (2) Black male, (3) Hispanic female, (4) Hispanic male, (5) White female, and (6) White male.

3.4 Contacting Attorneys

After assigning inquirer names to attorneys, we prepared for the emails to be sent out. The text of the emails makes initial contact inquiring into whether an attorney handles a particular kind of case, specifically a minor torts matter. In other words, each email is a request for general information rather than a specific request to engage an attorney in a matter. Other than the name of the sender, the text of the email is the same for all senders. The subject matter line also is the same for all emails.

We used two slightly different versions of the text, whereby the second version was sent only to a subsample consisting of 2,000 randomly selected attorneys contacted during the first round of the study.[9] The first version of the text contained a description of a simple slip-and-fall case in an unnamed supermarket that resulted in a serious ankle injury, medical bills, and several months of unemployment. The second version was similar to the first one with the exception that it did not name several months of unemployment as a result of the slip-and-fall. Instead, it suggested that the sender was a retired senior citizen. Both emails, in addition to the description of the case, asked whether the attorney's firm handled cases of that sort.[10]

The last step consisted of sending out the emails, and recording responses. The study was implemented using the mail merging tool YAMM.[11] Email addresses were sorted in groups of three (one from each racial/ethnic group) which were identical with the exception of the last name of the purposed inquirer. All emails sent on each given day of the study were sent from one group of three email addresses, with roughly different numbers of emails for each of the three addresses. Emails were sent from all three email addresses in close succession, and the order of the email addresses was changed each time emails from a certain group were sent, ensuring that the results were not influenced by the time and date attorneys were contacted.

9. See below Section 3.6.
10. See Section OA1 in the Online Appendix for the text of the emails.
11. https://yet-another-mail-merge.com/.

3.5 Coding of Attorneys' Responses

Of course, recording attorney's responses is not trivial. Unlike many lab experiments in which participants are required to perform a certain action that is directly quantifiable, the responses we receive come in the form of text. For the purpose of our main analysis, we register whether an attorney replied to a message. In other words, for each message we receive, we use an algorithm to determine whether a response is an error message or not. If the message is not identified as an error message, it is counted as a response by the attorney. This approach potentially disregards useful information in favor of ensuring the replicability of the results. Any attempt to determine whether a response is positive or not necessarily entails at least some level of subjectivity.

In an alternative version of the analysis, we use an outcome measure that evaluates whether a response indicated an attorney's willingness to help the inquirer obtain legal representation. We obtain this measure by hand-coding a randomly selected sample of 1,003 emails and using supervised machine learning to predict the nature of the remaining emails. We report details on this alternative measure and corresponding results in Section OA2 in the Online Appendix.

3.6 Multiple Rounds and Holdout Data

In this study, we explore heterogeneous treatment effects along multiple dimensions, implicitly testing a large number of hypotheses. Testing a large number of models on a finite data set creates risks of generating spurious results. We implement a so-called split-sample approach to ensure that results are valid and do not arise from overfitting model parameters to the data (Egami et al. 2018; Fafchamps and Labonne 2017).

With the increasing use of high-dimensional datasets in social science research, considerable efforts in data transformation and variable selection is often necessary to translate research questions into testable hypothesis (Grimmer and Stewart 2013). Although, in some contexts, theory can provide a unique mathematical model of a phenomenon under study, frequently there are many different models that can be used. The exploration of a large number of variables of interest also leads to the implicit testing of many models, which interferes with the ability to draw valid inferences from statistical constructs such as p-values. Flexibility in research design and data analysis raises problems of overfitting and the out-of-sample validity of statistical results.

Over time, various solutions to this problem have been proposed. In addition to split-sample approaches, such proposals include statistical corrections for multiple hypothesis testing and the submission of pre-analysis plans (see Fafchamps and Labonne 2017). However, these other solutions are often impracticable in studies that involve high-dimensional and/or unstructured data. Pre-analysis plans, for example, require researchers to work out in detail the operationalization of their research hypotheses before laying hands on the data. This feature imposes severe restrictions on the use of computational methods such as machine learning, which need to be calibrated on the data in order to attain their desired characteristics.

In recent years, the use of split samples has become particularly attractive, especially as the cost of data acquisition has declined. Under a split-sample method, researchers divide data into a training dataset and a testing dataset. The training dataset is used for exploration and to refine the statistical model and analyses. The testing data are used to replicate and validate the models constructed on the testing data. No changes to the specifications are allowed once the testing dataset is available to researchers. This approach is particularly popular in studies in which a sufficiently high number of observations is available to alleviate concerns about negative effects on the statistical power of tests.

We implement a split-sample validation method similar to that in Kleinberg et al. (2018). We constructed our two datasets during two experimental rounds that used roughly half of the email addresses at our disposal in each round. In the first round of the study (which ran between early June 2019 and mid-July 2019), we contacted 12,786 lawyers from all states but one.[12] Due to technical issues, we discarded 1,469 inquiries (of which 669 contained the alternative version of the email text), leaving us with 11,317 observations.[13] In a second round (running between mid-August 2019 and late September 2019), we contacted an additional set of 12,894 lawyers from all states. In this round of the study, no data was discarded. The first round data was used to construct and test our statistical models, and the second round is the held out data which we used to confirm our results. After completing the analysis of the first round data, we preregistered the study with aspredicted.org, a preregistration platform maintained by the Wharton Credibility Lab at the University of Pennsylvania.[14]

12. See above Section 3.3.

13. A number of times, our outreach seems to have triggered spam filters. In such a case, we discarded all inquiries sent out on that day, ensuring that these incidents did not overly affect inquiries by senders from one racial/ethnic group. Overall, we discarded 475 inquiries by Black senders, 499 inquiries by Hispanic senders, and 495 inquiries by White senders (a χ^2 test yields a p-value of .7134). In some cases, we also discontinued the use of the email address that triggered the filter.

14. The preregistration statement can be found in Section OA8 in the Online Appendix.

3.7 Limitations

While our study design allows us to measure differences in response rates for a nationwide sample of attorneys, it comes with at least two major weaknesses, which it shares with many other correspondence studies (see Bertrand and Duflo 2017).

One important limitation of our study is that it uses only a crude measure of legal representation as the main outcome variable. Our study investigates the initial point of interaction between attorneys and prospective clients, a setting that does not allow us to obtain direct evidence on differences in legal representation. Notwithstanding these limitations, the findings presented in this study still provide valuable insights about existing disparities in access to justice. Unless attorneys are more likely to pursue representation of Black or Hispanic clients once initial contact has been established, the differences documented here will translate into differences in legal representation. To the extent that attorneys' decisions at later points of the decision-making process disfavor members of minorities in a similar way to what we document here, these differences will be even more pronounced than the results presented in this paper. Accordingly, it seems plausible to interpret our results as a lower bound of the true extent of discrimination that members of racial/ethnic minorities face in obtaining legal representation.

Second, while we choose the names used for purported senders to signal membership in a certain racial/ethnic and gender groups, we cannot exclude the possibility that these names send other signals as well. For example, the names we selected could be perceived as related to socio-economic status, with the differential treatment we observe resulting from a reaction to differences in socio-economic status of senders rather than race or gender.

4 Data and Descriptive Statistics

Our datasets treat each inquiry as a separate observation; in total, they contain 24,211 such inquiries (11,317 from the first round, 12,894 from the second round). During the first round, 3,787 inquiries were sent using a Black name, 3,763 using a Hispanic name, and 3,767 using a White name. We received responses for 2,757 of these inquiries, corresponding to a response rate of 24.4%. In addition, for 3,314 emails, we received error messages indicating that the emails could not be delivered to the recipient. During the second round, 4,298 inquiries were sent using a name from each group, triggering a total of 3,136 responses (which corresponds to a response rate of 24.3%). In the second round, we received error messages for 3,452 inquiries.[15]

15. Almost all these error messages seemed to be due to faulty email addresses. As should be expected, such error messages were distributed more or less equally between different groups. In the first round, Black

In addition to email addresses used in the inquiries, we obtained a number of personal characteristics of the attorneys in our dataset from the commercial database we used to construct our sample. This information included a lawyer's name, the name of her firm, the address of the lawyer's offices, the number of attorneys working in the same firm and at the same office (both in ranges), and the year that the attorney was admitted to practice in different jurisdictions.

We used the attorneys' names (in the case of race/ethnicity, together with information on the racial/ethnic composition in the area in which an attorney's office is located) to obtain probability estimates for their race/ethnicity and gender. From this data, we create indicator variables for attorney gender and attorney race/ethnicity.[16] The names of the attorneys in our dataset suggest that an overwhelming majority of them is White: For 87.7% of the attorneys, our estimates indicate that there is a chance of >50% that an attorney belongs to this group, and more than 80% surpass the somewhat higher threshold of 65% that we use to construct ATT_WHITE. The names also suggest that an almost equally overwhelming majority of them is male: 85.2% of attorneys' first names are more commonly used by male individuals.[17]

We linked our data to NHGIS data reporting the demographic characteristics of all census blocks in the U.S. (Manson et al. 2018), giving us information on the racial/ethnic composition of the population in an area as well as basic economic information such as the unemployment rate and the median household income. We also obtained county-level information from Manson et al. (2019). Finally, we linked our data to information about the percentage of voters who voted for Donald Trump in the 2016 presidential election in the counties in which our attorneys' offices are located in MIT Election Data and Science Lab (2018), and obtained a measures for the ideological leanings of individuals ("CFscores") in the Database on Ideology, Money in Politics, and Elections (DIME) (Bonica 2016).[18]

senders received 1,128 error messages, followed by White senders with 1,096 and Hispanic inquirers with 1,090 (a χ^2 test yields a p-value of .7013). In the second round, Hispanic senders received the highest number of error messages (1,161) followed by Black (1,156) and White inquirers (1,135) (p-value: .798). Figure OA11 panel (b) in the Online Appendix report how the rates of responses and blocked emails varies by state.

16. We set the indicator variable ATT_MALE to be 1 if the estimated probability was greater than 0.5; otherwise we set it to be 0. We set ATT_WHITE to be 1 if the estimated probability was greater than .65; otherwise the value was set at 0.

17. Note that our attorney sample does not constitute a random sample of all attorneys listed in the database. There are two main reasons for this. First, we contact only attorneys who indicate that their practice area includes the representation of plaintiffs in private injury lawsuits. Second, as explained in Section 3.3 above, we oversample attorneys in small firms and solo practitioners in order to avoid contacting multiple attorneys from the same firm. Figures OA4 and OA5 in the Online Appendix detail how the characteristics of attorneys in our sample and the areas in which their offices are located differ from (a) the full set of private injury attorneys (whose email is listed in the database) and (b) a random selection of all attorneys with email addresses listed in the database from which we obtained our information.

18. CFscores capture the political preferences of individuals and organizations in a one-dimensional space

Table OA8 in the Online Appendix reports summary statistics.

5 Results

5.1 The Impact of Sender Identity on Response Rates

We begin by reporting results for the question of whether attorneys in the aggregate treat senders differently depending on their name. As explained above, we are primarily interested in differences between response rates of senders belonging to different racial/ethnic groups.

In line with the hypothesis that lawyers are less likely to respond to inquiries by members of racial/ethnic minorities, Black and Hispanic inquirers received fewer responses than White senders. This result appears in both rounds of the study, and it is independent of whether we choose as the base line all inquiries or only inquiries that did not yield an error message.[19] The bar plots in Figure 1 show the response rates for different sender groups, with error bars indicating 95% confidence intervals. The upper panels shows results for the first round data, the lower panels results for the second round data. Panels on the left use all inquiries as the base line, panels on the right all inquiries not flagged as error messages.

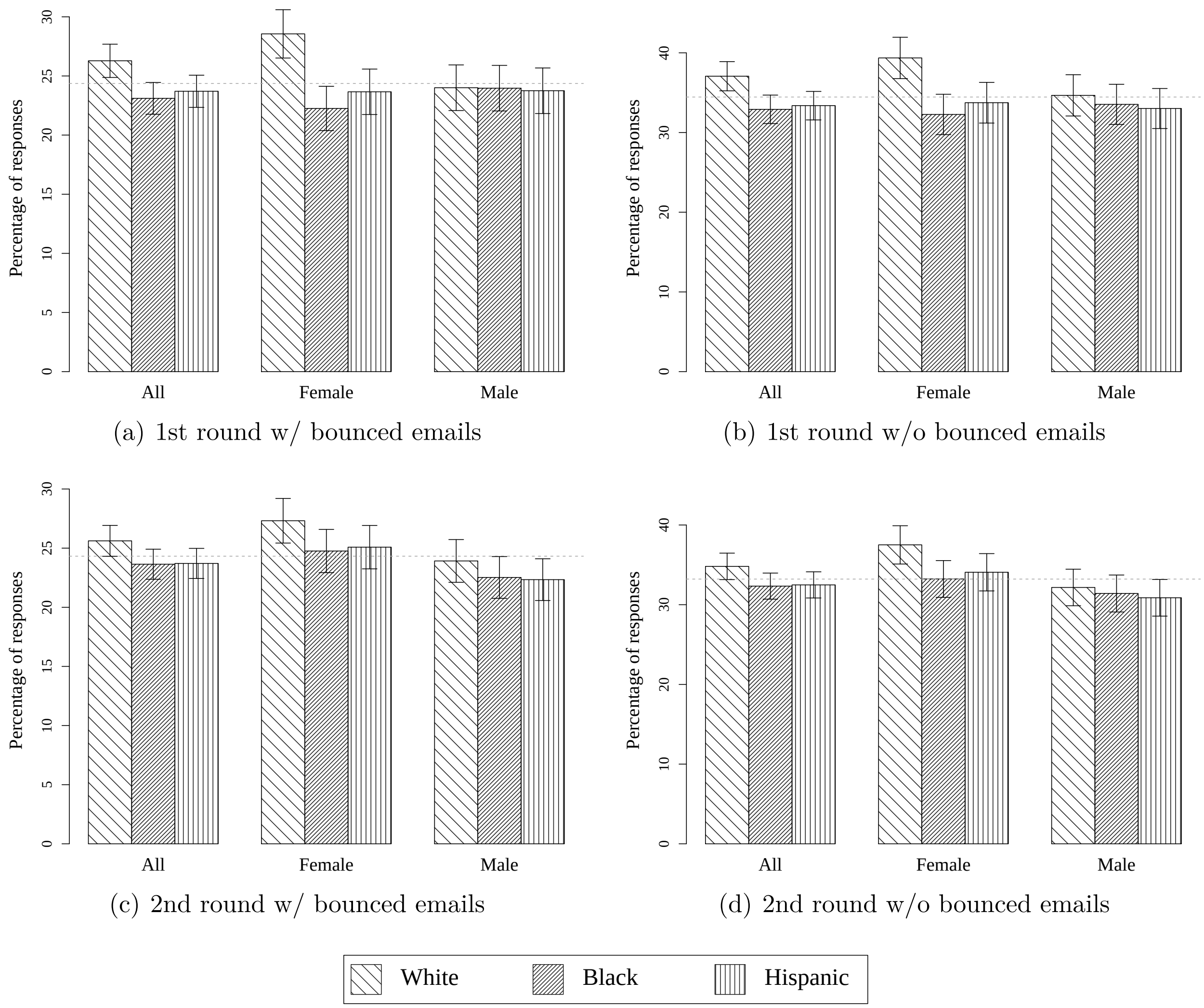
In absolute numbers, during the first round of the study, Black senders received 875 responses, Hispanic senders 892, and White senders 990. During the second round of the study, Black senders received 1,016 responses, Hispanic senders 1,019, and White senders 1,101. Overall, using all inquiries as the baseline, these numbers translate into a 23.4% response rate for Black senders, a 23.7% response rate for Hispanic senders, and a 25.9% response rate for White senders. These results indicate that Black senders are ∼11% less likely to receive a response to an inquiry than White senders. For Hispanic senders, this number is ∼9%.

The difference between the response rates for White inquirers on the one hand and Black and Hispanic inquirers on the other hand is statistically significant at the 1% level (a Fisher's Exact Test yields a p-value of .0008) in the first round and at the 5% level in the second

ranging from very liberal (highly negative) to very conservative (highly positive values). To obtain CFscores, we replicated the steps outlined in Bonica, Chilton, and Sen (2016) and Bonica and Sen (2017) to create a probabilistic algorithm that matches attorneys in our dataset to individuals in DIME based on their name, occupation, employer, and location. As DIME contains CFscores only for individuals who contributed to political campaigns, and as not all individuals contribute, we were unable to obtain ideology scores for a substantial shares of attorneys in our dataset. The final version of our linkage algorithm identifies matching observations in DIME for 45.3% of the attorneys in our dataset. This is comparably close to Bonica, Chilton, and Sen (2016) and Bonica and Sen (2017), who were able to identify matching observations in DIME for 43.4% of the attorneys in the Martindale-Hubbell database.

19. See Section 3.5 above. Technically, one can view the analysis of data that includes bounced emails as an intention-to-treat analysis and the analysis of data that excludes these observations as an as-treated analysis.

Figure 1: Response rates by gender and race/ethnicity group



(a) 1st round w/ bounced emails

(b) 1st round w/o bounced emails

(c) 2nd round w/ bounced emails

(d) 2nd round w/o bounced emails

round (p-value: .0157).[20] This result persists when bounced emails are excluded from the analysis (p-values: .0005 and .0206).[21]

To illustrate the substantive size of the effects, we calculate the number of inquiries senders from different groups need to send to obtain a certain number of responses. For example, our results imply that White senders need an average of 11.6 inquiries to receive three replies. For Black inquirers, this number is 12.8, for Hispanic senders, 12.7.

We also benchmark our results to the findings from similar studies in other fields. For

20. Here and in the following, we report conventional p-values. Our split sample design mitigates concerns that the models that we registered and tested in the second round were over-fit to the data. The p-values should be interpreted in light of the fact that each of the models were tested independently.

21. We obtain largely similar results when comparing White senders with senders from the two other groups separately.

this, we calculate response discrimination ratios (Gaddis et al. 2021) as the response rate for White senders divided by the response rates for Black and Hispanic senders. Combining the data from both rounds of the study, we obtain a discrimination ratio of 1.11 for Black senders and a discrimination ratio of 1.09 for Hispanic senders.

When compared with the results from studies in the hiring, housing, public services, education, and medical fields, our results fall right into the middle. As reported by a recent meta-study (Gaddis et al. 2021), Black senders face comparably high levels of discrimination (substantially higher than the ones we report here) in the hiring and rental housing contexts (discrimination ratios: 1.24-1.29 and 1.13-1.16). Hispanic senders also face substantial discrimination in the hiring context (discrimination ratio: 1.25-1.31), while discrimination against these senders in the rental housing context appears less severe (discrimination ratio: 1.04-1.07). Discrimination ratios for both Black and Hispanic senders in the context of public services, education, and healthcare all hover below 1.05, substantially lower than the estimates we obtain here. Our results are comparable in size to those found in a recent study examining differential response rates to Black and White senders within the general population (Block et al. 2021). In sum, these comparisons suggest that discrimination is more prevalent in attorneys' case selection decisions than it is in some professional contexts, but not as severe as in others.

Our study design also allows us to examine whether response rates varied depending on senders' gender identities. In both rounds of the study, female senders received more responses than male senders. In the second round of the study, this difference was substantial (1,658 for female as compared to 1,478 responses for male senders). However, because the difference between both groups in the first round of the study was rather small (1,407 for female as compared to 1,350 responses for male senders), we did not include gender effects among the hypotheses we planned to test using our second round data. Accordingly, while the data provides some support for the existence of a gender effect favoring females, we cannot corroborate the existence of this effect within the framework of this study.

Finally, we also investigate the relationship between the intersection of race/ethnicity and gender on the one hand and response rates on the other hand. Much of the differences between response rates for different racial/ethnic groups are a function of a preferential treatment of White female inquirers, a result that is notably absent from correspondence studies in other areas. White female senders receive by far the highest number of responses in both rounds of the study. The difference between responses for White female inquirers and all Black and Hispanic inquirers is highly significant in both rounds of the study (p-values: $< .0001$ and $.0005$), while the difference for White male inquirers is insignificant (p-values:

.5847 and .8206).[22] Note that this result is not driven by individual names: 4 out of 5 first names used for White female senders show higher response rates than all first names used for any other group.

All of the main results described in this section are robust to using an alternative outcome measure that captures whether an attorney responded to an inquiry with an email showing a willingness to help the inquirer obtain legal representation. Accordingly, we can rule out the possibility that the observed differences in response rates for different groups are due to certain groups receiving more responses declining legal representation than others. Detailed information on how we constructed the alternative outcome measure and more information on the results from this analysis can be found in Section OA2 in the Online Appendix.

Furthermore, all results are confirmed by regression analysis, the results of which we report in Section OA4 in the Online Appendix.

5.2 Treatment Effects and Attorney Characteristics

The extent and diversity of our data allows us to investigate whether and how the effects documented above vary based on attorney characteristics and attorney geography.

In a first step, we explore how the treatment of different senders varies between attorneys located in different areas. We find that a preferential treatment of White senders can be documented in a majority of U.S. states. At the same time, we find evidence of considerable geographic heterogeneity, but not necessarily one that tracks intuitive regional categories. We also investigate whether differences in the response rates for different groups vary systematically with geographic factors at the sub-state level, for example the economic well-being, racial composition, and political leanings of residents in particular areas. However, this analysis fails to produce evidence suggesting the existence of any such relationships. Detailed information on this analysis can be found in Section OA5 in the Online Appendix.

The remainder of this section presents the results of an investigation of how the treatment of different sender types varies between attorneys with different personal characteristics. In conducting this analysis, we attempt to shed at least some light on the underlying mechanisms responsible for the differential treatment of senders in different racial/ethnic groups.

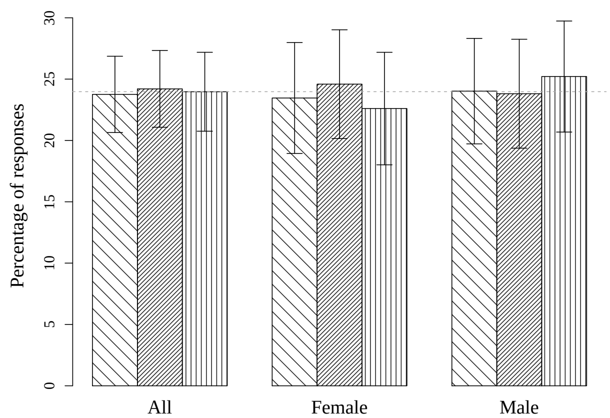
5.2.1 Descriptive Results

We start by reporting how the treatment of different groups varies by the attorneys' own predicted race. Figure 2 shows response rates for different groups of inquiries depending on

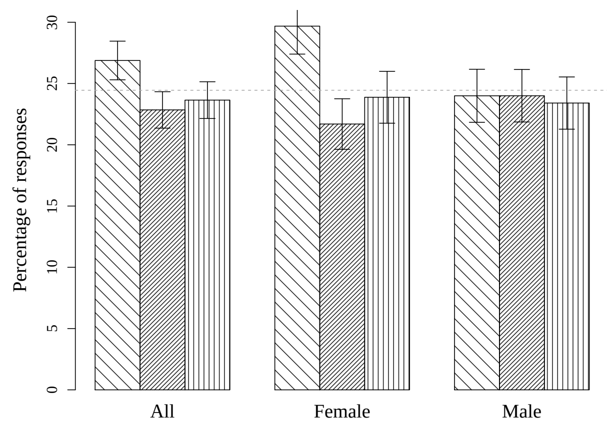
22. We obtain a similar result when comparing response rates for White female and White male senders to response rates for other senders in the same gender category.

the value of ATT_WHITE, our indicator variable for attorney race. The Figure furthermore differentiates between the data gathered in the two different rounds of the study. The upper panels show results for the first round data, the lower panels results for the second round data. Panels on the left show the responses to inquiries sent to non-White attorneys, panels on the right those to inquiries sent to White attorneys.

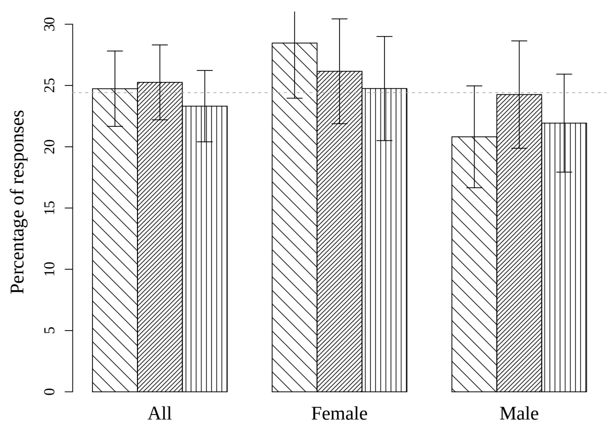
Figure 2: Differences between treatment effects by predicted attorney race



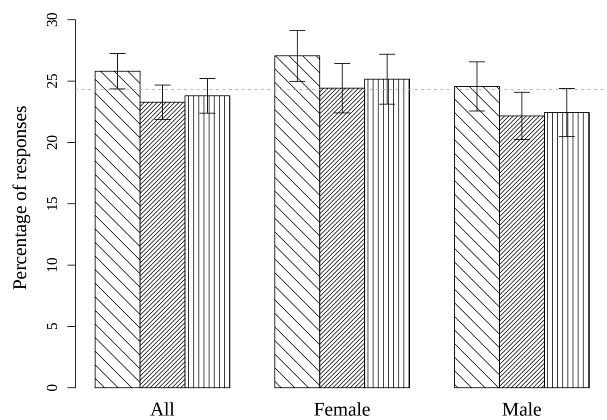
(a) Non-White attorneys – 1st round



(b) White attorneys – 1st round



(c) Non-White attorneys – 2nd round



(d) White attorneys – 2nd round

While White attorneys are less likely to respond to inquiries by Black and Hispanic senders than their non-White counterparts, they respond more often to White inquirers than the latter attorneys. On average, the chances of a White inquirer to receive a response is about two percentage points higher when the recipient is a White attorney than otherwise. This result holds for both rounds of the study.

However, there are some important differences between the data gathered in both rounds. In the first round, White attorneys treated White female senders particularly preferential.

In the second round, it was White male senders that were treated better by White attorneys as compared to their non-White counterparts.

A similar preferential treatment of White inquirers can be observed for other groups of attorneys, including male attorneys, attorneys who are solo practitioners, and older attorneys.[23] By contrast, measures of attorney ideology do not exhibit a systematic relationship with the treatment effect.[24]

5.2.2 Isolating the Effect of Attorney Characteristics

On its own, the results that White senders have a comparably higher probability of receiving a response when interacting with White attorneys does not demonstrate a preferential treatment of members of the attorneys' own groups. This is because the objective value of representing clients from different racial/ethnic groups as well as the names of attorneys might vary in different parts of the country. Maybe personal injury lawsuits brought by White plaintiffs are treated favorably by the legal systems in some states, but not in others. If our sample of attorneys from the first group of states contains a higher share of White attorneys than the sample of attorneys from other states, rational anticipation of the chances of success of lawsuits in different states could yield results like the ones described above. Technically, this problem can be described as a variation of omitted variable bias.

In order to evaluate whether the preferential treatment of White senders is more common among White attorneys, one therefore has to compare the responses of White attorneys with those of non-White attorneys from the same geographical area. In other words, what is needed is a way to hold geography constant.

In this study, we explore two different strategies to achieve this goal. First, we run a series of regressions that include, alongside the variables mentioned above, a range of different variables capturing socio-economic characteristics of the area in which an attorney is based as control variables.

However, this strategy (controlling for potential confoundering variables by way of including them as additional right-hand variables in a regression) comes with significant downsides. One is that it can only account for observable variables. Therefore, in the present context, it is impossible to rule out that unobserved geographical variables are responsible for any observed differences between attorneys from different racial/ethnic groups. Another problem is

23. Figures OA12—OA14 in the Online Appendix display information on these variables and their interaction with the treatment effect.

24. While the results from the second round of the study might suggest that attorneys whose campaign donations indicate more conservative political leanings treat White inquirers comparably better than their liberal counterparts, a similar effect could not be observed in the first round data. In that data, if anything, the preferential treatment of White female senders appears stronger for attorneys whose campaign donations indicate more liberal political leanings. For more information, see Figure OA15 in the Online Appendix.

that multiple regressions rely on linear extrapolations to obtain estimates of the effect of the variables of interest. This problem is particularly relevant if the distribution of covariates for "treated" and "untreated" observations differs substantially (Imbens and Rubin 2015, Ch. 12).

For these reasons, we also carry out an analysis that uses matching on geography to isolate the effects of personal attorney characteristics from potentially confounding effects of geography. In this analysis, we construct a version of the data that consists of pairs of attorneys (one with $ATT\_WHITE = 0$, the other with $ATT\_WHITE = 1$) who have their offices in the same area. We then test whether the effect in Figure 2 Panel (a) persists in this version of the dataset. Under the assumption that the objective value of lawsuits brought by members of different groups is the same for attorneys based in the same area, any remaining effect can be attributed to differences in the response behavior of different attorneys. This does not mean that it would be possible to infer that the different response rates are caused by the membership of an attorney in a racial/ethnic group. However, as our goal is only to exclude explanations that attribute differential response rates to factors that exist independently from the identity of attorneys participating in our study (and not to find evidence in favor of a particular explanation), we do not need to draw such an inference.

To understand what our matching procedure achieves, consider the following. For each pair of attorneys, because they are based in the same area, the expected, attorney-independent value of representing clients from different groups must be the same. Any observed difference in their behavior that cannot be explained as a result of chance therefore has to be related to the attorneys' personal characteristics (either the attorneys' group membership itself or another factor which is correlated with group membership). For the dataset as a whole, this approach eradicates any correlation that might exist in the data between (observed and unobserved) geographical variables and attorney name (see also Imbens and Rubin 2015, Ch. 18), which addresses concerns about omitted variable bias related to unobserved geographical variables.[25]

To construct the matched dataset, we implement the following steps for all attorneys with $ATT\_WHITE = 0$: First, we identify all attorneys with $ATT\_WHITE = 1$ and offices in the same county as the first attorney that had not already been matched. If there is more than one such candidate attorney in the dataset, we select the attorney whose office

25. This approach also ensures that geographic covariates for the treatment and control groups are more balanced than in the original dataset, implying that only observations from areas that can provide some information about the comparison between different types of attorneys are considered in the analysis. More precisely, under the assumption that relevant geographic variables are identical for attorneys in the same area, the matching strategy ensures that the distributions of (geographic) covariates for the different treatment and control groups (ATT_WHITE interacted with the type of sender) are identical in expectation.

is closest to the first attorney. If there is no such candidate in the dataset, we obtain a second list of candidates by identifying all attorneys that have their offices in the same state and within 10 miles from the first attorney. In case of multiple candidates, we again match based on physical proximity. If this second round does not yield a candidate, the observation is removed from the analysis. Overall, we are able to obtain matches for 4,444 of the 4,482 attorneys with $ATT\_WHITE = 0$. Of these 4,444 pairs, 4,297 consist of attorneys who have their offices in the same county.[26]

5.2.3 Original Hypothesis

As Figure 2 indicates, our analysis of the first round data suggested that the preferential treatment of White senders by White attorneys profited mostly senders with a name common among White females. White female senders constitute the only group that enjoyed higher response rates among White attorneys. By contrast, Black female (and, to a lesser extent, Hispanic male) senders had substantially lower response rates among White attorneys. The other sender groups (including White male senders) obtained roughly similar rates of responses from attorneys in both groups. Importantly, these results could also be observed in the matched data, which we interpreted as evidence that they were not due to spurious correlations with geographical variables. In addition, the results from regressions indicated that these results were statistically significant. This result persisted when we included a range of different variables capturing socio-economic characteristics of the area in which an attorney is based as control variables.[27]

Based on these results, we included in our preregistration the hypothesis that White attorneys are more likely to treat White female senders more favorably compared to other senders than non-White attorneys. However, as we document above, in the second round of the study, it was White male senders (and not White female ones) who were treated better by White attorneys as compared to their non-White counterparts. Accordingly, we were unable to confirm the original hypothesis with our second round data.

5.2.4 Further Exploration

While we were unable to replicate the results from the first round data using the second round data, this result does not offer conclusive evidence against the hypothesis that White

26. Figure OA16 in the Online Appendix shows how the covariate distributions of both groups attorneys change as a result of the matching procedure. While, for some variables, substantial differences can be observed pre-matching, the matching procedure results in subpopulations of attorneys with almost identical (geographical) covariate distributions.

27. Regression results are reported in Table OA9 in the Online Appendix.

(female) senders are treated preferentially by White attorneys. Because there were only a small number of attorneys who were not classified as White in our sample, the power of our test is limited. This implies that we cannot rule out the possibility that we would have been able to confirm our initial result if we had had more data available.

Furthermore, while we were unable to confirm our initial hypothesis that White female senders are treated better by White attorneys, all White senders taken together did receive a preferential treatment by White attorneys in the second round data as well. White senders are the only racial/ethnic group who show a higher response rate for inquiries sent to White as compared to non-White attorneys. All other racial/ethnic groups, by comparison, receive relatively more responses from non-White attorneys than they do from their White counterparts.

Against this background, we use the combined data from both rounds of the study in an attempt to generate a best guess as to the existence of a preferential treatment of White senders by White attorneys. In our analysis, which we report in Section OA6 in the Online Appendix, we find some support for the hypothesis that the observed differential treatment of senders is at least partly driven by a tendency of White attorneys to respond preferentially to inquiries by members of their own racial/ethnic group. These results suggest that the differential treatment cannot be explained by objective differences in the expected value of lawsuits brought by different client types. We stress that the results from this analysis were not corroborated by our split sample design, and we leave it to future research to determine whether these results can be replicated.[28]

6 Discussion

Providing a forum for the nonviolent resolution of disputes between private parties is a core government function, and access to such a forum is a basic civil right. In the United States, both public and private actors play a role in determining which disputes are translated into legal claims that are ultimately heard (Felstiner, Abel, and Sarat 1980). Courts affect legal claims through procedural and substantive rules as well as through their treatment of individual matters. A variety of social processes might affect the types of disputes that are litigated, but there are many reasons to believe that the relationship between clients and lawyers is pivotal. In particular, the ability of a potential litigant to find adequate legal

28. In unreported regressions, we also test whether signals about the within race socioeconomic status of senders as encoded in their first names (see Gaddis 2017a; Gaddis 2017b) are correlated with response rates. These tests fail to detect a systematic relationship between two variables, which suggests that objective differences in the expected value of lawsuits involving different clients might be of limited importance in attorney's decisions.

representation is an important step in the litigation process, and for many cases may amount to a *sine qua non*. Often, if no lawyer can be found, a claim simply cannot go forward.

This study examines the first point of contact between potential litigants and lawyers to examine whether certain demographic factors—as encoded in names—affect the likelihood that an attorney will pursue a relationship. In line with the findings of analogous studies in fields such as employment discrimination and access to public officials, the findings discussed above indicate that members of racial/ethnic minorities have a harder time establishing an initial contact with an attorney than Whites. While this effect is smaller than the effects documented in contexts such as hiring decisions, it is notably bigger than the discrimination that senders from minorities face in access to public services, education, and healthcare.

At the same time, the results from our study also differ in important ways from previous findings in other fields. Most importantly, we find evidence that the differences between response rates for senders from different racial/ethnic groups are mostly driven by a preferential treatment of senders who are White and female, a result that is notably absent from many correspondence studies in other fields. This result, which we had not anticipated in designing the study, points to the possibility that the disparate treatment we document is the result of other factors than the ones responsible for the discrimination observed in other areas.

Although the setup of this study does not provide an opportunity to explain why the preferential treatment of White senders is concentrated in White female senders, this result points to possible inter-sectional race and gender dynamics, a factor that has so far been largely ignored in correspondence studies. In other words, our results could indicate that decision-makers such as attorneys react to the race/ethnicity of others differently depending on the gender of the person they are interacting with. For example, some attorneys in our study might be less sensitive to the race/ethnicity of their prospective male clients, while they accord more weight to these factors in deciding whether to work with female clients. Or, some of the attorneys in our sample (the majority of which are White and male) may receive certain non-monetary benefits from interacting with senders who are White and female but not others.

Insofar as correspondence studies in other fields have ignored this dimension, this omission might be considered a shortcoming that researchers should address in future research.

In addition to our finding regarding differences in response rates for members of different racial/ethnic groups, we also find tentative evidence of in-group/out-group bias: the effects we document are heterogeneous based on the personal characteristics of the attorney, even when comparing attorneys based in the same geographical area. The latter finding might suggest that the observed differences in the treatment of different groups are not merely a reaction to anticipated differences in the expected payoff from lawsuits involving different plaintiffs.

Our findings are subject to several important limitations. First, while we treat our (tentative) finding that the preferential treatment of White senders is more pronounced among White attorneys as evidence against certain forms of "rational" statistical discrimination, other forms of rational behavior seem compatible with these patterns. Most importantly, the objective value of lawsuits involving clients with specific characteristics might vary between attorneys, causing some attorneys to prefer cooperation with specific clients even though other attorneys do not share the same preferences. This could be true, for example, if attorneys received higher reputational benefits from successfully suing on behalf of members of their own racial/ethnic group, or (conversely) if society in certain areas stigmatized interactions between members of different groups. Another potential source of such a variation in value could be that the expected chance of succeeding would be higher for certain attorney-client pairings than for others. However, there is no apparent reason to believe that members of minorities had a better chance to succeed in court if minority attorneys represented them or that white plaintiffs fared better in court when represented by white attorneys.

Second, our study shows a differential treatment of senders in personal injury lawsuits only. A notable feature of this context is that lawyers' fees are often paid on a contingent basis. We leave it to future research to explore whether our results replicate in areas such as criminal law, in which attorneys usually demand up-front retainers.[29] We also cannot rule out the possibility that the attorneys contacted in this study constitute an unrepresentative sample of the population of private injury attorneys in the United States, and that some or all of the results presented in this analysis do not hold true for this population.

Differential access to legal representation could have a variety of social and economic consequences. Perhaps most obviously, access to the courts to gain compensation from past harm is a good with economic value, and if access is granted or denied based on demographic characteristics, then that good will be unequally distributed. Such differential treatment could exacerbate existing social and economic inequalities. In addition, the basic legitimacy of the court system, and the rule of law more generally, may be undermined if certain groups of people consistently find that their efforts to achieve judicial remedies for private harms are stymied.

These consequences raise the question whether there are potential remedies available to

29. Results from a field experiment involving criminal defense attorneys in California (Libgober 2020) suggest that members of minorities might face similar discrimination in these areas. However, it should be noted that this study was unable to replicate these effects for attorneys in Florida.

tackle this problem. Furnishing clients with a right to bring suit against an attorney whose decision to decline representation because of the client's racial/ethnic status will probably be of little avail. This is because clients will usually not have the information needed to establish whether an attorney exhibits a pattern of preferential treatment of members of other groups. Establishing forms of legal aid that include the financing of lawsuits brought by members of minorities that have a lower chance to obtain legal representation by an attorney in principle seems more effective. However, it would require administrators to take over the role of attorneys in screening cases for whether they have a chance to result in a successful lawsuit or not. If administrators performed worse in this task than attorneys do, a higher share of bogus claims would end up in court, wasting not only the resources allocated to such a program, but also other court resources. A third way to tackle this problem would be to strive for a more diverse bar. If there are enough attorneys from different racial/ethnic groups in all parts of the country, the problem of differential access to legal representation would be severely curtailed.

7 Conclusion

This paper reports results from a nationalwide study of the effects of demographic information (as encoded in names) on attorney behavior at the initial inquiry stage of client intake for personal injury claims. Our study implements an email audit including more than 24,000 attorneys from all across the United States and documents whether response rates vary with sender names suggesting senders belonging to different racial/ethnic groups. We find significant results suggesting a discriminatory effect preferring some members of the majority race/ethnicity. We also find tentative evidence that these effects vary depending on personal characteristics of the attorneys. The latter finding might suggest that the observed differences in the treatment of different groups are not merely a reaction to anticipated differences in the expected payoff from lawsuits involving different types of plaintiffs, but that attorneys' selection decisions are at least partly a function of their personal characteristics. Mechanisms that could explain this latter effect include "taste-based" discrimination and ill-informed beliefs about members of other groups.

References

Abrams, David S., Marianne Bertrand, and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race." *Journal of Legal Studies* 41 (2): 374–384.

Agan, Amanda, and Sonja Starr. 2018. "Ban the Box, Criminal Records, and Statistical Discrimination: A Field Experiment." *The Quarterly Journal of Economics* 133 (1): 191–235.

———. 2020. "Employer Neighborhoods and Racial Discrimination." NBER Working Paper 28153. Accessed October 14, 2022. https://www.nber.org/papers/w28153.

Alesina, Alberto, and Emiliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." *American Economic Review* 104 (3397-3433).

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100 (401): 464–477.

Arnold, David, Will Dobbie, and Crystal S. Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics* 133 (4): 1885–1932.

Ayres, Ian. 1991. "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations." *Harvard Law Review* 104:817–872.

Ayres, Ian, and Peter Siegelman. 1995. "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations." *The American Economic Review* 85 (3): 304–321.

Baert, Stijn. 2018. "Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005." Chap. 3 in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance,* edited by S. Michael Gaddis, 63–77. Cham, Switzerland: Springer.

Baradaran, Shima, Michael Findley, Daniel Nielson, and J. C. Sharman. 2013. "Does International Law Matter." *Minnesota Law Review* 97 (3): 743–837.

Baradaran, Shima, Michael Findley, Daniel Nielson, and Jason Sharman. 2014. "Funding Terror." *University of Pennsylvania Law Review* 162 (3): 477–536.

Becker, Gary S. 1971. *Economics of Discrimination.* 2nd Revised edition. 178. University of Chicago Press.

Bertrand, Marianne, and Esther Duflo. 2017. "Field Experiments on Discrimination." Chap. 8 in *Handbook of Field Experiments,* edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:309–393. North Holland.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *The American Economic Review* 94 (4): 991–1013.

Block, Ray, Charles Crabtree, John B. Holbein, and J. Quin Monson. 2021. "Are Americans less likely to reply to emails from Black people relative to White people?" *Proceedings of the National Academy of Sciences* 118 (52): e2110347118.

Bonica, Adam. 2016. *Database on Ideology, Money in Politics, and Elections: Public version 2.0 [computer file].* Stanford, CA: Stanford University Libraries. https://data.stanford.edu/dime.

Bonica, Adam, Adam S. Chilton, and Maya Sen. 2016. "The Political Ideologies of American Lawyers." *Journal of Legal Analysis* 8 (2): 277–335.

Bonica, Adam, and Maya Sen. 2017. "The Politics of Selecting the Bench from the Bar: The Legal Profession and Partisan Incentives to Introduce Ideology into Judicial Selection." *Journal of Law and Economics* 60:559–595.

Braucher, Jean, Dov Cohen, and Robert M. Lawless. 2012. "Race, Attorney Influence, and Bankruptcy Chapter Choice." *Journal of Empirical Legal Studies* 9 (3): 393–429.

Broockman, David E. 2013. "Black politicians are more intrinsically motivated to advance Blacks' interests: A field experiment manipulating political incentives." *American Journal of Political Science* 57 (3): 521–536.

Butler, Daniel M., and David E. Broockman. 2011. "Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators." *American Journal of Political Science* 55 (3): 463–477.

Cotropia, Christopher Anthony, Jonathan S. Masur, and David L. Schwartz. 2019. "Gender Discrimination in Online Markets." University of Chicago Coase-Sandor Institute for Law & Economics Research Paper No. 864. Accessed December 30, 2021. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3200409.

Crabtree, Charles. 2018. "An Introduction to Conducting Email Audit Studies." Chap. 5 in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance,* edited by S. Michael Gaddis, 103–117. Cham, Switzerland : Springer.

Crabtree, Charles, and Volha Chykina. 2018. "Last Name Selection in Audit Studies." *Sociological Science* 5:21–28.

Eagly, Ingrid V, and Steven Shafer. 2015. "A National Study of Access to Counsel in Immigration Court." *University of Pennsylvania Law Review* 164 (1): 1–91.

Egami, Naoki, Christian J. Fong, Justin Grimmer, Margaret E. Roberts, and Brandon M. Stewart. 2018. "How to Make Causal Inferences Using Texts." Accessed December 30, 2021. https://arxiv.org/pdf/1802.02163.

Einstein, Katherine Levine, and David M. Glick. 2017. "Does Race Affect Access to Government Services? An Experiment Exploring Street-Level Bureaucrats and Access to Public Housing." *American Journal of Political Science* 61 (1): 100–116.

Fafchamps, Marcel, and Julien Labonne. 2017. "Using Split Samples to Improve Inference on Causal Effects." *Political Analysis* 25 (4): 465–482.

Felstiner, William LF, Richard L Abel, and Austin Sarat. 1980. "The Emergence and Transformation of Disputes: Naming, Blaming, Claiming..." *Law and Society Review* 15 (3/4): 631–654.

Frankenreiter, Jens, and Michael A. Livermore. 2020. "Computational Methods in Legal Analysis." *Annual Review of Law and Social Science* 16:39–57.

Gaddis, S. Michael. 2017a. "How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies." *Sociological Science* 4:469–489.

———. 2017b. "Racial/Ethnic Perceptions from Hispanic Names: Selecting Names to Test for Discrimination." *Socius* 3:1–11.

———. 2018. "An Introduction to Audit Studies in the Social Sciences." Chap. 1 in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance,* edited by S. Michael Gaddis, 3–44. Cham, Switzerland: Springer.

Gaddis, S. Michael, Edvard Nergå rd Larsen, Charles Crabtree, and John B. Holbein. 2021. "Discrimination Against Black and Hispanic Americans is Highest in Hiring and Housing Contexts: A Meta-Analysis of Correspondence Audits." Accessed January 12, 2022. https://doi.org/10.17605/OSF.IO/4EQN2.

Greiner, D. James, and Cassandra Wolos Pattanayak. 2012. "Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make?" *Yale Law Journal* 121:2118–2214.

Greiner, D. James, Cassandra Wolos Pattanayak, and Jonathan Hennessy. 2013. "The Limits of Unbundled Legal Assistance: A Randomized Study in a Massachusetts District Court and Prospects for the Future." *Harvard Law Review* 126:901–989.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis,* 1–31.

Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences : an Introduction.* New York: Cambridge University Press.

Jensen, Eric, Nicholas Jones, Kimberly Orozco, Lauren Medina, Marc Perry, Ben Bolender, and Karen Battle. 2021. "Measuring Racial and Ethnic Diversity for the 2020 Census." *Census.gov Random Samplings,* https://www.census.gov/newsroom/blogs/random-samplings/2021/08/measuring-racial-ethnic-diversity-2020-census.html.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human Decisions and Machine Predictions." *The Quarterly Journal of Economics* 133 (1): 237–293.

Libgober, Brian. 2020. "Getting a Lawyer While Black: A Field Experiment." *Lewis and Clark Law Review* 24:forthcoming.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2018. *IPUMS National Historical Geographic Information System: Version 13.0 [Database].* Minneapolis: University of Minnesota. http://doi.org/10.18128/D050.V13.0.

———. 2019. *IPUMS National Historical Geographic Information System: Version 14.0 [Database].* Minneapolis: University of Minnesota. http://doi.org/10.18128/D050.V14.0.

Miles, Thomas J. 2012. "Racial Disparities in Wiretap Applications before Federal Judges." *Journal of Legal Studies* 41 (2): 419–458.

MIT Election Data and Science Lab. 2018. *County Presidential Election Returns 2000-2016.* V. V6. Harvard Dataverse. https://doi.org/10.7910/DVN/VOQCHQ. https://doi.org/10.7910/DVN/VOQCHQ.

Moser, Petra. 2012. "Taste-based discrimination evidence from a shift in ethnic preferences after WWI." *Explorations in Economic History* 49 (2): 167–188.

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62 (4): 659–661.

Posner, Richard A. 1989. "An Economic Analysis of Sex Discrimination Laws." *University of Chicago Law Review* 56:1311–1335.

Shayo, Moses, and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *The Quarterly Journal of Economics* 126:1447–1484.

Spamann, Holger, and Lars Klöhn. 2016. "Justice is Less Blind, and Less Legalistic, then We Thought: Evidence from an Experiment with Real Judges." *The Journal of Legal Studies* 45 (2): 255–280.

Stangor, Charles. 2009. "The Study of Stereotyping, Prejudice, and Discrimination Within Social Psychology. A Quick History of Theory and Research." Chap. 1 in *Handbook of Prejudice, Stereotyping, and Discrimination,* edited by Todd D. Nelson, 1–22. New York: Psychology Press.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statisticval Society, Series B (Methodological)* 58 (1): 267–288.

Underhill, Kirsten. 2019. "Price and Prejudice: An Empirical Test of Financial Incentives, Altruism, and Racial Bias." *Journal of Legal Studies* 48:245–274.

White, Ariel R., Moah L. Nathan, and Julie K. Faller. 2015. "What Do I Need to Vote? Bureaucratic Discretion and Discrimination by Local Election Officials." *American Political Science Review* 109 (1): 129–142.

Yang, Crystal. 2015. "Free at Last? Judicial Discretion and Racial Disparities in Federal Sentencing." *Journal of Legal Studies* 44:75–111.

OA1 Text of Inquiries

This section reports the text of the email inquiries described in Section 3.4 of the paper. The following text was used in 10,786 inquiries sent during the first round of the study and in all inquiries sent during the second round.

> **Subject: Supermarket accident**
>
> *Dear Ms./Mr. [ATTORNEYNAME] / Hello:*
>
> *My name is [SENDERNAME]. I am contacting you to ask if your law firm handles the case of a very serious ankle injury that included two surgeries, with a plate and screws, and several months of incapacitation and rehabilitation. The injury happened in a grocery store due to VERY unsafe conditions. In addition to medical bills, there were several months of no employment because of this and pain and lack of mobility still makes working difficult.*
>
> *Thank you,*
>
> *[SENDERNAME]*

The following text was used in 2,000 inquiries sent during the first round of the study. We did not use this version of the text during the second round.

> **Subject: Supermarket accident**
>
> *Dear Ms./Mr. [ATTORNEYNAME] / Hello:*
>
> *My name is [SENDERNAME]. I am contacting you to ask if your law firm handles the case of a very serious ankle injury that included two surgeries, with a plate and screws, and several months of incapacitation and rehabilitation. The injury happened in a grocery store due to VERY unsafe conditions. Especially as a retired senior citizen, pain and lack of mobility make it impossible to engage in normal activities.*
>
> *Thank you,*
>
> *[SENDERNAME]*

OA2 Alternative Analysis Using Positive Responses as the Outcome Measure

Our main analysis measures whether attorneys responded to outreach emails or not. In doing so, we essentially treat every email received in response to our outreach as a positive reaction on the attorney's part. While ensuring easy replicability of our results and avoiding the need for judgment calls about whether an email suggests a willingness on the part of an attorney to represent the purported inquirer, this approach also discards potentially useful information. Most importantly, if attorneys are more likely to send rejection emails to some sender types than others (instead of not answering an email at all), this could explain at least parts of the results reported in Section 5 of the paper.

In response to this issue, we repeat the analyses above using a different outcome measure that evaluates whether a response indicated an attorney's willingness to help the inquirer obtain legal representation. Accordingly, this outcome measure treats "negative responses" equivalent to instances in which the outreach email did not trigger a response at all.

OA2.1 Using Machine Learning to Predict Negative Responses

To obtain this alternative outcome measure, we hand-code 1,003 randomly selected emails from among the emails that we received in response to inquiries sent during the first round of the study and that were not flagged as error messages.[30] For this, we use a coding scheme that distinguished between whether a response indicates an attorney's willingness to work with a client (Category 1), a response referred the inquirer to another attorney (Category 2), the response indicates an attorney's unwillingness to represent a client (Category 3), or the response falls in none of these categories (Category 4).[31]

Table OA1 contains example emails from each category and indicates the frequency with which each of these categories occur in the sample. The majority of responses ($\sim 78\%$) received indicates an attorney's willingness to at least explore representing the inquirer, while less than 7% contains a clear refusal to work with them.

In the following analysis, we treat responses as positive responses if they fall into Categories 1 or 2, and as negative responses otherwise.

For the remaining email responses, we use a machine learning algorithm to predict from the vocabulary used in emails whether unlabeled emails contain negative responses or not. We select and fine-tune the algorithm on the basis of results obtained from 10-fold cross validation. We implement the following steps to obtain these predictions:

1. We convert all emails into vector representations using a simple "bag-of-words" approach. For this, we obtain dummy indicators for whether a word is featured in the email or not, which we rescale using the "inverse document frequency" (IDF) algorithm (*see* Frankenreiter and Livermore 2020).

30. *See* Section 3.5 of the paper.
31. One of the co-authors (Frankenreiter) did the manual coding in August 2019.

Table OA1: Response Categories.

Category	Description	Example email	# in sample	Treated as
1	Indicates willingness to consider representing the client	*I shall discuss the case with you; give me a call at [PHONE NUMBER]. [ATTORNEY SIGNATURE].*	782	Positive
2	Refers the client to another attorney	*Dear Mr. Battle- I no longer o [sic] Personal Injury cases. However, I can get your details and refer you to one of a few very good injury attorneys in my circle of attorney friends. What is your phone nimber [sic]?*	74	Positive
3	Declines representation	*Dear Ms. Alston, I do not handle these types of claims. Thank you for contacting my office. [ATTORNEY NAME].*	70	Negative
4	None of the above	*I have received your message. I will be preparing a Court of Appeals argument and will be unable to respond to your message until Thursday. Thank you for your understanding. [ATTORNEY NAME].*	77	Negative

2. We reduce the dimensionality of the ensuing dataset using principal component analysis (PCA). We use the first 50 factor loadings as "features" in our classification algorithm.

3. We train a LASSO classifier (Tibshirani 1996) to predict whether the text of the email indicates a negative response or not.

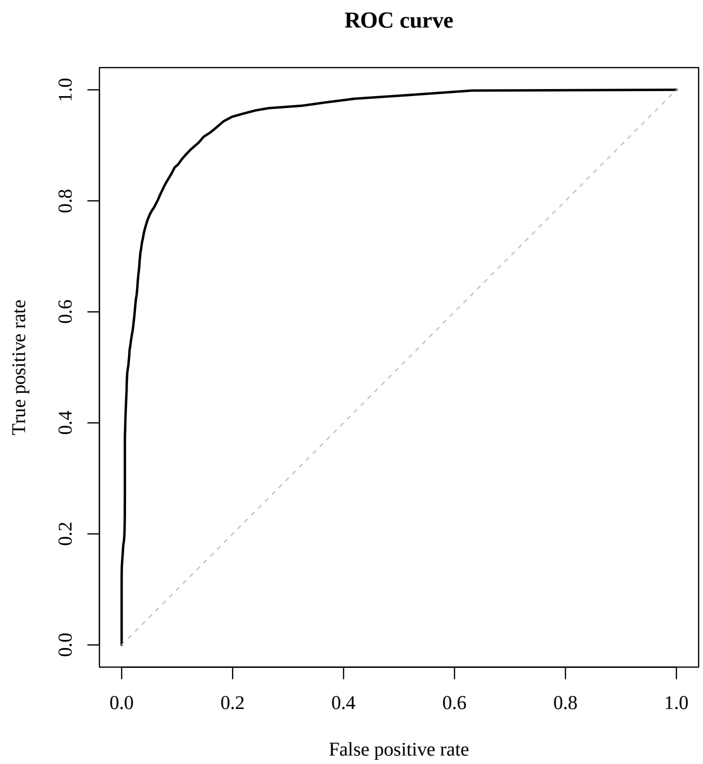
4. We use this algorithm to generate predictions for the unlabeled part of our dataset.

To evaluate the performance of this algorithm, consider Figure OA1. This figure contains a so-called receiver operating characteristic (ROC) curve that describes the predictive accuracy of a binary classifier in dependence on the discrimination threshold used. A curve that bends far to the left (like the one in the graph) is indicative of a strong predictive accuracy.

Because our dataset contains more positive than negative responses, we calibrate the final version of the algorithm so that it classifies a number of emails as negative responses that corresponds to the share of negative responses in the coded part of the dataset. Results from our 10-fold cross validations confirm that the classification algorithm performs reasonably well in this configuration. In 20 repetitions using different (random) sample splits, we achieve an average accuracy of $\sim$ .927 and an average F-1 score of $\sim$ .752.[32]

32. These numbers imply that, on average, the algorithm correctly identifies 111 of the 147 negative responses in the labeled part of our dataset.

Figure OA1: Performance of LASSO classifier predicting negative responses



OA2.2 Results

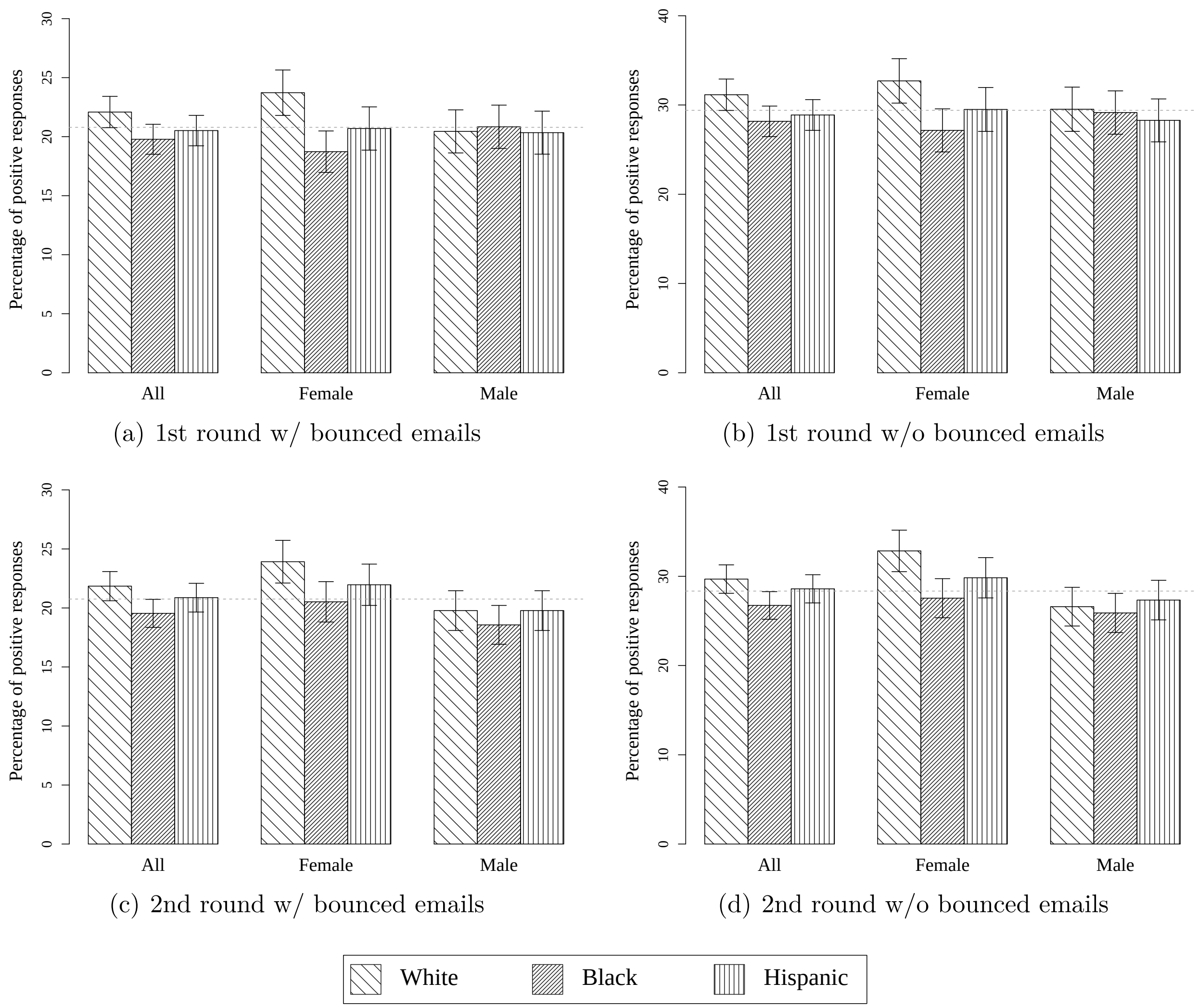
OA2.2.1 Differences in Response Rates

Next, we repeat the analyses reported in Section 5 of the paper using this alternative outcome measure. While some effects attenuate, all major results are robust to this change.

Figure OA2 replicates Figure 1. As can be seen from this figure, White inquirers receive the highest share of positive responses (832 or 22.1% during the first round, 939 or 21.8% during the second round), followed by Hispanic (772 or 20.5% and 897 or 20.9%, respectively) and Black (749 or 19.8% and 840 or 19.5%) senders. A Fisher's Exact test for differences between response rates for Black and Hispanic senders on the one hand and White senders on the other hand indicates that this difference is statistically significant in both rounds of the study (p-values: .017 and .032). Like in the analysis reported in Section 5.1 of the paper, almost all of this difference is driven by differences in responses to inquiries sent under female identities.[33]

33. Fisher's Exact tests comparing response rates for all Black and Hispanic inquirers with White female inquirers only yield highly significant results (p-values: .0008 and .0002), while the same test yields insignificant results (p-values: .773 and .674) if White male senders are substituted for White female senders.

Figure OA2: Response rates by gender and race/ethnicity group



(a) 1st round w/ bounced emails

(b) 1st round w/o bounced emails

(c) 2nd round w/ bounced emails

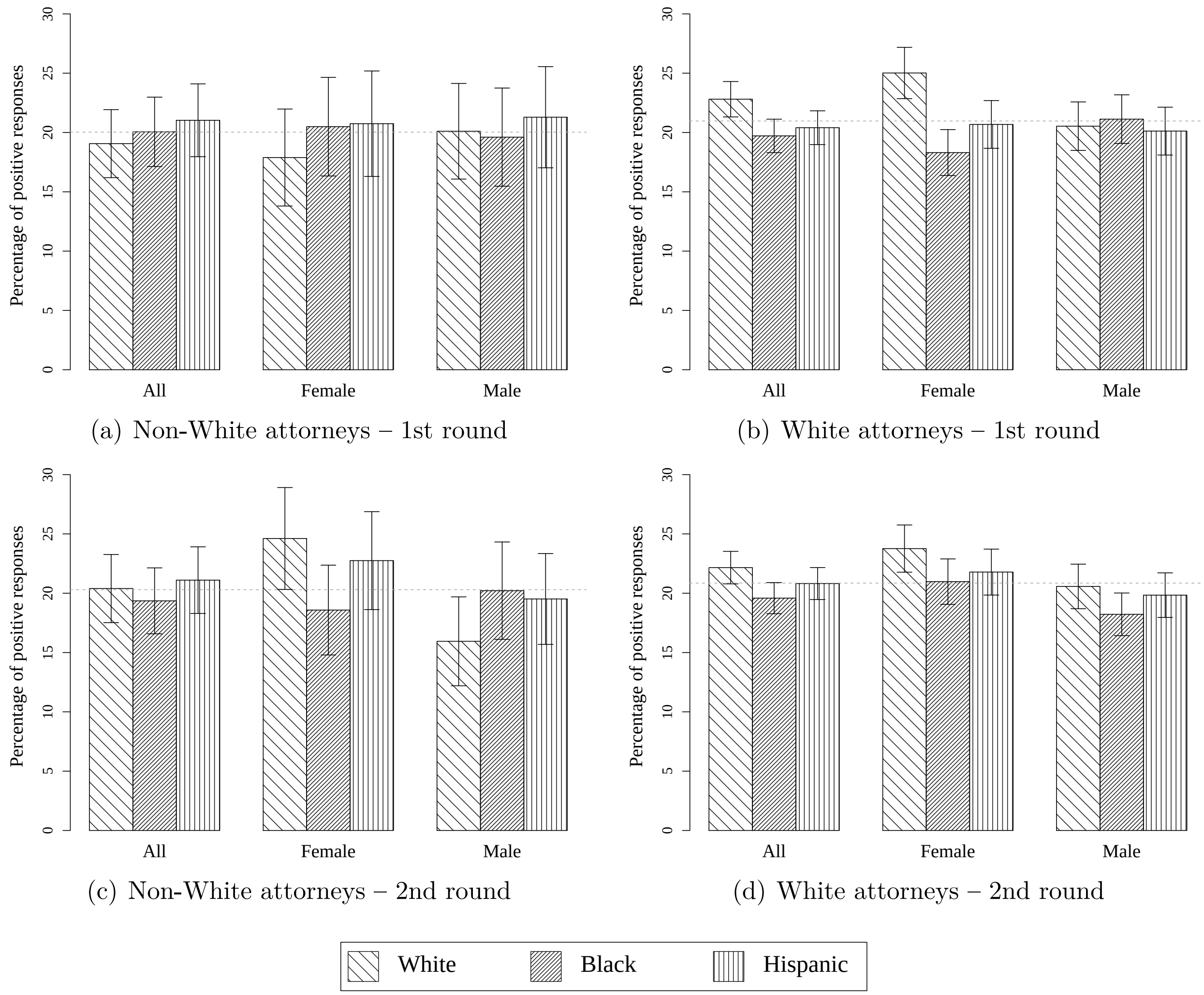
(d) 2nd round w/o bounced emails

OA2.2.2 Evidence of In-Group Bias

In addition, we replicate the results reported in Section 5.2 of the paper to rule out the possibility that our findings there are solely a function of our choice of outcome variable. Figure OA3 replicates Figure 2. Like in the previous analysis, it is apparent that the differential treatment of senders from different racial/ethnic groups (and, in particular, the preferential treatment of White female senders) are almost exclusively driven by attorneys with names most common among White people.

We also repeat the various steps undertaken to rule out the potential confounding influence of geographical factors, including the matching strategy described in Section 5.2.2 of the paper. The results from this analysis, which we report in Section OA6, provide additional evidence that the differential treatment of senders from different racial/ethnic groups are almost exclusively driven by attorneys with names most common among White people is independent from the choice of the outcome variable.

Figure OA3: Differences between treatment effects by attorney race



(a) Non-White attorneys – 1st round

(b) White attorneys – 1st round

(c) Non-White attorneys – 2nd round

(d) White attorneys – 2nd round

OA3 Information on Attorney Sample

Table OA2: Information on personal injury lawyers by state.

	Population estimate for 2017	Number lawyers in data	Lawyers per 100k inhabitants	Number lawyers w/ email	Percent. lawyers w/ email
Alabama	4,874,747	1,916	39.30	1,154	60.23
Alaska	739,795	279	37.71	185	66.31
Arizona	7,016,270	1,444	20.58	817	56.58
Arkansas	3,004,279	867	28.86	445	51.33
California	39,536,653	7,045	17.82	3,115	44.22
Colorado	5,607,154	920	16.41	608	66.09
Connecticut	3,588,184	1,284	35.78	922	71.81
Delaware	961,939	263	27.34	179	68.06
District of Columbia	693,972	344	49.57	205	59.59
Florida	20,984,400	5,754	27.42	4,673	81.21
Georgia	10,429,379	2,709	25.97	2,041	75.34
Hawaii	1,427,538	373	26.13	279	74.80
Idaho	1,716,943	351	20.44	187	53.28
Illinois	12,802,023	4,430	34.60	2,359	53.25
Indiana	6,666,818	1,874	28.11	952	50.80
Iowa	3,145,711	800	25.43	429	53.62
Kansas	2,913,123	779	26.74	386	49.55
Kentucky	4,454,189	893	20.05	585	65.51
Louisiana	4,684,333	2,813	60.05	1,583	56.27
Maine	1,335,907	154	11.53	108	70.13
Maryland	6,052,177	843	13.93	594	70.46
Massachusetts	6,859,819	3,137	45.73	1,621	51.67
Michigan	9,962,311	2,104	21.12	1,526	72.53
Minnesota	5,576,606	1,602	28.73	1,075	67.10
Mississippi	2,984,100	1,039	34.82	583	56.11
Missouri	6,113,532	2,269	37.11	1,351	59.54
Montana	1,050,493	501	25.57	284	56.69
Nebraska	1,920,076	627	32.65	379	60.44
Nevada	2,998,039	675	22.51	404	59.85
New Hampshire	1,342,795	506	37.86	257	50.79
New Jersey	9,005,644	3,754	41.68	1,993	53.09
New Mexico	2,088,070	659	31.56	336	50.97
New York	19,849,399	6,874	34.63	3,663	53.29
North Carolina	10,273,419	1,181	11.50	902	76.38
North Dakota	755,393	184	24.36	134	72.83
Ohio	11,658,609	3,300	28.31	1,803	54.64
Oklahoma	3,930,864	1,430	36.38	806	56.36
Oregon	4,142,776	963	23.24	708	73.52
Pennsylvania	12,805,537	4,940	38.58	2,589	52.41
Rhode Island	1,059,639	577	54.45	324	56.15
South Carolina	5,024,369	851	16.94	611	71.80
South Dakota	869,666	444	51.05	260	58.56
Tennessee	6,715,984	1,754	26.12	1,121	63.91
Texas	28,304,596	6,611	23.36	3,758	56.84
Utah	3,101,833	511	16.47	349	68.30
Vermont	623,657	310	49.71	182	58.71
Virginia	8,470,020	1,912	22.57	1,129	59.05
Washington	7,405,743	2,208	29.81	1,262	57.16
West Virginia	1,815,857	851	46.86	478	56.17
Wisconsin	5,795,483	898	15.49	514	57.24
Wyoming	579,315	238	41.08	120	50.42
TOTAL	325,719,178	89,045	27.34	52,328	58.77

Notes. Population estimates obtained from U.S. Census Bureau, available online at https://www2.census.gov/programssurveys/popest/tables/20102017/state/totals/nstest2017-01.xlsx (last accessed January 29, 2019).

Figure OA4: Representativeness of attorney sample (1)

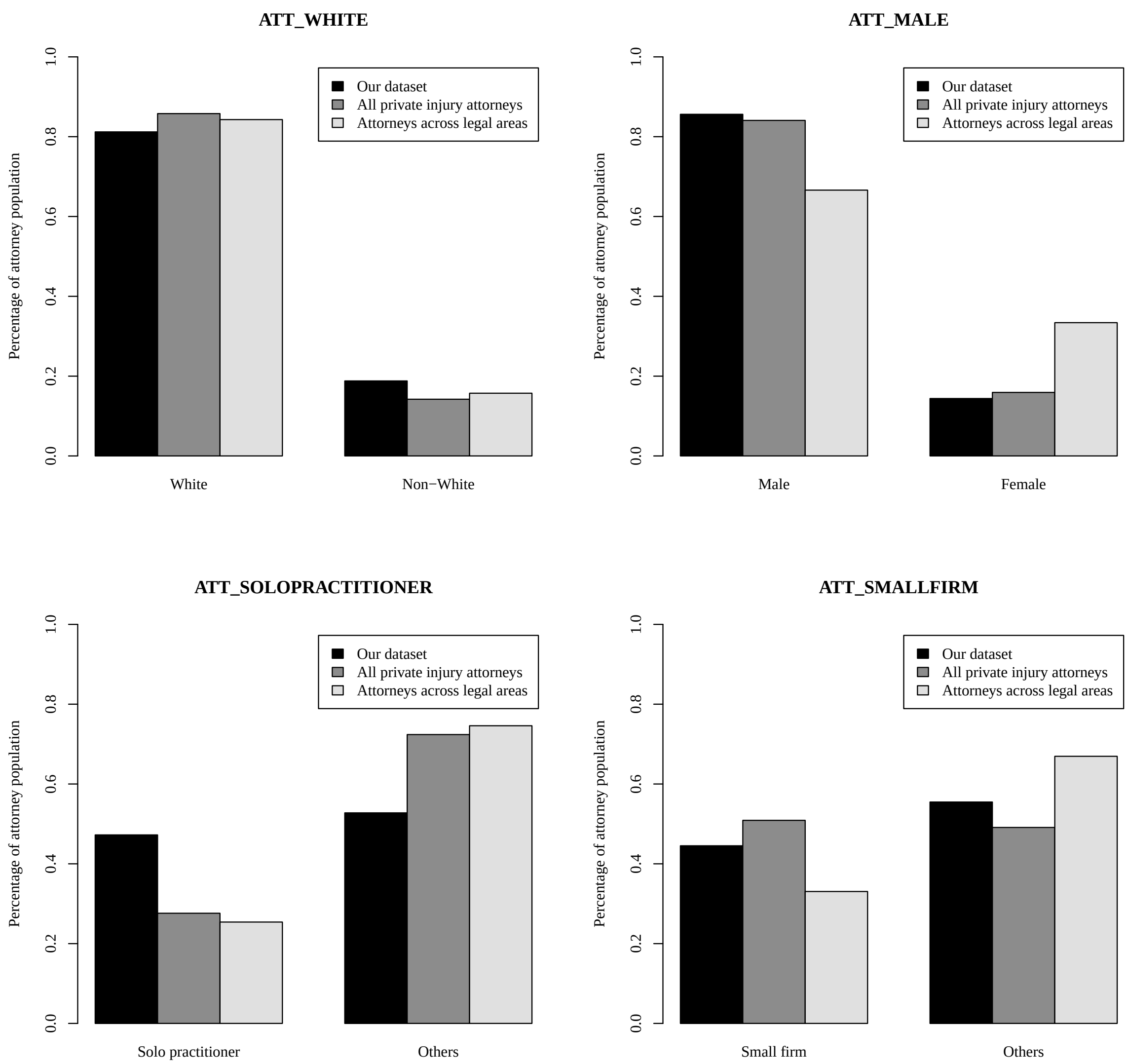
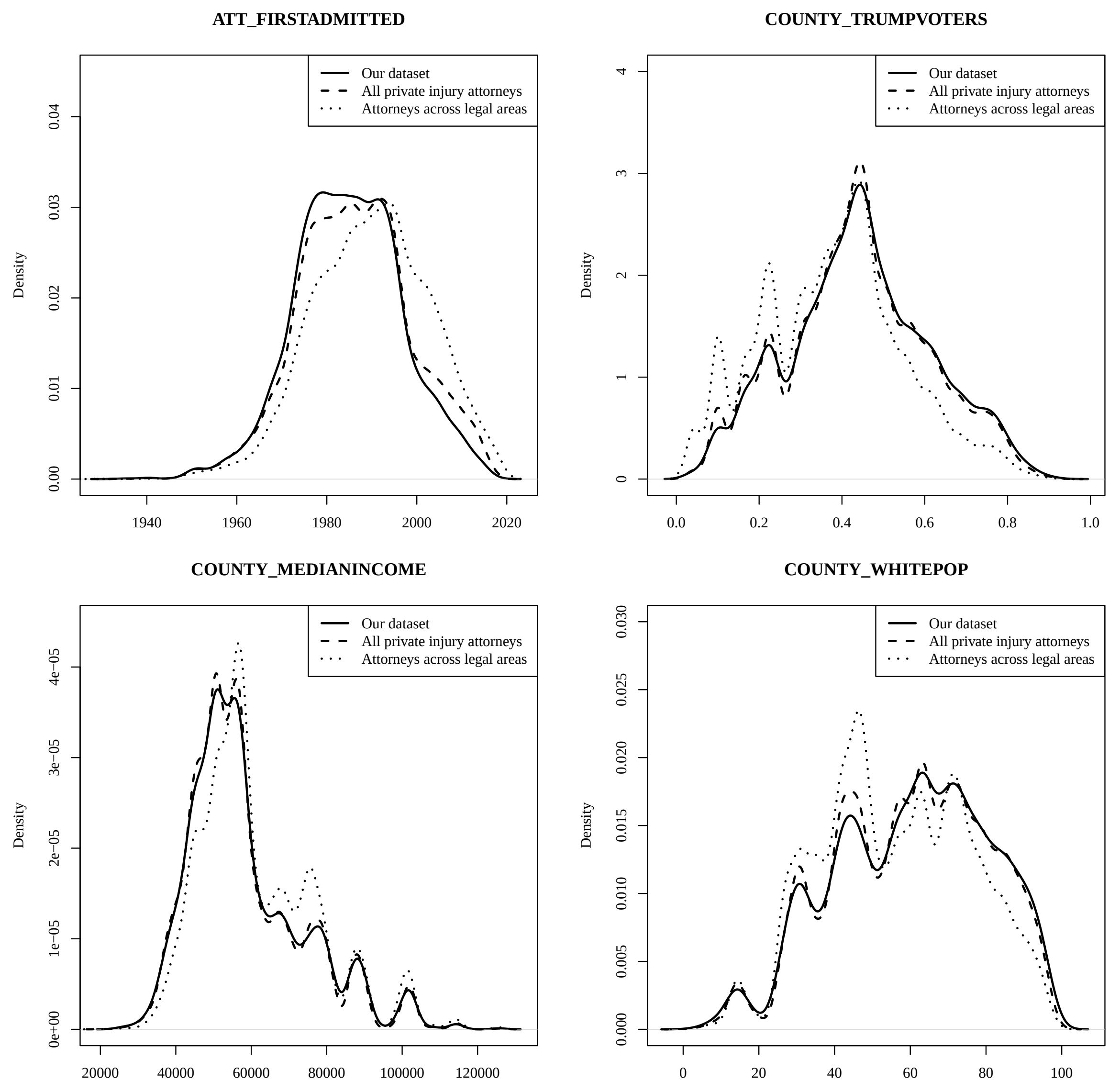


Figure OA5: Representativeness of attorney sample (2)



OA4 Regression Analysis

In this section, we report the results from a series of regressions aimed at confirming the existence of the main effect documented in this paper, namely the impact of senders' purported race/ethnicity and gender on response rates.

Table OA3 reports the results from a number of different logit regressions. In both panels in Table OA3, Columns (1) and (2) report results of regressions using the first round data, columns (3) and (4) results for the second round data, and columns (5) and (6) results for the combined data from both rounds. The odd-numbered columns report results using all inquiries, while the results reported in even-numbered columns were obtained from regressions using only emails that did not result in an error message. All regressions include state fixed effects. Standard errors are clustered at the state level.[34]

The regressions reported in Panel (a) compare the response rates for Black and Hispanic senders with those for White senders. As can be seen from the table, the estimates for both Black and Hispanic are negative and significant in all specifications. This confirms that members of minority groups indeed have a harder time initiating a relationship with attorneys. Furthermore, the negative effect for MALE confirms that attorneys preferentially respond to female senders.

The introduction of separate variables for each combination of racial/ethnic group and gender in the even-numbered columns confirms the finding that the difference in responses between different groups is largely driven by differences in responses to female inquirers.

The size of the estimates is also substantially significant. For example, the point estimate of -.227 for Black female senders in column (9) suggests that an attorney who responds with a probability of 30% to senders using a name common among White females will only respond to a sender writing under a name common with Black females with a probability of 25.4%.

Because of the random assignment of sender names to attorneys, our estimates for the main treatment effects would be unbiased even in absence of controls for attorney characteristics. Nevertheless, as described above, all regressions reported in Table OA3 include state fixed effects. Under randomized assignment, the inclusion of these variables should not substantially affect the size of the estimated coefficients. In a series of unreported regressions, we confirm that the inclusion of fixed effects does not meaningfully alter the size of our estimates, suggesting that our method of randomization worked as expected.

Substantially similar results can be obtained when using the alternative outcome variable described in Section OA2 *in lieu* of our main outcome variable. Table OA4 reports the results of regressions that mirror those reported in Table OA3, with the exception that we use positive responses as the outcome variable. All estimates for differences between Black senders and White or White female senders remain significant. The same is not true for Hispanic senders, which suggests that attorneys are less likely to respond to inquiries by perceived Black senders than to those by senders perceived to belong to other racial/ethnic minorities.

34. Using "standard" Huber-White standard errors instead of clustered standard errors does not materially affect the results.

Table OA3: Logit Regressions.

(a) Aggregated Group Categories

	1st Round		2nd Round		Combined Data	
	(1)	(2)	(3)	(4)	(5)	(6)
BLACK	-.174***	-.194***	-.108*	-.113*	-.138***	-.146***
	(.041)	(.041)	(.050)	(.055)	(.033)	(.037)
HISPANIC	-.139*	-.165*	-.104*	-.107*	-.120**	-.131**
	(.064)	(.075)	(.050)	(.045)	(.036)	(.038)
MALE	-.053	-.053	-.153***	-.160***	-.105***	-.112***
	(.037)	(.044)	(.038)	(.040)	(.025)	(.027)
_Intercept	-1.41***	-.467***	-1.31***	-1.01***	-1.35***	-.801***
	(.107)	(.126)	(.065)	(.072)	(.082)	(.088)
N	11,317	8,003	12,894	9,442	24,211	17,445

(b) Disaggregated Group Categories

	1st Round		2nd Round		Combined Data	
	(1)	(2)	(3)	(4)	(5)	(6)
BLACK-FEMALE	-.339***	-.330**	-.134+	-.184**	-.227***	-.247***
	(.091)	(.098)	(.073)	(.068)	(.063)	(.061)
BLACK-MALE	-.245***	-.253***	-.259***	-.273**	-.250***	-.262***
-	(.061)	(.070)	(.072)	(.085)	(.046)	(.053)
HISPANIC-FEMALE	-.247*	-.248**	-.117+	-.148*	-.182**	-.193***
-	(.085)	(.095)	(.071)	(.066)	(.051)	(.051)
HISPANIC-MALE	-.259**	-.277**	-.270***	-.300***	-.261***	-.287***
	(.077)	(.087)	(.072)	(.066)	(.050)	(.050)
WHITE-MALE	-.242***	-.200**	-.180**	-.235***	-.207***	-.222***
	(.068)	(.071)	(.068)	(.065)	(.048)	(.047)
_Intercept	-1.38***	-.460***	-1.45***	-1.13***	-1.41***	-.863***
	(.102)	(.113)	(.047)	(.043)	(.077)	(.080)
N	11,317	8,003	12,894	9,442	24,211	17,445

Notes. Dependent variable: binary variable indicating whether a response was received. Odd-numbered columns report results from regressions including bounced emails, even numbered columns results from regressions excluding these emails. Standard errors clustered at the state level in parentheses. Controls for different emails and state fixed effects included in all regressions. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Reference categories: Panel (a): WHITE. Panel (b): WHITE-FEMALE.

Table OA4: Regression results.

	1st Round		2nd Round Data		Combined Data	
(a) Aggregated Group Categories						
	(1)	(2)	(3)	(4)	(5)	(6)
BLACK	-.143***	-.154***	-.142**	-.151**	-.141***	-.147***
	(.035)	(.039)	(.047)	(.051)	(.029)	(.031)
HISPANIC	-.095	-.110	-.059	-.056	-.075+	-.078+
	(.067)	(.078)	(.059)	(.055)	(.039)	(.041)
MALE	-.033	-.029	-.169***	-.174***	-.105***	-.109***
	(.043)	(.049)	(.044)	(.047)	(.028)	(.031)
_Intercept	-1.93***	-1.09***	-1.68***	-1.40***	-1.79***	-1.28***
	(.115)	(.134)	(.074)	(.081)	(.089)	(.095)
StateF.E.	Y	Y	Y	Y	Y	Y
N	11,317	8,003	12,894	9,442	24,211	17,445

	1st Round		2nd Round Data		Combined Data	
(b) Disaggregated Group Categories						
	(1)	(2)	(3)	(4)	(5)	(6)
BLACK-FEMALE	-.305***	-.286**	-.199**	-.250**	-.246***	-.262***
	(.078)	(.086)	(.075)	(.073)	(.059)	(.058)
BLACK-MALE	-.174*	-.165*	-.324***	-.340***	-.250***	-.256***
	(.067)	(.077)	(.079)	(.090)	(.047)	(.054)
HISPANIC-MALE	-.179*	-.152	-.112	-.138*	-.142**	-.143**
	(.085)	(.093)	(.075)	(.070)	(.054)	(.053)
HISPANIC-MALE	-.202*	-.208*	-.245**	-.265***	-.223***	-.238***
	(.083)	(.094)	(.075)	(.071)	(.051)	(.053)
WHITE-MALE	-.195*	-.144+	-.245***	-.298***	-.221***	-.230***
	(.076)	(.079)	(.070)	(.070)	(.052)	(.054)
_Intercept	-1.89***	-1.07***	-1.82***	-1.52***	-1.84***	-1.33***
	(.110)	(.119)	(.048)	(.044)	(.085)	(.086)
StateF.E.	Y	Y	Y	Y	Y	Y
N	11,317	8,003	12,894	9,442	24,211	17,445

Notes. Dependent variable: binary variable indicating whether a positive response was received. Odd-numbered columns report results from regressions including bounced emails, even numbered columns results fromregressions excluding these emails. Standard errors clustered at the state level in parentheses. Controls for different emails and state fixed effects included in all regressions. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Reference categories: Panel (a): WHITE. Panel (b): WHITE-FEMALE.

OA5 Geographic Variation

In this section, we explore how the treatment of different senders varies between attorneys located in different areas. This analysis is motivated by the assumption that the relative value of representing clients from different racial/ethnic groups might vary across the country, most importantly because the legal system in some parts of the country might treat members of certain groups more or less favorably than others. Here and in the following, we focus on differences between the response rate for White senders and the response rate for Black and Hispanic senders combined. The reason for this is the fact that the response rates for the latter groups were similar in both rounds of the study.

As can be seen from Figure OA6, there is considerable geographical variation in differences in response rates between states. The plots in this figure display differences in response rates for purported inquirers who are members of different race/ethnicity groups by state. Each state is represented by one squared "bin", with the surface area of the bin equivalent to the overall numbers of inquiries sent to attorneys located in this state. The color of the bin and the direction of the lines in the bin indicate the gap between the response rates for inquiries by purported White inquirers and the response rates for inquiries by various groups of other inquirers, calculated in log odds ratios. A red bin (with downward-facing lines) indicates a higher response rate for White inquirers, while a blue bin (with upward-facing lines) indicates a higher response rate for other inquirers. Panel (a) shows differences in response rates between White senders and the combined data for Black and Hispanic senders, calculated on the basis of the data gathered during both rounds of the study. Because log odds ratios defy a straightforward interpretation, we provide two examples to illustrate the magnitude of these effects. The state of Michigan is depicted in medium red in the graph. Its log odds ratio is .383, indicating a substantially higher response rates for White senders as compared to the other groups combined. In fact, the response rates for White inquirers was 21.5%, while the response rates for the other groups was 15.7% on average. Pennsylvania, by contrast is depicted in a lighter red, corresponding to a log odds ratio of 0.135 and response rates of 29.1% and 26.4%, respectively.

Overall, the graphic suggests considerable geographic heterogeneity, but not necessarily one that tracks intuitive regional categories. However, the preferential treatment of White senders is not restricted to only a few localities: 27 out of 51 states show a difference in log odds in favor of White senders of at least .15.[35]

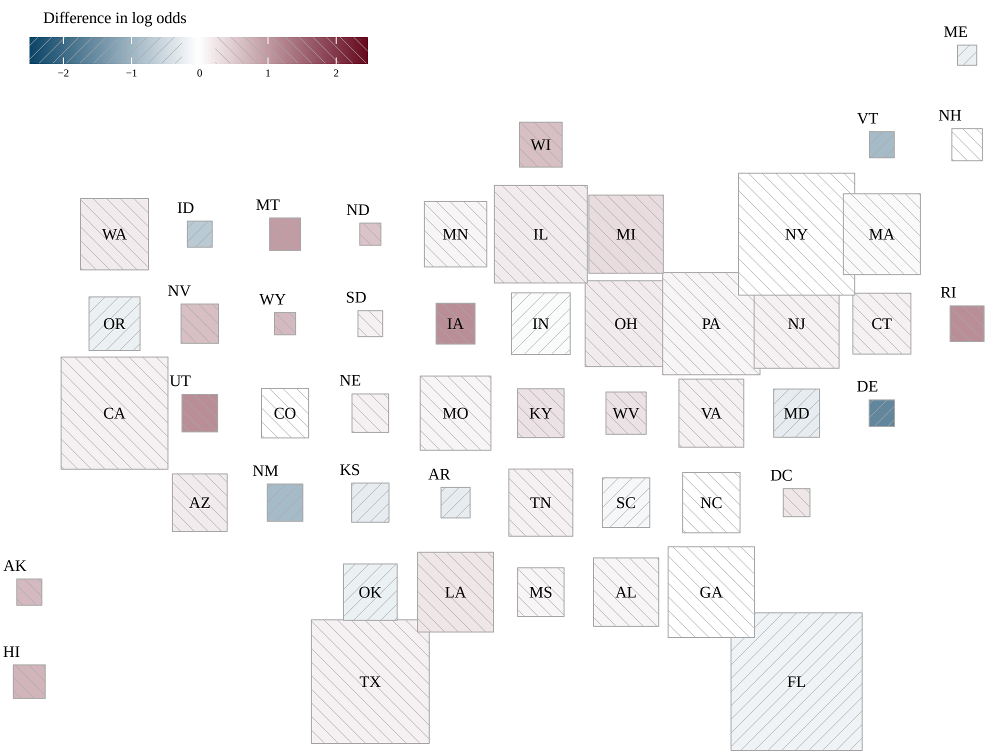
Panels (b) and (c) show the same results separately for Black and Hispanic senders. The geographic patterns of discrimination for both groups appear to be roughly similar.

Figure OA7 displays the differences in response rates between White female senders and other senders. This graphic shows comparable geographic patterns to the ones in Figure OA6, although the differences in response rates between White female and other senders are generally higher than the differences in response rates between White senders on the one hand side and Black and Hispanic senders on the other hand side.

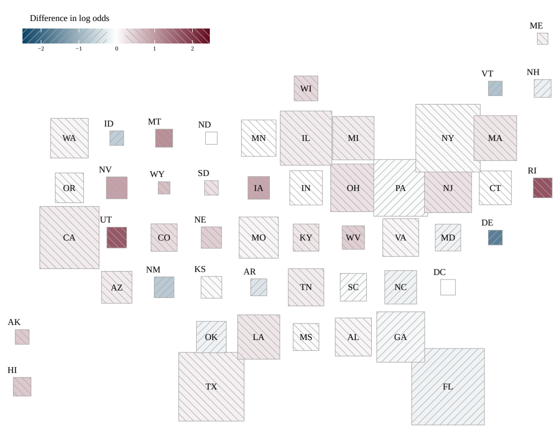
We also investigate whether differences in the response rates for different groups vary systematically with geographic factors at the sub-state level. We test variables indicative

35. Results for states with a small number of attorneys in the dataset (of which New Mexico is an example, as can also be seen from the size of the bins in Figure OA6) should be taken with a grain of salt, as a small number of inquiries of course increases the probability of more extreme result.

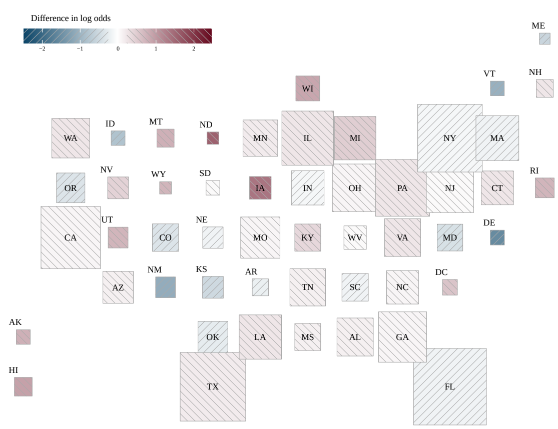
Figure OA6: Differences between treatment effects by state



(a) White vs. Black and Hispanic senders

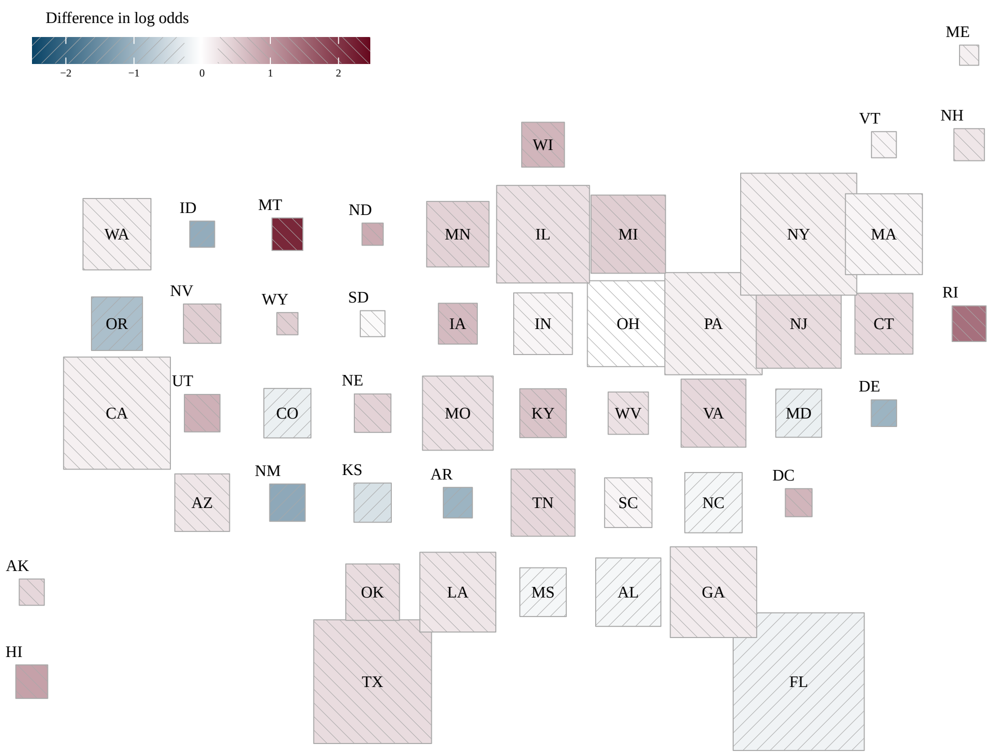


(b) White vs. Black senders

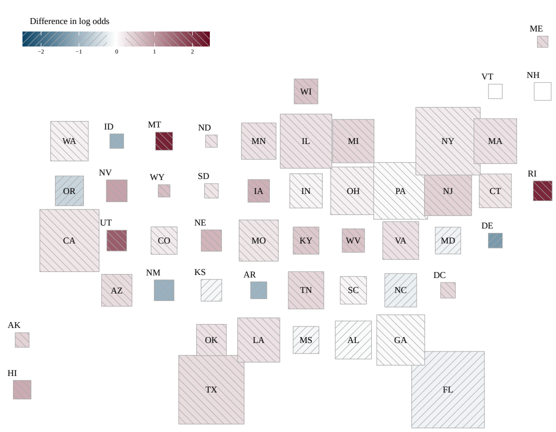


(c) White vs. Hispanic senders

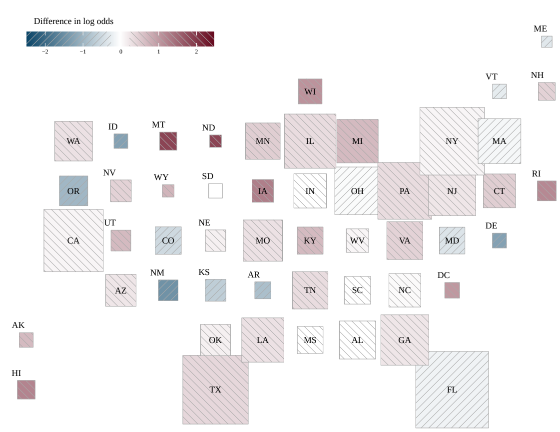
Figure OA7: Differences between treatment effects by state



(a) White female vs. Black and Hispanic senders



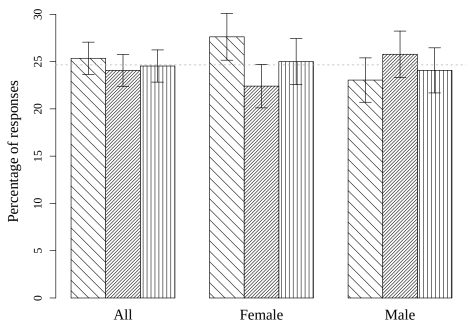
(b) White female vs. Black senders



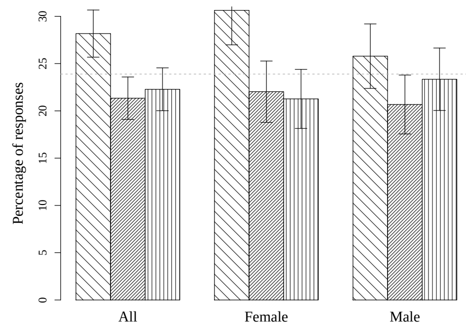
(c) White female vs. Hispanic senders

of the economic well-being of residents in particular areas (median income, poverty rates), the composition of the population (percentage of Whites in a certain area) as well as the economic fortunes of different groups of the population (the difference between the overall rate of poverty and the poverty rates of non-Hispanic Whites in an area). Most variables do not show a systematic relationship with the treatment of different groups. As Figure OA8 shows, we discovered a relationship between the percentage of votes for presidential candidate Donald Trump in the 2016 election in the county in which the attorney's office is located and the treatment of Black and Hispanic vs. White senders in the first round data: Both White female and White male senders fared considerably better in areas with high numbers of Trump voters. However, the same relationship did not exist in the second round data.

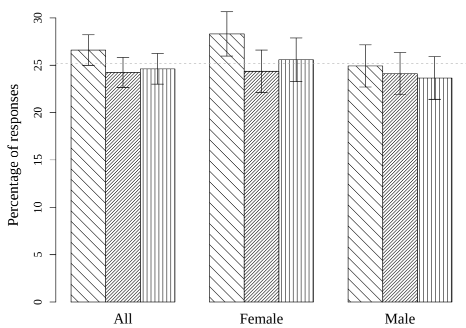
Figure OA8: Differences between treatment effects by % of Trump voters



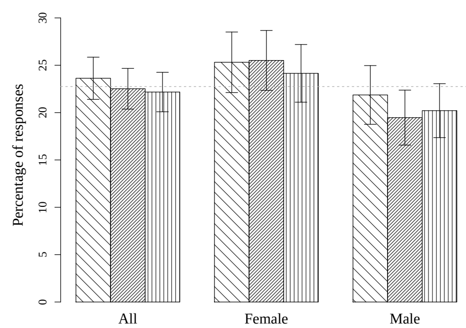
(a) Fewer Trump voters – 1st round



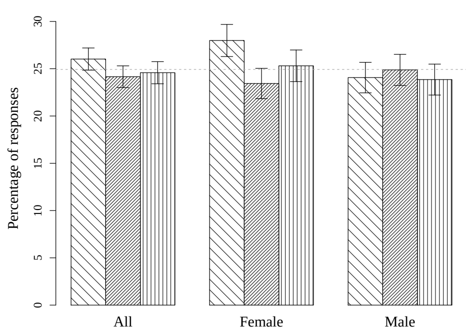
(b) More Trump voters – 1st round



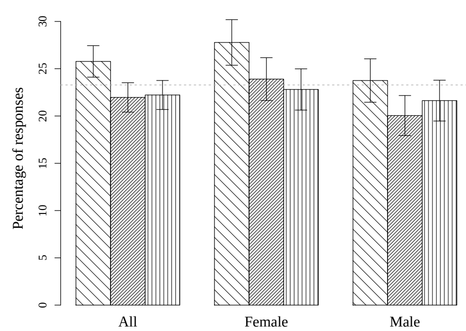
(c) Fewer Trump voters – 2nd round



(d) More Trump voters – 2nd round



(e) Fewer Trump voters – combined



(f) More Trump voters – combined

OA6 Exploratory Analysis of Attorney Race Effects

In this section, we document the results from an analysis using the combined data from both rounds of the study to generate a best guess as to the existence of a preferential treatment of White senders by White attorneys. We stress that the results in this section were not corroborated by our split sample design, and we leave it to future research to determine whether these results can be replicated.

Figure OA9 panels (a) and (b) display response rates for all inquiries sent to non-White and White attorneys in both rounds of the study. This graph confirms the finding in Section 5.2.1 in the paper that White senders receive a comparably higher share of responses from White attorneys. When we combine the data from both rounds of the study, this effect can be observed for both White female and White male senders.

In a first step, we use regression analysis to confirm that the difference in the treatment of White senders by attorneys from different groups is statistically significant. For this, we replicate the regressions in Table OA3 panel (b) columns (5) and (6), using ATT_WHITE as well as an interactions between ATT_WHITE and an indicator for White sender as additional independent variables. The interaction term between ATT_WHITE and WHITE is the main variable of interest in these regressions.

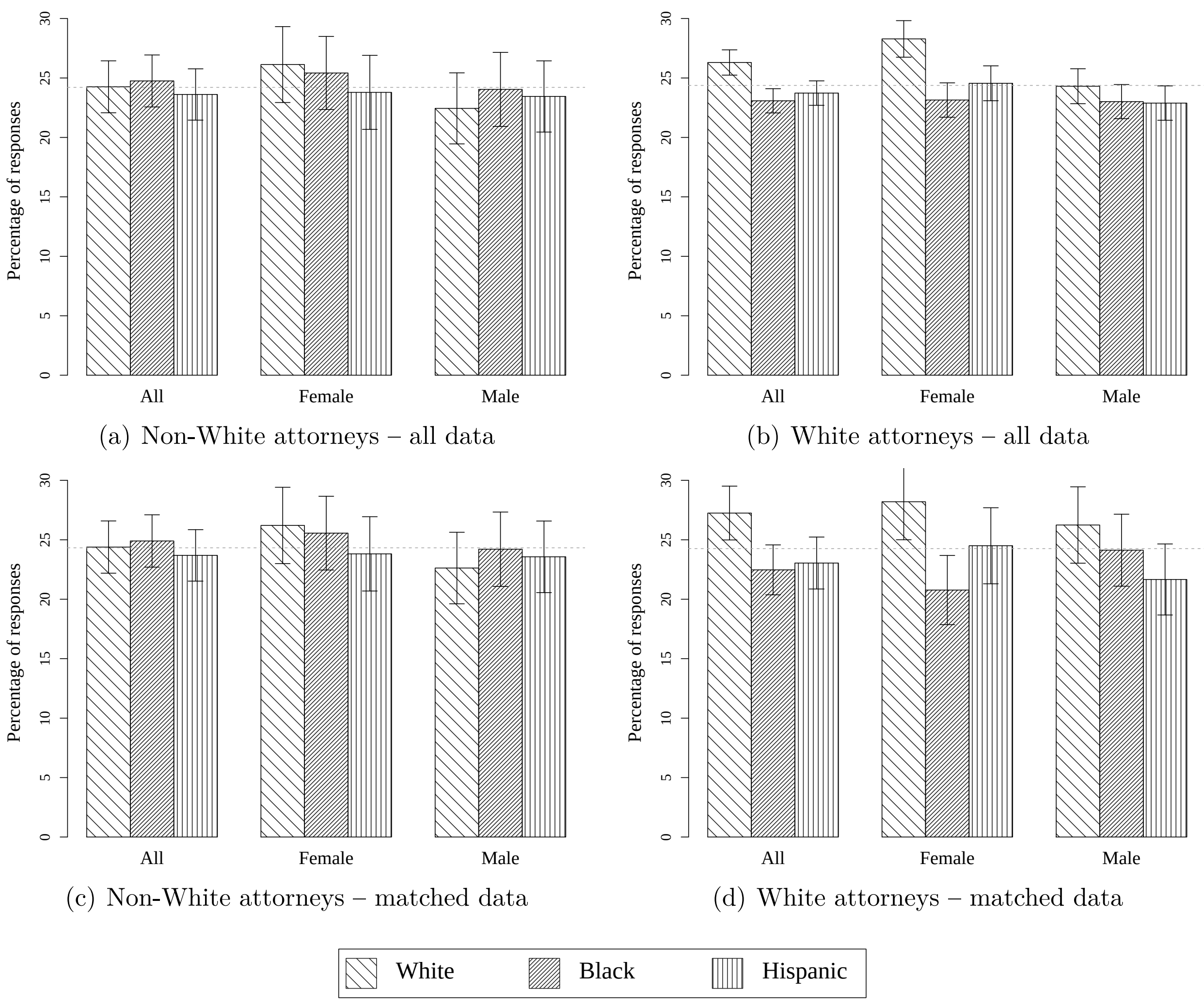
The results from these regressions are reported in columns (1) and (4) in Table OA5. The coefficients for the interaction term ATT_WHITE*WHITE is positive and (at least weakly) significant. The point estimates of .154 and .181 also confirm that this effect is substantial. For example, the regression model in column (1) predicts that, if a White sender receives an average of 30% responses from non-White attorneys in a particular state, the probability of receiving a response from a White attorney is 32.3%. By comparison, for senders from other groups, the probability of receiving a response from a White attorney is lower than the probability of receiving a response from a non-White attorney.

In Table OA6, we report the results from a series of regressions that include other attorney characteristics as well as interactions between these attorney characteristics and WHITE as additional control variables. It can be seen that the inclusion of most of these variables does not meaningfully alter the size of the estimated coefficient for ATT_WHITE*WHITE. The only attorney characteristic whose inclusion affects this estimate is attorney age. This result suggests that the estimated effect for ATT_WHITE*WHITE as reported in Table OA5 might partly capture the fact that older attorneys (who also mostly have names more common among Whites) treat White inquirers preferentially. However, as we are not postulating a causal relationship between attorney race and the treatment of different sender groups, this finding does not threaten the validity of our results.

This result persists when we use our matching strategy to isolate the effects of personal attorney characteristics from potentially confounding effects of geography. Following the steps described in Section 5.2.2, we assemble a dataset that consists of 4,444 non-White attorneys and 4,444 White attorneys who have their offices in the same geographic area. Figure OA9 panels (c) and (d) display response rates for different different sender types for these two groups. The differences in the treatment of White senders that could be observed in panels (a) and (b) persists despite the considerable drop in observations.

Columns (7) and (8) in Table OA5 report the results of regressions using the data obtained from the matching procedure. In these regressions, the point estimates for WHITE*ATT_WHITE

Figure OA9: Differences between treatment effects by predicted attorney race (combined data)



(a) Non-White attorneys – all data

(b) White attorneys – all data

(c) Non-White attorneys – matched data

(d) White attorneys – matched data

are larger than in the ones that use the unmatched data. Also, despite the reduction of the size of the dataset, the effects are significant.

Table OA5: Logit Regressions.

	Full Dataset (Combined Data)						Matched Data	
	Including Bounced Emails			Excluding Bounced Emails				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BLACK-FEMALE	-.102	-.141	-.248	-.100	-.114	-.002	-.096	-.107
	(.087)	(.270)	(.640)	(.090)	(.238)	(.672)	(.101)	(.108)
BLACK-MALE	-.123	-.162	-.263	-.113	-.126	-.005	-.040	-.069
	(.089)	(.267)	(.643)	(.092)	(.234)	(.675)	(.105)	(.108)
HISPANIC-FEMALE	-.057	-.096	-.204	-.046	-.060	-.050	-.043	-.004
	(.092)	(.270)	(.644)	(.095)	(.239)	(.676)	(.109)	(.114)
HISPANIC-MALE	-.132	-.171	-.251	-.136	-.150	-.005	-.128	-.117
	(.083)	(.267)	(.643)	(.086)	(.236)	(.676)	(.102)	(.105)
WHITE-MALE	-.208***	-.208***	-.214** *	-.223***	-.226***	-.222***	-.145+	-.157+
	(.048)	(.048)	(.050)	(.051)	(.051)	(.053)	(.080)	(.082)
ATT_WHITE	-.043	-.039	-.047	-.028	-.021	-.043	-.087	-.083
	(.051)	(.051)	(.055)	(.051)	(.051)	(.543)	(.060)	(.064)
WHITE*	.154+	.143+	.091	.181*	.168+	.113	.233*	.243*
ATT_WHITE	(.083)	(.082)	(.082)	(.088)	(.088)	(.089)	(.109)	(.118)
_Intercept	-1.27***	-1.25***	-1.11*	-.721**	-.710**	-.054	-1.06***	-.624***
	(.164)	(.232)	(.486)	(.164)	(.214)	(.077)	(.150)	(.163)
Controls include interactions between WHITE and								
- State F.E.	N	Y	Y	N	Y	Y	N	N
- Geographical characteristics	N	N	Y	N	N	Y	N	N
N	24,128	24,128	22,896	17,392	17,392	16,491	8,888	6,471

Notes. Dependent variable: binary variable indicating whether a response was received. Standard errors clustered at the county level in parentheses. Controls for different emails and state fixed effects included in all regressions (state fixed effects only in (1)–(6)). $+p < 0.1$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$.

When controlling for geographical characteristics by way of regression analysis, this result appears attenuated in some specifications. Columns (2) and (5) report results for regressions that include interactions between state fixed effects and WHITE. The inclusion of the variables controls for geographic factors that influence differential response rates at the state level. As can be seen, the estimates are almost identical to those reported in columns (1) and (4). This is true both for the point estimates and the estimated standard errors.

However, when we include additional control variables capturing socio-economic characteristics of the area in which an attorney is based (columns (3) and (6)), the effect appears substantially attenuated. This attenuation also results in p-values above all common thresholds for statistical significance. At the same time, the fact that our preferred matching strategy (which "controls" for the same variables) yields different results points to the possibility that these regressions might underestimate the true relationship between ATT_WHITE*WHITE and the probability of a response.

Note that our results are robust to using the alternative outcome variable described in Section OA2 *in lieu* of our main outcome variable. Figure OA10 and Table OA7 replicate the analyses presented and Figure OA9 and Table OA5, with substantially similar results.

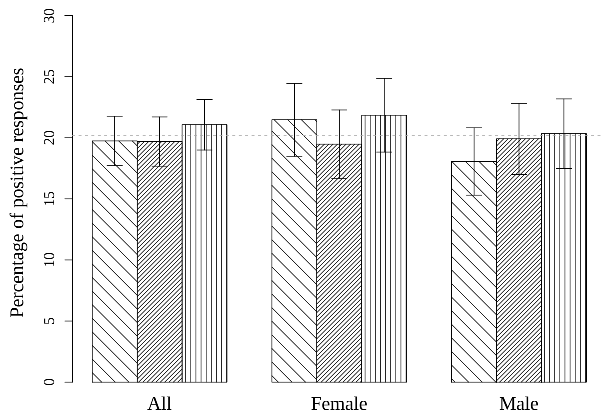
Overall, the results presented in this section lend support to the hypothesis that the observed differential treatment of senders is at least partly driven by a tendency of White attorneys to respond preferentially to inquiries by members of their own racial/ethnic group. In other words, these results suggest that the differential treatment cannot be explained by objective differences in the expected value of lawsuits brought by different client types.

Table OA6: Logit Regressions.

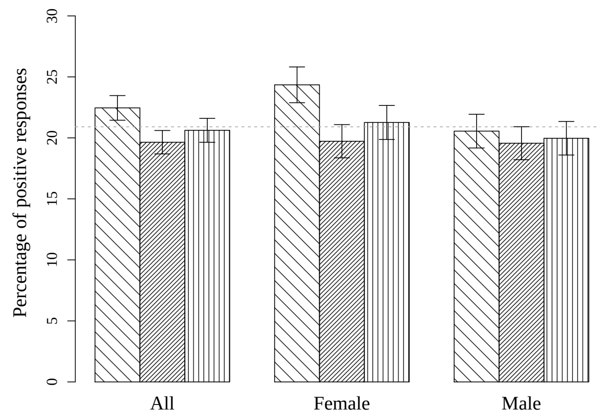
	Full Dataset (Combined Data)											
	Including Bounced Emails						Excluding Bounced Emails					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
BLACK-FEMALE	-.102	-.021	-.045	-.138	-.007	-.030	-.100	-.015	-.049	-.134	-.065	-.015
	(.087)	(.131)	(.090)	(.095)	(.134)	(.197)	(.090)	(.138)	(.092)	(.097)	(.141)	(.210)
BLACK-MALE	-.123	-.060	-.065	-.166+	-.016	-.056	-.113	-.047	-.063	-.150	-.035	-.008
	(.089)	(.129)	(.129)	(.097)	(.138)	(.201)	(.092)	(.135)	(.093)	(.100)	(.145)	(.213)
HISPANIC-FEMALE	-.057	.016	.002	-.113	.054	.049	-.046	.035	.004	-.104	.010	.072
	(.092)	(.133)	(.094)	(.101)	(.140)	(.205)	(.095)	(.139)	(.097)	(.104)	(.149)	(.220)
HISPANIC-MALE	-.132	-.062	-.074	-.180+	-.058	-.086	-.136	-.063	-.085	-.176+	-.102	-.056
	(.083)	(.126)	(.087)	(.092)	(.143)	(.205)	(.086)	(.133)	(.091)	(.093)	(.149)	(.221)
WHITE-MALE	-.208***	-.204***	-.209***	-.235***	-.257***	-.279***	-.223***	-.221***	-.229***	-.244***	-.237**	-.258**
	(.048)	(.049)	(.048)	(.051)	(.067)	(.075)	(.051)	(.051)	(.051)	(.056)	(.074)	(.084)
ATT_WHITE	-.043	-.056	-.044	-.030	-.058	-.045	-.028	-.035	-.028	-.025	-.075	-.076
	(.051)	(.050)	(.051)	(.055)	(.070)	(.075)	(.051)	(.050)	(.051)	(.055)	(.074)	(.079)
WHITE*	.154+	.145+	.156+	.112	.199+	.149	.181*	.177*	.184*	.147	.247+	.204
ATT_WHITE	(.083)	(.084)	(.083)	(.091)	(.118)	(.128)	(.088)	(.089)	(.088)	(.095)	(.126)	(.136)
ATT_MALE	-	.168***	-	-	-	.175+	-	.166**	-	-	-	.141
		(.054)				(.093)		(.057)				(.096)
WHITE*	-	.086	-	-	-	-.052	-	.086	-	-	-	.028
ATT_MALE		(.103)				(.172)		(.108)				(.181)
ATT_SOLOPRACTITIONER	-	-	-.056	-	-	-.066	-	-	.060	-	-	.025
			(.038)			(.061)			(.039)			(.064)
WHITE*	-	-	.138*	-	-	.183+	-	-	.125+	-	-	.150
ATT_SOLOPRACTITIONER			(.066)			(.100)			(.070)			(.111)
ATT_OLDER	-	-	-	-1.11***	-	-.941***	-	-	-	-.901***	-	-.804***
				(.097)		(.122)				(.101)		(.127)
WHITE*	-	-	-	.357*	-	.228	-	-	-	.308*	-	.248
ATT_OLDER				(.144)		(.178)				(.149)		(.183)
ATT_CONSERVATIVE	-	-	-	-	-.170**	-.143*	-	-	-	-	-.120*	-.101
					(.057)	(.061)					(.058)	(.063)
WHITE*	-	-	-	-	.124	.142	-	-	-	-	.004	.031
ATT_CONSERVATIVE					(.095)	(.102)					(.102)	(.108)
_Intercept	-1.27***	-1.49***	-1.31***	-1.21***	-1.44***	-1.44***	-.721**	-.958***	-.789***	-.650***	-.882**	-.893**
	(.164)	(.178)	(.164)	(.188)	(.253)	(.307)	(.164)	(.183)	(.164)	(.187)	(.254)	(.311)
N	24,128	23,677	24,128	20,533	10,841	9,390	17,392	17,069	17,392	14,530	8,079	6,898

Notes. Dependent variable: binary variable indicating whether a response was received. Standard errors clustered at the county level in parentheses. Controls for different emails and state fixed effects included in all regressions. Additional variables are defined as follows: $ATT\_WHITE = 1(ATT\_PROBMALE > .5)$; $ATT\_OLDER = 1(FIRSTADMITTED \leq 1970)$; $ATT\_CONSERVATIVE = 1(ATT\_CFSCORE > -.665)$. $+p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

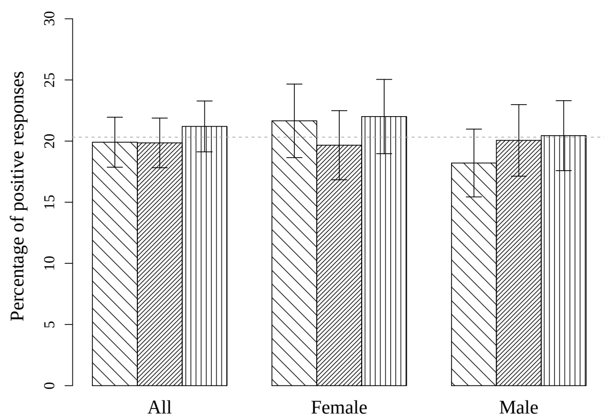
Figure OA10: Differences between treatment effects by attorney race (combined data)



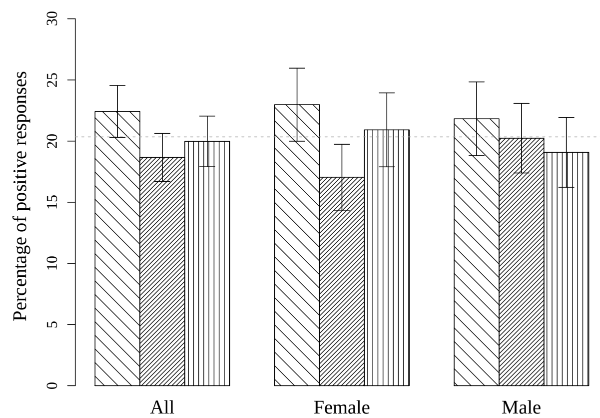
(a) Non-White attorneys – all data



(b) White attorneys – all data



(c) Non-White attorneys – matched data



(d) White attorneys – matched data

Table OA7: Logit Regressions.

	Full Dataset (Combined Data)						Matched Data	
	Including Bounced Emails			Excluding Bounced Emails				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BLACK-FEMALE	-.094	.088	-.147	-.090	-.136	.100	-.131	-.141
	(.095)	(.295)	(.642)	(.098)	(.260)	(.666)	(.110)	(.115)
BLACK-MALE	-.098	-.084	-.143	-.084	-.143	.118	-.014	-.036
	(.093)	(.291)	(.642)	(.096)	(.255)	(.664)	(.107)	(.110)
HISPANIC-FEMALE	.010	.192	-.049	-.029	-.255	.214	.064	.114
	(.099)	(.294)	(.645)	(.102)	(.259)	(.666)	(.114)	(.117)
HISPANIC-MALE	-.068	.114	-.096	-.063	-.164	.159	-.040	-.017
	(.090)	(.292)	(.644)	(.092)	(.256)	(.668)	(.105)	(.107)
WHITE-MALE	-.221***	-.221***	-.228***	-.231***	-.233***	-.228***	-.138+	-.145+
	(.052)	(.052)	(.054)	(.054)	(.054)	(.057)	(.081)	(.083)
ATT_WHITE	-.024	-.021	-.036	-.007	-.002	-.033	-.077	-.070
	(.060)	(.060)	(.064)	(.061)	(.061)	(.063)	(.071)	(.077)
WHITE*	.185+	.175+	.117	.211*	.197+	.139	.225+	.227+
ATT_WHITE	(.095)	(.094)	(.091)	(.100)	(.101)	(.097)	(.119)	(.129)
_Intercept	-1.57***	-1.69***	-1.51**	-1.05***	-1.20***	-1.13+	-1.39***	-.994***
	(.165)	(.263)	(.553)	(.165)	(.244)	(.582)	(.160)	(.171)
Controls include interactions between WHITE and								
- State F.E.	N	Y	Y	N	Y	Y	N	N
- Geographical characteristics	N	N	Y	N	N	Y	N	N
N	24,128	24,128	22,896	17,392	17,392	16,491	8,888	6,471

Notes. Dependent variable: binary variable indicating whether a positive response was received. Standard errors clustered at the county level in parentheses. Controls for different emails and state fixed effects included in all regressions (state fixed effects only in (1)–(6)).
$+p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

OA7 Additional Tables and Graphics

Table OA8: Summary statistics

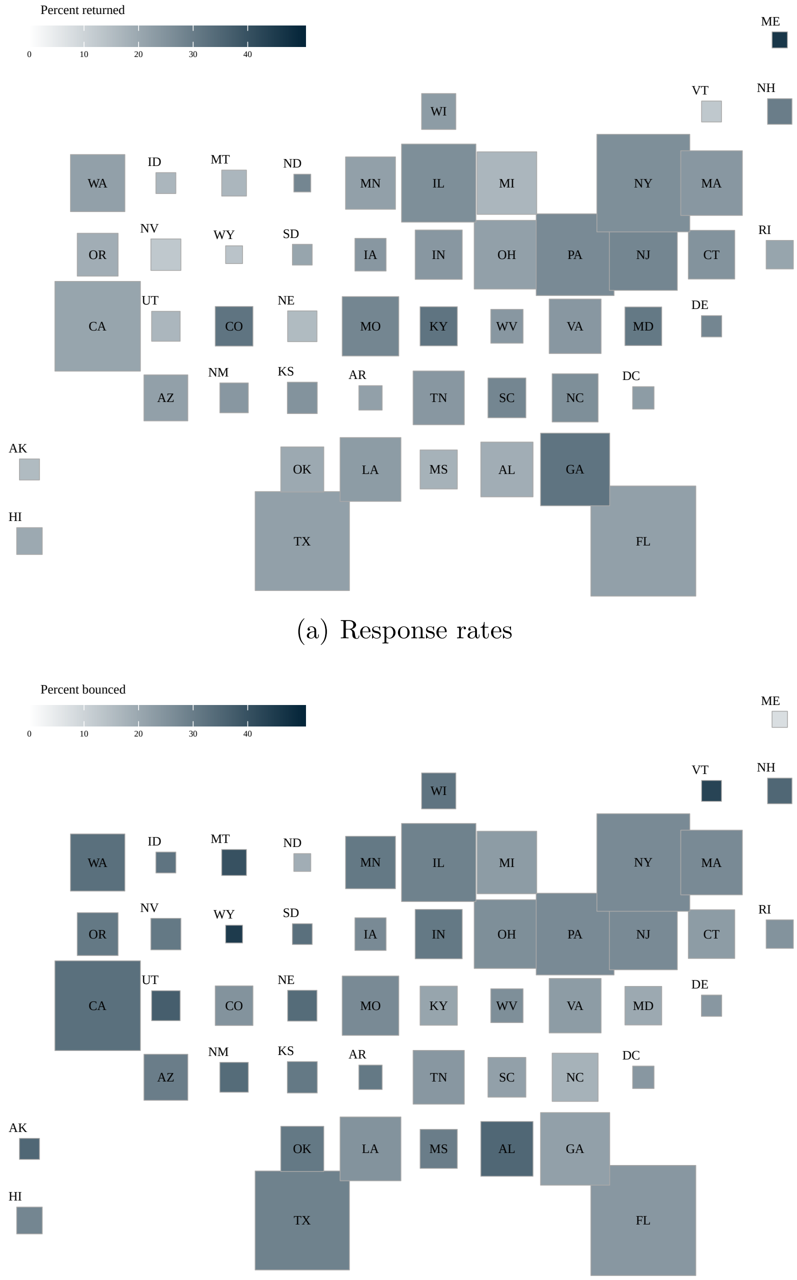
	Obs.	Mean	Std. Dev.	Min	25th Perc.	Median	75th Perc.	Max
Features of the inquiry								
BLACK	24,211	.334	.472	0	0	0	1	1
HISPANIC	24,211	.333	.471	0	0	0	1	1
WHITE	24,211	.333	.471	0	0	0	1	1
MALE	24,211	.500	.500	0	0	0	1	1
SCENARIO2	24,211	.055	.228	0	0	0	0	1
Demographic characteristics								
CENSUSBLOCK_MEDIANINCOME	23,299	64,747	39,557	5,174	36,103	56,597	83,018	250,001
CENSUSBLOCK_WHITEPOP	24,154	.731	.206	0	.611	.778	.890	1
CENSUSBLOCK_UNEMPLOYED	24,138	.066	.064	0	.024	.051	.088	1
COUNTY_MEDIANINCOME	24,128	58,413	15,153	18,972	48,104	55,277	66,529	125,672
COUNTY_WHITEPOP	24,128	.605	.200	.008	.455	.622	.764	.985
COUNTY_UNEMPLOYED	24,128	.074	.019	.006	.061	.070	.085	.211
COUNTY_POVERTYLEVEL	23,842	.127	.058	.022	.085	.123	.157	.600
COUNTY_POVERTYDIFFERENCES	23,842	.045	.042	-.016	.013	.035	.065	.391
COUNTY_FOREIGNBORN	24,128	.137	.105	.002	.056	.107	.211	.522
COUNTY_TRUMPVOTER	24,128	.419	.163	.041	.314	.416	.528	.895
Attorney characteristics								
ATT_FIRSTADMITTED	20,606	1985	12.0	1922	1977	1985	1993	2018
ATT_SOLOPRACTITIONER	24,211	.410	.492	0	0	0	1	1
ATT_PROBMALE	23,760	.847	.346	0	.992	.996	.997	1
ATT_PROBWHITE	24,211	.810	.239	0	.740	.917	.972	1
ATT_CFSCORE	10,891	-.424	.810	-3.81	-1.03	-.655	.129	2.698
ATT_WHITE	24,211	.815	.388	0	1	1	1	1

Table OA9: Logit Regressions.

	Full Dataset (First Round Data)						Matched Data	
	Including Bounced Emails			Excluding Bounced Emails				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WHITE-MALE	.060	-.215	-.315	.109	-.461	-2.01	.048	.066
	(.149)	(.351)	(1.14)	(.156)	(.358)	(1.34)	(.159)	(.172)
MINORITY1-FEMALE	-.034	-.310	-.406	-.018	-.591	-2.13	-.064	-.023
	(.155)	(.352)	(1.14)	(.164)	(.359)	(1.34)	(.151)	(.164)
MINORITY1-MALE	.060	-.216	-.335	.063	-.508	-2.073	.073	.036
	(.155)	(.354)	(1.14)	(.164)	(.360)	(1.34)	(.186)	(.195)
MINORITY2-FEMALE	.047	-.229	-.355	.067	-.503	-2.072	-.005	.081
	(.161)	(.356)	(1.14)	(.169)	(.361)	(1.35)	(.171)	(.180)
MINORITY2-MALE	.054	-.222	-.312	.040	-.531	-2.058	.030	.064
	(.145)	(.348)	(1.14)	(.147)	(.352)	(1.34)	(.148)	(.153)
ATT_WHITE	-.036	-.036	-.067	-.005	-.004	-.053	-.042	-.039
	(.060)	(.060)	(.067)	(.065)	(.066)	(.073)	(.075)	(.083)
WHITE-FEMALE*	.371*	.378*	.290+	.381*	.390*	.305+	.364+	.406+
ATT_WHITE	(.154)	(.155)	(.156)	(.165)	(.173)	(.176)	(.193)	(.212)
_Intercept	-1.43***	-1.20**	-.994	-.793***	-.321	1.26	-1.29***	-.825***
	(.231)	(.372)	(1.06)	(.223)	(.357)	(1.25)	(.198)	(.222)
Controls include interactions between WHITE-FEMALE and								
- State F.E.	N	Y	Y	N	Y	Y	N	N
- Geographical characteristics	N	N	Y	N	N	Y	N	N
N	11,278	11,272	10,696	7,984	7,983	7,572	4,212	3,027

Notes. Dependent variable: binary variable indicating whether a response was received. Standard errors clustered at the county level in parentheses. Controls for different emails and state fixed effects included in all regressions. $+p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure OA11: Response rates and rates of bounced emails by states



(a) Response rates

(b) Bounce rates

Figure OA12: Differences between treatment effects by attorney gender

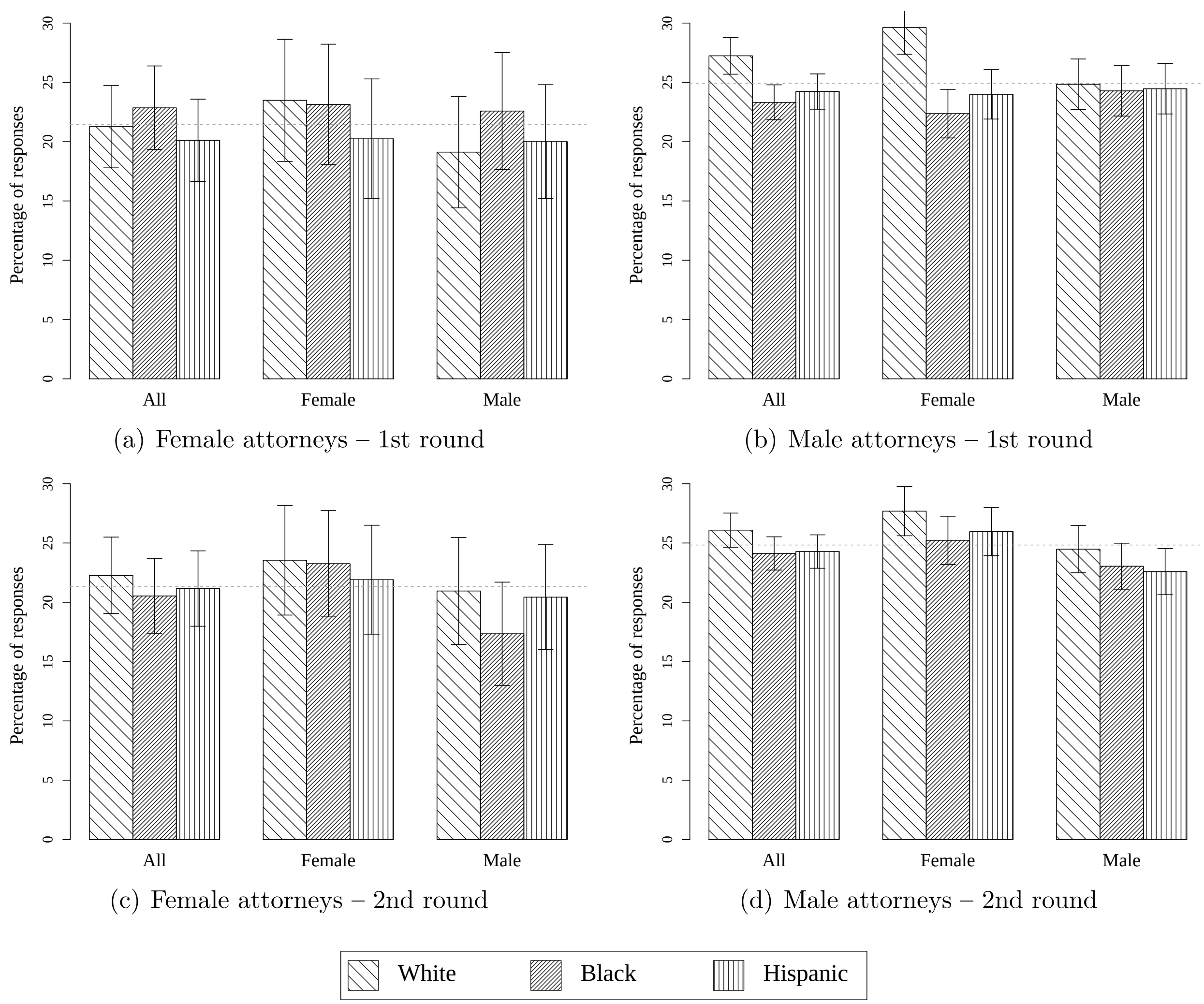
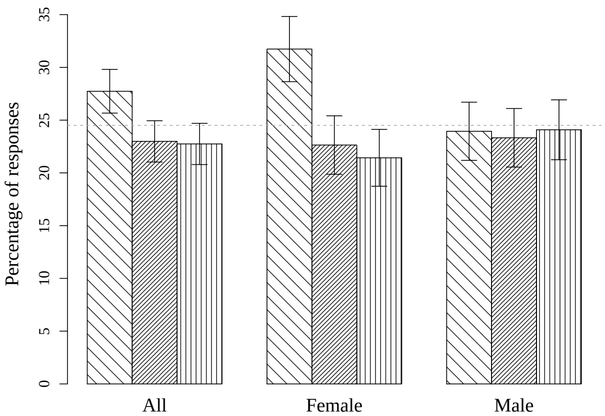
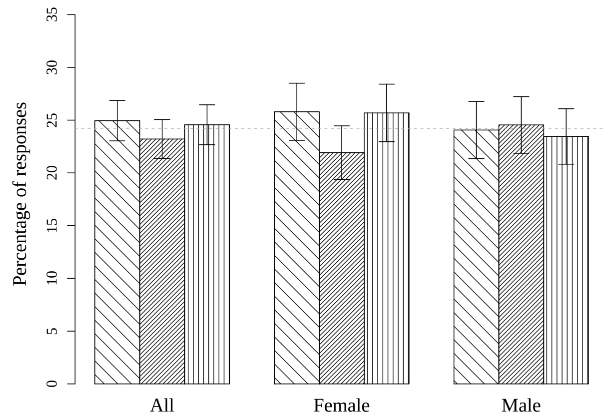


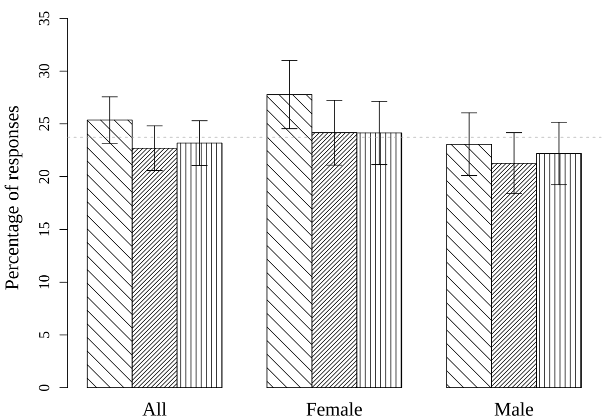
Figure OA13: Differences between treatment effects by firm size



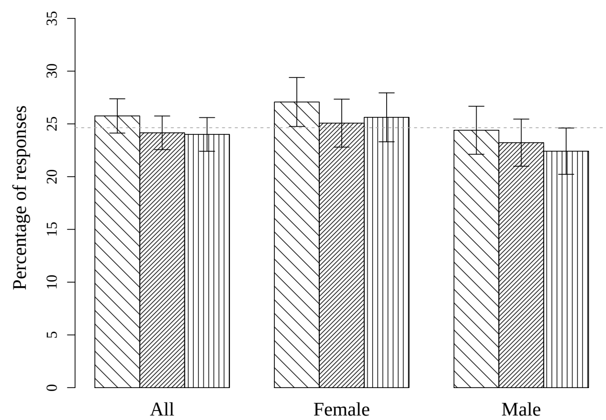
(a) Solo practitioners – 1st round



(b) Other attorneys – 1st round

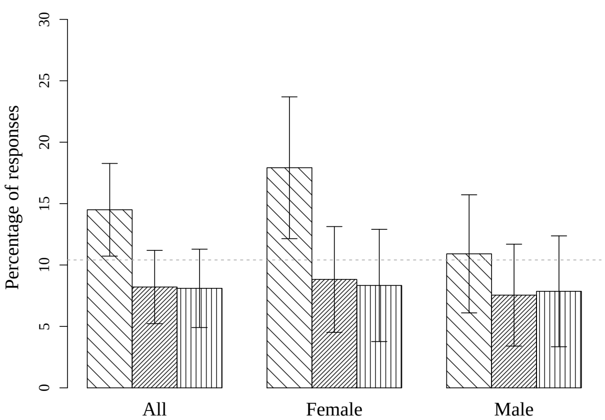


(c) Solo practitioners – 2nd round

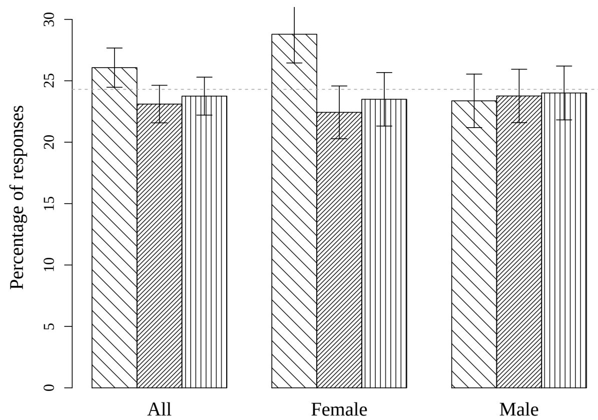


(d) Other attorneys – 2nd round

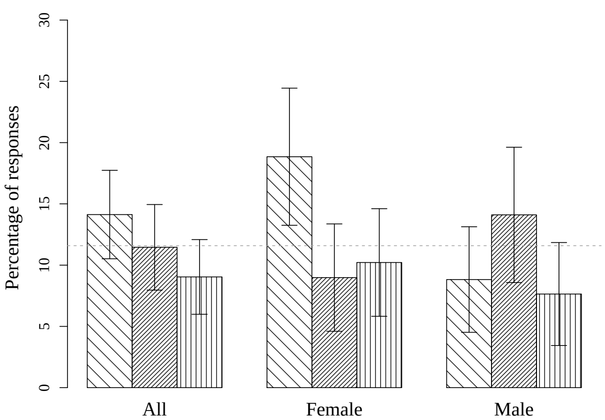
Figure OA14: Differences between treatment effects by attorney age



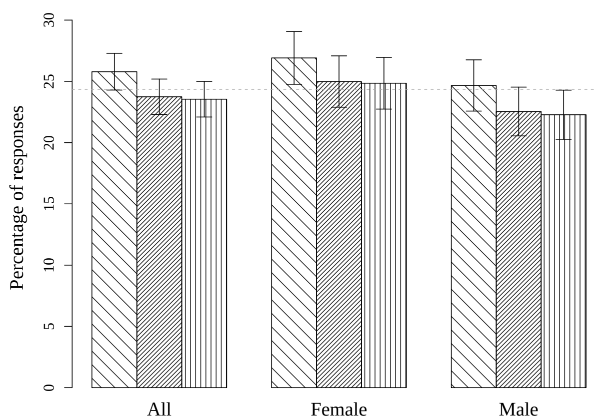
(a) Solo practitioners – 1st round



(b) Other attorneys – 1st round

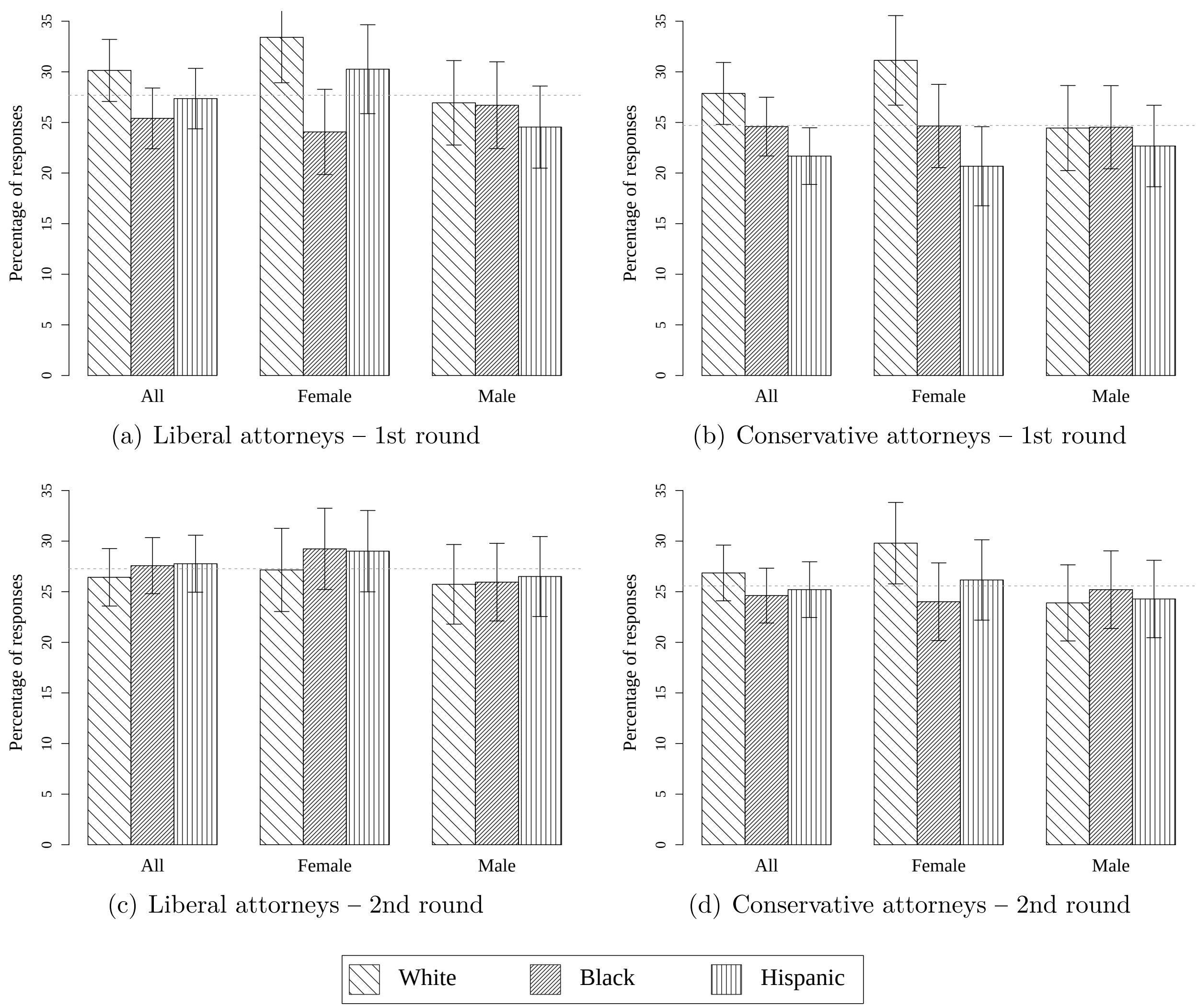


(c) Solo practitioners – 2nd round



(d) Other attorneys – 2nd round

Figure OA15: Differences between treatment effects by attorney ideology



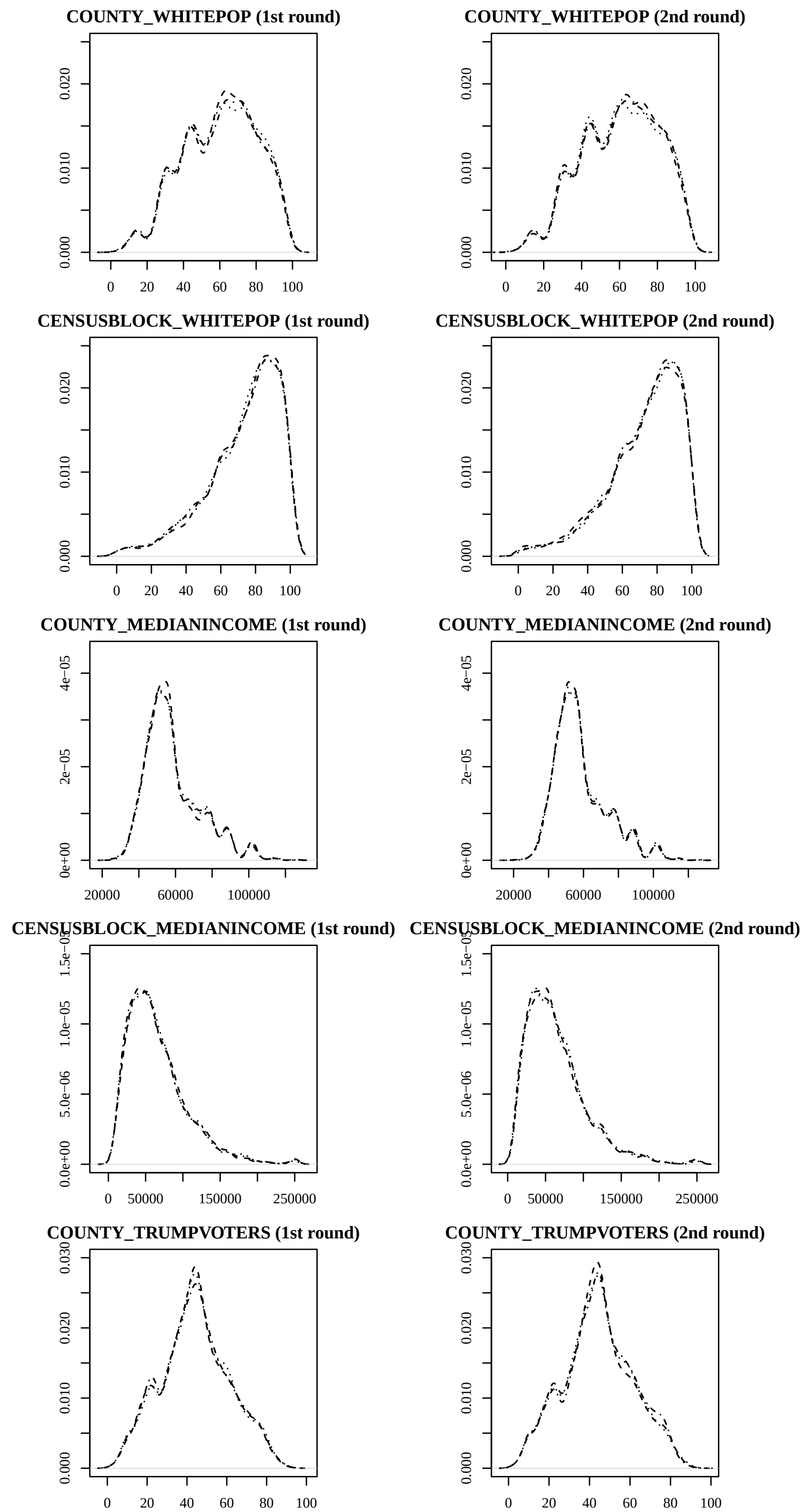
(a) Liberal attorneys – 1st round

(b) Conservative attorneys – 1st round

(c) Liberal attorneys – 2nd round

(d) Conservative attorneys – 2nd round

Figure OA16: Covariate distributions pre and post matching



OA8 Preregistration Statement

CONFIDENTIAL - FOR PEER-REVIEW ONLY

Are Lawyers' Case Selection Decisions Biased? (#66602)

Created: 05/23/2021 08:44 AM (PT)
Shared: 05/23/2021 09:05 AM (PT)

1) Have any data been collected for this study already?
It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid pre-registration nevertheless.

2) What's the main question being asked or hypothesis being tested in this study?
This is a field experiment (conducted via email) exploring how demographic information (as encoded in the names of potential clients) affects how attorneys respond to initial inquiries by clients. Our main research question is the following: (1) Are attorneys more likely to respond to inquiries from (perceived) White potential clients than to inquiries from (perceived) Black or Hispanic clients? In addition, we explore two research questions related to the mechanism behind this effect: (2) Is this effect mostly driven by a favorable treatment of White female clients, as opposed to White male clients? (3) Are attorneys who are statistically likely to be White (based on their names) more likely to treat White (female) clients more favorably compared to other clients than other attorneys?

3) Describe the key dependent variable(s) specifying how they will be measured.
Dummy variable indicating whether we receive an email in response to our outreach that was not flagged (by an algorithm written by us in python) as an error message

4) How many and which conditions will participants be assigned to?
Six conditions: WHITE-FEMALE, WHITE-MALE, MINORITY1-FEMALE, MINORITY1-MALE, MINORITY2-FEMALE, MINORITY2-MALE (MINORITY1 indicates a sender name more common among Black persons, MINORITY2 a sender name more common among Hispanic persons).

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
Main research question: Fisher's Exact Test comparing the response rates for all WHITE senders with the response rates for all MINORITY1 and MINORITY2 senders.

Research question (2): Fisher's Exact Test comparing the response rates for WHITE-FEMALE senders with the response rates for all MINORITY1 and MINORITY2 senders and Fisher's Exact Test comparing the response rates for WHITE-MALE senders with the response rates for all MINORITY1 and MINORITY2 senders (based on our exploratory study, we expect the second test not to yield a significant result).

Research question (3): Logit regression including, as independent variables: (i) dummies for treatment group, (ii) dummy for whether attorney has name that is more common among White people (ATT_WHITE), (iii) interaction between ATT_WHITE and WHITE-FEMALE. Our main variable of interest is the interaction term described under (iii).

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
n/a (no observations will be excluded in the analysis of the second round data)

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
13,044 email inquiries / observations (see also below under 8)

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
We implement a split-sample validation method. We constructed our two datasets during two experimental rounds that used roughly half of the email addresses at our disposal in each round.

The first round of the study ran between early June 2019 and mid-July 2019. The first round data (which includes 11,317 observations) was used in an exploratory analysis to construct and test our statistical models.

In a second round (running between mid-August 2019 and late September 2019), we contacted 13,044 lawyers. Data gathered in the second round has to date not been analyzed in any way. This dataset is the dataset for which we preregister the study.

Version of AsPredicted Questions: 2.00

Available at https://aspredicted.org/blind.php?x=8hn5ps