Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

An Experiment in Hiring Discrimination via Online Social Networks

Alessandro Acquisti, Christina Fong

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 66, No. 3, March 2020, pp. 1005–1024
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

An Experiment in Hiring Discrimination via Online Social Networks

Alessandro Acquisti,[a] **Christina Fong**[b]

[a] Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213; [b] Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Contact: acquisti@andrew.cmu.edu, https://orcid.org/0000-0001-6582-6178 (AA); fong2@andrew.cmu.edu, https://orcid.org/0000-0003-2703-422X (CF)

Received: March 11, 2016
Revised: August 19, 2017; June 7, 2018
Accepted: November 6, 2018
Published Online in Articles in Advance: November 1, 2019

https://doi.org/10.1287/mnsc.2018.3269

Abstract. We investigate whether personal information posted by job candidates on social media sites is sought and used by prospective U.S. employers. We create profiles for job candidates on popular social networks, manipulating information protected under U.S. laws, and submit job applications on their behalf to more than 4,000 employers. We estimate employer search activity and bias in interview callbacks. We find evidence of employers searching online for the candidates. At the national level, we find no significant difference in the callback rates for a Muslim versus a Christian candidate, or for a gay versus a straight candidate. However, employers in Republican areas exhibit significant bias against the Muslim candidate relative to the Christian candidate. This bias is significantly larger than the bias in Democratic areas. The results on callback bias are robust to using state- and county-level data, to controlling for firm, job, and geographical characteristics, to including additional interaction effects in the empirical specification, and to several estimation strategies. The results suggest that the online disclosure of certain personal traits can influence the hiring decisions of some U.S. employers, but the likelihood of hiring discrimination via online searches varies across employers.

History: Accepted by Chris Forman, information systems.
Funding: This study is based on work partly supported by the National Science Foundation [Grant CNS-1012763]; Carnegie Mellon CyLab; TRUST (Team for Research in Ubiquitous Secure Technology), which received support from the National Science Foundation [Grant CCF-0424422]; and the IWT SBO Project on Security and Privacy for Online Social Networks (SPION).
Supplemental Material: The e-companion is available at https://doi.org/10.1287/mnsc.2018.3269.

Keywords: online social networks • information systems • privacy • field experiments • economics • discrimination

1. Introduction

The rise of internet and social media services such as online social networks has created new channels through which employers and job candidates can find information about each other. Those channels can facilitate and improve the matching between firms and workers. However, job seekers also reveal, online, information that would not be easily discovered during the interview process, and which may be even illegal for employers to request or use in the hiring process. Thus, although new online tools can facilitate labor market matching, they may also create a new arena for labor market discrimination.

To date, no field experiments have demonstrated how social media profiles affect the hiring behavior of U.S. firms. In surveys, employers admit to using various online services to research job candidates.[1] However, the Equal Employment Opportunity Commission (EEOC) has cautioned firms about risks associated with searching online for protected characteristics[2] and thus may have dissuaded some firms from using social media in the hiring process. Some states have even drafted bills limiting employers' ability to access candidates' online information.[3] Thus, whether hiring bias results from personal information posted online remains an open question.

Furthermore, in surveys, employers claim to use social media merely to seek job-relevant information about candidates. However, much more private information can be gleaned from the online presence of prospective hires. On a social media profile, a status update can reveal a candidate's place of worship, a comment can suggest sexual orientation, and a personal photo can reveal ethnic origins. In a country known for its cultural assimilation of immigrants (Vigdor 2008), private differences that traditionally may have been scrubbed out for work or education might now be more visible online. Whether U.S. employers react to such online personal information, rather than merely to the online professional information they may seek, is not known.

We present a randomized field experiment testing the joint hypothesis that firms (i) search online for information about job applicants, and (ii) change their hiring behavior according to manipulated online personal information. The experiment relies on a methodology consisting of the creation and careful design of online presences of fictional individuals. We design social media profiles for four job candidates—a Muslim versus a Christian candidate, and a gay versus a straight candidate—to manipulate personal information that may be hard to discern from résumés and interviews and that may be protected under either federal or some state laws (henceforth referred to as "protected information").[4] We manipulate the candidates' personal information exclusively via their online profiles, using material revealed online by actual members of popular social networking sites and job seeking sites. Candidates' professional background and résumés are kept constant across conditions.

After vetting the realism and quality of candidates' online profiles in a randomized online pilot experiment (henceforth, the online pilot), we submit résumés and cover letters on behalf of those four candidates to more than 4,000 U.S. job openings, with a single application sent to each employer (henceforth, the field experiment). The résumés and letters contain no references or links to the candidates' manipulated personal information. Therefore, to be treated by our manipulation of online profiles, the employer must independently choose to search online for information about the candidate using the name indicated on the submitted résumé. The main dependent variable in the field experiment is the number of interview invitations each candidate receives (i.e., callbacks). We compare the callback rate for the Christian candidate with that for the Muslim candidate, and the callback rate for the straight candidate with that for the gay candidate. We control for independent variables used in related literature (e.g., Bertrand and Mullainathan 2004, Tilcsik 2011), including firm characteristics, job characteristics, and geographical characteristics of the job's location. We also use Google and LinkedIn data to estimate the frequency with which employers search the candidates online. Following survey evidence of more negative attitudes to both Muslims and gay people among Republican survey respondents (Arab American Institute 2012; Pew Research Center 2012, 2013) and results from previous experimental work (e.g., Tilcsik 2011) showing geographical variation in hiring bias, as well as patterns found in our own online pilot experiment, we test for stronger bias in firms located in states and counties that have a high fraction of Republican voters.

Nationwide, we detect no significant difference in callback rates for the gay candidate compared with the straight candidate or for the Muslim candidate compared with the Christian candidate. However, we find significant bias against the Muslim candidate relative to the Christian candidate in areas with strong Republican party support. The bias in these Republican areas is significantly larger than the bias in areas with strong Democratic party support. In states with a high fraction of Republican voters, the callback rates for the Muslim and Christian candidates are 2.27% and 17.31%, respectively. In states with a high fraction of Democratic voters, the callback rates for the Muslim and Christian candidates are 11.74% and 11.61%, respectively. Similarly, in strongly Republican counties, the callback rates for the Muslim and Christian candidates are 6.25% and 22.58%, respectively. In other words, in those counties, Muslim candidates are nearly four times less likely to be called back for interviews than equally qualified Christian candidates. In strongly Democratic counties, the callback rates for the Muslim and Christian candidates are, respectively, 12.37% and 12.13%. Note that our results focus on aggregate voting behavior in geographical *areas* (such as counties and/or states) with multiple, possibly uncontrolled, sources of cross-sectional heterogeneity—not on individual voters. Our experimental design does not identify the individual making the callback decision, nor does it allow us to observe individual employers' traits (such as his or her political identification). Thus, we do not make conclusions regarding causality from voting patterns to discrimination. That noted, the results are consistent with those of the online pilot, in which we find significant bias against the Muslim candidate, relative to the Christian candidate, among subjects with hiring experience who self-identified as Republican and Christian. Furthermore, the significance of the political interaction is robust to the inclusion of firm characteristics and a host of additional controls (such as how rural is the area, median incomes, and the percentage of the local population that is Muslim), to different estimation strategies, and to various robustness tests, including different categorizations of Republican, mixed, and Democratic states or counties based on electoral results and Gallup Organization surveys.

In short, at the nation-wide level, the information we manipulated does not seem to significantly affect employers' behaviors. However, the evolution of social media usage among job seekers and job providers does make available information that used to be protected under the law, and this information seems to influence the decision of *some* employers (with probability varying across employers in manners consistent with previous research on bias and discrimination). Such findings make a number of contributions to the literature. First, the findings highlight an emerging tension between modern information systems and institutional regulation written for a pre-internet world. The latter aims to preclude certain information from being used in the hiring process; the former can effectively bypass legislation by allowing individuals to make their information openly available to others online. A similar tension is being studied in the growing empirical (Miller and Tucker 2009, Goldfarb and Tucker 2011) and theoretical (Acquisti et al. 2016) information systems literature on privacy and its economic aspects (Stigler 1980, Posner 1981). Second, the paper tests the viability of a methodology—the manipulation of internet profiles—for field experiments on digital discrimination in fields that include, and extend beyond, labor market research. One possible advantage of this method, which we exploit here, is its suitability for investigating discrimination associated with traits that may be protected under a country's laws or that may be difficult to realistically manipulate in résumé-only or audit studies (because most job candidates may refrain from revealing certain types of personal information on résumés). Some of the disadvantages we confronted in deploying this methodology, however, include the significant efforts required to acquire appropriate sample sizes, the looseness of experimental treatments (relative to tightly controlled laboratory experiments), and the trade-offs between ecologically valid setup and challenges associated with precisely testing causal mechanisms. Third and finally, this paper illustrates a wrinkle in the literature investigating the role of information in economic outcomes. Economists have long been interested in the role of information (Stigler 1962) and signaling (Spence 1973) in job market matching. Recent work has highlighted how hiring mechanisms that *reduce* candidates' information available to employers (for example, by making the candidate anonymous) may *increase* interview opportunities for certain categories of applicants (Goldin and Rouse 2000, Aslund and Skans 2012) and may raise social welfare (Taylor and Yildirim 2011). In this paper, conversely, we find that internet and social media platforms can affect interview opportunities for some categories of applicants at the expense or benefit of others, by making more information about candidates available to employers. The extent to which this new information channel will improve labor search efficiency (Kroft and Pope 2014) and reduce labor market frictions (by allowing better matching of employers and employees) or will in fact lead to more discrimination is likely to be an issue of increasing public policy relevance.

2. Background

2.1. Public Disclosures, Social Media, and Social Networking Sites

The rise of social media has both fueled and been fueled by an arguably unprecedented amount of public sharing of personal information. The shared information frequently includes disclosures and revelations of a personal, and sometimes surprisingly candid, nature. In certain cases, personal information is revealed through fully identified profiles.[5] Other times, users provide personal information under pseudonyms that may still be identified.[6] Some social media users take advantage of privacy settings to manage and restrict their online audiences (Stutzman et al. 2012). Others think they do but actually fail to protect their information.[7]

Employers can access shared personal information in many ways. Some job candidates make their online profiles (on social media or blogging platforms) openly accessible to strangers.[8] Others are more selective, but sensitive information such as religious affiliation, sexual orientation, or family status may still be indirectly inferable from seemingly more mundane data.[9] Finally, some employers engage in social engineering (such as using "friends of friends" connections to view a candidate's profile) or even ask for the candidates' passwords to access their profiles. Although the legal consensus seems to suggest that seeking information online about job candidates (or employees) may not violate U.S. law, such searches do raise privacy and legal issues, such as the discriminatory behavior that may result from discovering protected information (Sanders 2011; Sprague 2011, 2012).

Above-cited surveys suggest that U.S. employers have been using social media to screen prospective job candidates. However, no prior controlled experiment has measured the frequency of firms' usage of online profiles in hiring decisions and how profile information actually affects those decisions. The study that comes closest to our experiment is by Manant et al. (2014), who investigate the role of social media in the French labor market. Other than differences in the country of study (United States versus France) and sample size (approximately 4,000 versus approximately 800 employers), Manant et al. (2014) focus on a different research question than our study: our paper focuses on testing the joint hypothesis that firms search online for information about job applicants and change their hiring activities according to the personal information they find, whereas Manant et al. (2014) focus on investigating the impact of different search costs for finding candidates' information via social media. In addition, Bohnert and Ross (2010) use a survey-based experiment to investigate how the content of social networking profiles can influence evaluations of job candidates. Garg and Telang (2017) analyze how job applicants can use LinkedIn to find connections that may lead to a job offer. Kluemper et al. (2012) assess the relationship between the content of a person's Facebook profile and her future job performance.

2.2. Discrimination and Résumé Studies

Following Bertrand and Mullainathan (2004), experiments using written applications to real employers have found evidence of discrimination against people with various traits. Pertinent prior literature has found discrimination against job candidates with Muslim rather than Swedish names in Sweden (Carlsson and Rooth 2007), candidates who openly signal Muslim beliefs (relative to other religious beliefs) on their résumés in New England (Wright et al. 2013), and candidates who explicitly or implicitly signal same-sex sexual orientation (Weichselbaumer 2003, Ahmed and Hammarstedt 2009, Tilcsik 2011)[10] but no evidence of discrimination against Muslims (and little systematic discrimination by caste in new economy sectors) in India (Banerjee et al. 2009).

One crucial difference between existing résumé studies and our approach is that we focus on employers' volitional choices to search online for candidates' information. A second important difference is that we focus on candidates revealing information in personal online social profiles, rather than volunteering personal traits in a professional context. This approach facilitates the investigation of discrimination based on protected information that candidates do not frequently provide in their résumés or during interviews.

It is important to note the difference between our question and more frequently asked questions in the economics literature on discrimination: does discrimination exist, to what extent, and in what form? That literature has developed sophisticated methods for testing whether discrimination is actually present and for empirically separating taste-based from statistical discrimination (Mobius and Rosenblat 2006, Neumark 2012, Ewens et al. 2014). We do not attempt to separate bias from search, let alone separate taste-based discrimination from statistical discrimination or other forms of bias. We instead focus on the more basic question of whether social media profiles can impact hiring via the combined effects of online search and bias.

3. Design Considerations

We test for the existence of hiring discrimination that stems jointly from two, possibly dependent, actions: first, each employer decides whether to search online for a candidate's information; and, second, each employer who finds the candidate's profile is unknowingly treated to randomized personal information and chooses whether to interview the candidate. Using a between-subject design, we test for an effect of randomized social network profiles on interview invitations (callbacks) following job applications. Employers' responses are included in the analysis whether they search online for the candidate. Thus, for a randomized candidate profile to have an effect, the employer would first have to search the candidate's name online and then react to the social network profile. Given that our design tests jointly for online searching and for discrimination on the basis of manipulated online profiles, a null effect in the field experiment may signal low levels of either of these behaviors.

Furthermore, the experiment could fail to reject the null hypothesis in a number of additional ways. First, employers may search the candidates online but fail to find our profiles. Our design addresses this concern by using name-selection criteria that produce high-ranked search results, and by checking, for the duration of the experiment, and across a variety of platforms, that candidate name searches produce the desired results (see Section 4.2). Another possibility is that the manipulations may fail to signal the traits we intended to manipulate. To address this possibility, before conducting the field experiment, we use an online pilot to test whether we successfully manipulated beliefs about the candidates' personal traits. A further concern is that the candidates may be either under- or over-qualified, creating floor or ceiling effects in callback rates. We address this possibility by designing the résumés and professional backgrounds of the candidates in concert with human resources professionals and testing their quality during the online pilot. An additional possibility is that employers may not pursue the candidates because of suspicions that they are fictional. Consequently, we design résumés and profiles using information that real individuals (including members of job search sites and members of the same social network the candidates were purported to belong to) had also made publicly available on their profiles or on job search sites. Furthermore, we chose privacy and visibility settings that were common among the profiles of real members. The online pilot tests the success of our efforts at creating realistic profiles by asking questions aimed at discerning whether online subjects had doubts about the veracity of the candidates (see Section 5.1). Finally, a null effect may be consistent with scenarios in which employers may search, but at a later stage in the hiring process (for instance, after interviews with the candidates). Our design would not capture this behavior and would not be able to distinguish this scenario from one with little or no discrimination.

4. Experimental Design

Our experimental design manipulates personal information that is unlikely to be volunteered by candidates in résumés and that may be risky under federal or state law for employers to inquire about during interviews, but which may be obtained online. Building on studies by Weichselbaumer (2003), Carlsson and Rooth (2007), and Tilcsik (2011), we focus on sexual orientation and religious affiliation.[11]

We implemented a between-subjects design with four treatment conditions consisting of social media profiles for a single job candidate. We manipulated religious affiliation (a Christian versus a Muslim male) and sexual orientation (a gay versus a straight male).[12] Thus, the experimental conditions represent a range of traits that include federally protected information (religious affiliation)[13] and information protected only in certain states (sexual orientation).[14] Although these traits enjoy different levels of protection under U.S. law, job candidates can signal this information online both in an explicit manner (e.g., self-descriptions on a profile indicating one's sexual orientation) and in an implicit manner (e.g., a primary profile photo suggesting the individual's religious affiliation).

For each of the four candidates, we designed (i) a résumé; (ii) a profile on LinkedIn (a popular online professional network commonly used by human resource professionals and job seekers, henceforth referred to as "PN," as in "Professional Network"); and (iii) a profile on Facebook (a popular social networking site commonly used for socializing and communicating, henceforth referred to as "SN," as in "Social Network").[15] We designed candidates' résumés and PN profiles to show identical information across conditions, except for the candidates' names. In contrast, the candidates' SN profiles manipulated information about the candidates' personal traits using various profile fields, including those that some employers admit to using during the hiring process in semistructured interviews (De La Llama et al. 2012). Each applicant's name corresponds to an experimental condition; the names link together a candidate's résumé, PN profile, and SN profile. Because a single name links all of these materials, we can submit just one application to each employer in our sample.[16] Submitting multiple candidates' applications to the same employer would increase the risk of an employer detecting social media profiles that have identical photos and generic information yet come from candidates with different names and condition-specific information.

4.1. Design Approach and Priorities

We designed profiles that (i) were realistic and representative of the population of SN users with the traits we were manipulating, and (ii) held constant direct signals of productivity outside of beliefs stemming from the trait itself. As described in the rest of this section, we populated the résumés and online profiles using existing information posted online by actual SN members demographically similar to the candidates, and by individuals who listed their résumés on job searching sites. The SN profiles are the vehicles for the manipulations, and responses to these profiles comprise our study's core. The remainder of this section discusses our design in more detail.

4.2. Candidates' Information

4.2.1. Names. We designed first and last names representing a U.S. male for each of the conditions used in the experiments. We chose first names common among U.S. males in the same age group as our candidates and assigned identical first names to candidates in matching pairwise conditions (the gay and straight candidates were both named Mike, and the Muslim and Christian candidates were both named Adam).[17] We then designed last names by altering letters of existing but similarly low-frequency U.S. last names. The last names were chosen to have the same number of syllables, identical lexical stress, and similar sounds across matching pairwise conditions (see Table A.1 in the e-companion for more detail). A potential concern is that Christian and Muslim names often differ. However, roughly 35% of the approximately 2.35 million American Muslims are, like our candidate, U.S.-born, and many have Anglo-Saxon names.[18]

We iteratively tested several combinations of first and last names until we found a subset that satisfied three criteria. Criterion 1 was that the exact first and last name combination would be unique on SN and PN: no other profile with that name should exist on the network. Criterion 2 was that SN and PN profiles designed under each name would appear among those names' top search results conducted with the most popular search engines (Google, Bing, and Yahoo), as well as searches conducted from within the SN and the PN social networks. We continuously monitored the fulfillment of these criteria for the experiment's duration.[19] Criterion 3 was the most critical: names alone should not elicit statistically significant differences in perceptions of the manipulated traits. We conducted two checks of this criterion. First, we tested that names, by themselves, did not influence perceptions of the traits manipulated in the profiles. We recruited 496 subjects from Amazon Mechanical Turk (MTurk) and presented each with one first and last name combination, randomly chosen from the list of possible names.[20] Each subject then responded to a questionnaire on associations between that name and the traits manipulated in field experiment. We selected names for the field experiment that did not elicit statistically significant differences in perceptions of traits (Table A.5 in the e-companion) and then randomly assigned them to candidates. The second check of criterion 3 combined names, résumés, and professional online profiles and checked that names and professional information did not elicit differential propensity to invite a candidate to an interview (see Section 5.1).

4.2.2. Email Addresses and Telephone Numbers. For inclusion in the résumés and cover letters sent to companies, we designed email addresses using a consistent format across the profiles and registered a telephone number for the applicants. We used automated message recordings for the number's voice messaging service. This message was identical across conditions and recorded with the voice messaging service's standard greeting. We also used the same postal address for all candidates, corresponding to a residential area of a midsize North American city.

4.2.3. Résumés. The résumés contained information depicting a professional and currently employed candidate. Each résumé contained the candidate's contact information, educational background, and work experience, as well as technical skills, certifications, and activities. The résumés were held constant across conditions, except for the names of the applicants and their email addresses.[21] Hence, professional and educational backgrounds did not vary across experimental conditions, thus holding constant the candidates' job market competitiveness. The information included in the résumé was modeled after résumés found on websites such as Monster.com and Career .com for job seekers demographically similar (in terms of age and educational background) to the candidates. The résumé represented a candidate with a bachelor's degree in computer science and a master's degree in information systems.[22] Creating a single SN and PN profile per candidate constrained the types of jobs to which we could apply. Hence, all résumés needed to be consistent with the constant information provided in the candidates' social media profiles (for instance, all candidates' profiles exhibited the same master's degree in information systems). Therefore, the openings (and respective résumés) tended to be technical, managerial, or analytic in nature. Two human resource recruiters vetted the résumés for realism, professionalism, and competitiveness before they were tested in the online pilot. A design objective (also tested during the online pilot) was to create a sufficiently competitive candidate for stimulating the level of interest necessary to generate an online search, but not so competitive as to outweigh any potential effect arising from an employer's perusal of the candidate's online profile.

The résumés did not include links to the candidates' personal profiles or references to the personal traits we manipulated on the SN. The experimental design relied entirely on the possibility that employers would autonomously decide to seek information about applicants online, searching for their names either on popular search engines or directly on popular social networks.

4.2.4. Professional Network Profiles. We designed PN profiles for each name, maintaining identical profile information across conditions. The content of the profiles (Table A.3 in the e-companion) reflected the information provided in the résumés. To increase realism, we also designed additional PN profiles for other fictional individuals and connected them to the candidates' profiles, so that they would become "contacts" for our candidates. We took great care to avoid any possible linkage between the actual candidates' profiles.[23] The number of PN connections was identical across all the profiles and chosen to be as close as possible to the median number of connections of a U.S.-based profile on the PN professional networking site at the time of the experiment.[24]

4.2.5. Social Network Profiles. SN profiles served as the principal manipulation vehicle in the field experiment. We paired each candidate with an SN profile created under the candidate's name. As detailed in the rest of this section, we used a number of strategies to create realistic and balanced profiles and capture the phenomenon of online social networking in an ecologically valid way: we populated our candidates' profiles with data extracted from actual profiles of social network users who were demographically similar to the candidates; we made sure that the overall amount of public self-disclosure in our profiles would be equivalent to the amount of self-disclosure in actual social network profiles of users demographically similar to our candidates; we designed profiles to include a combination of information that changed according to the experimental condition and information that was held constant across all conditions; and finally, we posted the same amount of manipulated and constant information for each candidate (therefore, the Christian and Muslim candidates' profiles, and the gay and straight candidates' profiles, presented equivalent information about the strength of their religion or sexual orientation).

First, we downloaded the public profiles of 15,065 members of the same Facebook college network from which the candidates purportedly graduated. The vast majority of those profiles belonged to individuals with a similar age range (their 20s), current location (a North American midsize city), and educational background as our candidates. As further detailed below, among that set of profiles we then focused on those with the same gender (male) as our candidates and who self-disclosed traits matching the ones manipulated in the experimental conditions (that is, Christian, Muslim, straight, or gay). Then, we designed several components of each SN profile on the basis of the information we had mined from real online profiles: personal information (such as current location, location of birth, education, and employment history); closed-ended text fields (such as interests, activities, books, movies, or television shows); open-ended text fields (such as those providing a summary self-description of the individual and his favorite quotations); friends list; primary profile photo and secondary photos; and background image. On the basis of the literature that emerged during the design phase of the experiments,[25] however, we decided not to make all this information publicly available and therefore visible to a human resources professional. To avoid potential confounding effects from over-disclosure, we refrained from showing certain fields that only a minority of users of the network actually publicly show, such as posts, status updates, and friends list.

Table A.3 in the e-companion presents the resulting information included in each of the four SN profiles. By design, some of these profile components (the non–condition-specific traits) were constant across the different profiles and thus across the different experimental conditions. Other components (the condition-specific traits) were manipulated across conditions. We chose a mix of condition-specific and non–condition-specific traits that replicated the combination and balance we observed in the actual profiles we mined and analyzed (as described above). The design choice of manipulating fields representing different types of personal information represents an experimental design trade-off: relative to tightly controlled laboratory experiments, this manipulation operates through multiple subtle differences across profiles; on the upside, the manipulation increases the realism and ecological validity of our experimental design—which we tested in our pilot experiment. We discuss each type of component in the following subsections.

4.2.6. Information Kept Constant Across SN Profiles. By design, we made the candidates' primary profile photo, secondary photos, current location, home town, age, education, employment history, and friend list constant across conditions. Basic personal information (such as the SN member's current location and city of birth, educational background, employment history, home town, and age) was made publicly visible on the profile and kept constant across the conditions. That information was made consistent with the content of the résumé. The candidate was presented as U.S.-born and English-speaking.[26]

The photo on the candidates' social media profiles was of a Caucasian male with dark hair and brown eyes. We picked the photo after recruiting nonprofessional models on Craigslist and asking them to submit a portfolio of personal photos. Approximately 40 subjects (recruited via MTurk) rated all models who submitted their photos along two seven-point Likert scales, indicating their perceived attractiveness and professionalism. We selected one male model who received median perceived attractiveness and professionalism ratings.

We designed numerous "friends" profiles to connect to the candidates' profiles. Again, the number and identities of friends were identical across conditions.[27] The set of friends showcased a mix of names, backgrounds, and profile visibility settings. We set the friends list to "private"—that is, not visible to an employer—because of extant research suggesting that the overwhelming majority of members of the network do not publicly disclose their friend lists.[28]

4.2.7. Information Manipulated Across SN Profiles. Some fields in the profiles were manipulated across conditions: closed-ended text fields (e.g., interests), open-ended text fields (e.g., quotations), and the background image. Furthermore, we manipulated the candidates' sexual orientation by filling out the field "interested in" (either male interested in females or interested in males), and we manipulated religious affiliation through the "religion" field (either Christian or Muslim, with no specific denomination).

We abided by a number of principles in designing the manipulated information. First, we constructed "baseline" profile data, which were constant across all conditions, including information such as specific interests and activities statistically common among the SN profiles we had previously mined ("baseline information"). We then augmented the profiles with additional data (such as additional interests or activities) specific to the traits that we wanted to manipulate. Both baseline and manipulated information were extracted from real, existing SN profiles. The profiles therefore represented realistic pastiches of *actual* information retrieved, combined, and remixed from existing social media accounts.[29] We took care to avoid confounding the trait we were manipulating with signals of worker quality beyond signals inherent in the manipulated trait itself. Furthermore, the different candidates' profiles were designed to disclose the same amount of personal information, including the same quantity and types of data revealing their sexual orientation or religious affiliation, so as not to create profiles with unbalanced condition-specific disclosures. For similar reasons, we took great care to ensure that the information revealed by the candidates would not be construed as over-sharing, or as a forced caricature of what a Muslim (or Christian, or gay, or straight) profile should look like. The combination of baseline and treatment information of true existing profiles was one of the strategies we used to meet this goal; using only information existing in other SN profiles, and making certain fields publicly inaccessible, were two others. The online pilot experiment, with its open-ended questions about the profiles, tested whether our goal was attained (see Section 5.1 and Appendix A in the e-companion).

Closed-ended text fields, such as interests and activities, were extracted using a combination of manual and statistical analysis from real SN profiles of people who shared demographic characteristics with the candidate (for the "baseline information") and who exhibited the same characteristics we manipulated (for "treatment information"). For example, when text involved countable objects such as one's favorite books or movies, we calculated the most popular items listed by the overall population of SN profiles and used that as the "baseline information" for the candidate. Then we repeated the operation, focusing on the subset of the public disclosures of more than 15,000 SN profiles that displayed the same traits we manipulated, to create the "treatment information." For instance, we constructed a portion of the profile text indicating activities and interests for the Christian male profile using statistical analysis of the entire sample of more than 15,000 profiles that also included non-Christian ones; the remaining portion of the profile text concerning his activities and interests was constructed using statistical analysis of the subsample of profiles of Christian males at his university. If our sample did not provide enough information (for instance, movies) for individuals with a given trait, we complemented the statistical analysis with a manual analysis of the same SN profiles. We also extracted open-ended fields (such as personal self-descriptions, personal quotations, or nontextual fields such as profile background images) through manual analysis of existing profiles, because the extreme variety of styles and contents across open-ended texts made a statistical approach unfeasible.

5. Results

5.1. Online Pilot Experiment

Before submitting résumés to actual job openings at U.S. firms, we conducted a pilot experiment, consisting of a randomly assigned questionnaire with online participants and a between-subjects design. The online pilot was designed to test whether the treatment conditions successfully manipulate relevant beliefs, such as the religious affiliation or sexual orientation of the candidates. Furthermore, its open-ended questions were designed to test the perceived realism of the profiles. Finally, the pilot experiment was designed to test whether, in absence of the manipulated personal profiles (SN), the candidates' names, résumés, and professional profiles elicited different propensities to invite the candidate for an interview. Details of the design, analysis, and findings of the online pilot are presented in Appendix A in the e-companion and are summarized here.

We recruited more than 1,750 U.S. residents as participants using Amazon Mechanical Turk. Participants were presented with the candidates' social profiles and résumés prepared for the actual field experiment. Participants included individuals with previous hiring experience—something we exploited in the analysis of the results. Participants were randomly assigned to one of four conditions (gay, straight, Christian, or Muslim candidate) and were provided links to one candidate's résumé, PN profile, and SN profile. The survey instrument had four elements: (i) introduction and randomized manipulation; (ii) measurement of perceived employability; (iii) measurement of beliefs for the purpose of manipulation check; and (iv) open-ended questions and demographic characteristics. Participants were asked to evaluate the candidate. The main dependent variables were hypothetical willingness to call the candidate for an interview and perceptions of the candidate's suitability for the job.

The online pilot found, first, that the treatment conditions successfully manipulated relevant beliefs, such as the religious affiliation and sexual orientation of the candidates. Second, open-ended questions checked for perceived realism of the profiles; they provided no evidence of doubt, among the participants, that the candidates were real. Third, the online pilot tested whether, in absence of links to the manipulated personal profiles (SN), the candidates' names, résumés, and professional profiles elicited different propensities to invite the candidate for an interview; we found no evidence that the candidates' names and professional materials elicited different responses in the absence of the manipulated online social media profiles. Finally, responses of hypothetical hiring behavior and judgments of employability provided evidence to complement the findings of our field experiment. Consistent with our field findings (presented further below in this section), manipulated profiles in the online pilot elicited no bias against the gay candidate, relative to the straight candidate. However, among subjects with hiring experience, we found highly significant bias against the Muslim candidate relative to the Christian candidate, especially among those who self-identify as Republican and Christian.

5.2. Field Experiment

The field experiment consisted of a between-subjects design in which each employer was randomly assigned to receive one job application from either the Muslim, Christian, gay, or straight candidate.

We extracted job openings from Indeed.com, an online job search site that aggregates jobs from several other sites.[30] We selected positions that fit the candidates' backgrounds—namely, positions that required either a graduate degree or some years of work experience. For each position, we sent the most appropriate version of our résumé. (Recall from Section 4.2

that we designed different versions of the résumé to fit 10 different job types covering a combination of technical, managerial, and analytic positions.)

We defined several criteria that jobs and companies had to pass for us to apply to them. We focused on private sector firms. Primarily, the job had to be related to the candidates' background and level of experience, although we also included (and controlled for) positions for which the candidates could be considered slightly over- or under-qualified. In addition, we carefully avoided sending two applications to the same company, or to companies that were likely to share HR resources such as databases of applicants (for instance, parent companies of firms to which we had already applied).[31] We also excluded staffing companies, companies located in the same geographic region as the candidates' reported current location, and companies with 15 or fewer employees (to limit the costs imposed on them by the process of searching and vetting fictional job candidates).[32] All applications (résumés and cover letters) were submitted online, either by email or through employer-provided web forms.

We recorded the city and state listed on job postings, and, when possible, we recorded the city and state where the job would be located. When job location was not provided, we used the location of the company's headquarters. We obtained this measure for all observations and used it for our state-level analysis. From the company name, we were able find the street address of the company headquarters for all but a few hundred observations. We used ArcGIS[33] to match this street address to its county. We then merged our data with county-level data from the American Community Survey[34] based on the county where the company headquarters is located.

Unlike laboratory experiments in fields with a long history of collecting data under similar conditions, formal sample power calculations raised a challenge in our context: existing literature provided little guidance regarding empirical estimates of employers' probabilities of searching online for job candidates. Without those, we could not precisely anticipate what treatment rates to expect. Samples used in previous résumé studies (see Section 2.2) have ranged from a few hundred to a few thousand employers. We applied to every U.S. job opening that fit our predefined criteria of a reasonably suitable position for our candidates, until finding jobs that met said criteria became increasingly arduous and infrequent. This amounted to 4,183 U.S. job applications (or roughly 1,045 employers per experimental condition) collected from early 2013 through the summer of 2013. Ten applications failed to meet criteria we had defined in our procedures for acceptable and complete applications, leaving us with 4,173 usable applications. Although it was still possible that the sample might be underpowered, ultimately we matched the higher end of the sample size spectrum of previous résumé studies. Summary statistics for our sample can be found in Table A.13 in the e-companion.

5.2.1. Search Trends. Although the focus of the experimental design was capturing employers' callback behavior, tracking and identifying employers' searches and visits to the manipulated profiles would be desirable. Without access to the social networks' proprietary traffic data, however, it is not possible to observe how many employers visited the candidates' profiles and thus got treated to the experimental manipulation. Hence, the experimental design does not allow us to capture directly the number of times the candidates' profiles were searched for or perused. In fact, in principle, *all* employers who received the candidates' applications could have searched for the candidates directly via the SN (thus remaining undetectable to us). In practice, that seems unlikely. It is possible, however, to obtain rough estimates of frequency of employers' online searches of job candidates using various indirect approaches. We summarize two approaches and their associated findings in this section and present detailed assumptions, calculations, and additional results in Appendix B of the e-companion.

We used a combination of two data sources to estimate the frequency with which employers searched for the candidates online. One data source consisted of Google AdWords "Keyword Tool" statistics. These publicly accessible statistics capture the number of times a certain term is searched on Google from various locations. We used this tool to estimate the number of times the exact names of the candidates were searched from U.S. IP (Internet Protocol) addresses. The second data source consisted of statistics provided by the PN network (LinkedIn) via so-called "Premium" accounts. If a user subscribes to a Premium account, that user will be able to get information such as the count of visits to its PN profiles, and in some cases the actual identity of the visitors. We subscribed to Premium accounts for each of our candidates' profiles, to track visits to those profiles.

Each of these sources of data is imperfect. Google Keyword Tool statistics provide aggregate monthly means for searches of a given term, rather than raw data. Furthermore, if the mean is higher than zero but below 10 searches per month, no exact count is provided. Similarly, "Premium" PN accounts do not actually track all visits to that account (for instance, in our tests, visits from subjects that had not logged onto the network went undetected and uncounted; visitors who only viewed the summary profile of the candidate, rather than his full profile, also were not detected or counted; in addition, certain Premium accounts may allow "invisible" views of other LinkedIn profiles). Nevertheless, considered together, these data sources do offer a rough estimate of the proportions of employers searching the candidates online.

We tracked Google Keyword Tool statistics and LinkedIn data over a period of several months. On the basis of the documented searches we detected on Google and on the PN, we can estimate a minimum lower threshold of employers who searched for the profiles at 10.33%, and the likely proportion of employers who searched at 28.82% (see Appendix B of the e-companion for details). These estimates seem consistent with the results of a 2013 CareerBuilder survey of 2,291 hiring managers and HR managers, according to which 24% claimed to "occasionally" use social media sites to search for candidates' information; 8% answered "frequently"; and only 5% answered "always."[35]

The search rates highlighted in this section are sample-wide, aggregate estimates, because (as noted) in most cases we could not identify the specific employers visiting the profiles. The subset of cases for which we could directly capture the identity of an employer searching for our profiles represents a small sample size (n = 121) but, unsurprisingly, represents employers who are strongly interested in our candidates. The overall callback rate for the four candidates in this subsample was 39.67%. The callback rates for the straight and gay candidates were 31.25% and 40.00%, respectively (χ^2 p-value = 0.472). The callback rates for the Christian and Muslim candidates were 54.84% and 32.14%, respectively (χ^2 p-value = 0.080).

5.2.2. Employer Callbacks. Our primary dependent variable is a callback dummy that equals one if the candidate was contacted for an interview and zero if the candidate received a rejection, no response, or a scripted request for more information. Of the 4,173 observations from both manipulations, 11.20% were callbacks, 15.86% were rejections, 69.29% had no response, and 3.32% were scripted requests for more information.

5.2.2.1. Timing of Callbacks. Figure A.1 (in the e-companion) shows the geographical distribution of applications across the United States. Figure A.2 (in the e-companion) provides a distribution of the timing of responses we received across all candidates, where a response is an interview invitation, a rejection, or a request for more information. This figure does not include the 70% of employers who did not respond to the application. Most responses came within the first 14 days.[36] Roughly 77% of responses arrived within one month from the application and roughly 90% within two months. The distribution of timing of responses follows the same patterns across the four candidates.

5.2.2.2. Callback Rates: Empirical Approach. Prior evidence (including both public surveys and field experiments) suggests a link between socio-economic traits (captured at the individual or the geographical levels) and the probability of bias toward members of certain groups. In particular, various sources suggest a link between political orientation and possible discrimination: surveys show more bias against Muslims and gay people among Republicans than Democrats (Arab American Institute 2012, Pew Research Center 2013); related field experiment evidence by Tilcsik (2011) links hiring bias toward gay people to geographical measures of political party support; and the results of our own online pilot experiment suggest a link between political orientation and attitudes toward Muslim job candidates. In communities where individuals are more likely to express bias, employers themselves may be more likely to share those attitudes, or may simply face incentives to hire candidates whose traits match those preferred by existing employees or community members. Thus, we test for stronger bias in certain areas—specifically, among firms located in states and counties with a high fraction of voters in the 2012 Presidential election who voted for the Republican candidate, Mitt Romney (we consider possible other sources of cross-sectional heterogeneity in Section 5.2.3, which discusses the robustness of our results).

Because we cannot identify the individual making the callback decision, we cannot observe his or her traits, such as his or her political identification. Following the literature, we use instead characteristics of the firm's geographical area. Our measure of political party support follows Tilcsik (2011), who used Presidential election data in his audit study, and the Gallup Organization, which regularly produces lists of the 10 most Republican and 10 most Democratic states based on survey data. Using Presidential election data from Leip (2013), we thus indicate the 10 states with the highest fractions of 2012 Romney voters as Republican states and the 10 states with the lowest as Democratic states. We refer to the remaining states as politically mixed. We indicate Republican and Democratic counties as those with the same cutoff values for Romney vote fractions that were used to indicate the Republican and Democratic states.

We control for a host of other variables used in comparable résumé studies, particularly Bertrand and Mullainathan (2004) and Tilcsik (2011), including employer characteristics (such as industry or number of employees), job characteristics (such as job-specific requirements), and regional characteristics based on location of the job (for instance, variations in state and local policies to protect gay people). Finally, we include controls that are specific to our experimental design's online dimension, such as a measure of Facebook penetration across states (a proxy for the likelihood that an individual in a particular state may be inclined to search online for the candidate's social media profile).

5.2.2.3. Callback Rates: Results. Table 1 presents summary statistics on callbacks for the whole sample (column 1) and by political party support (columns 2–7). The nationwide callback rates for the straight and gay candidates are 10.63% and 10.69%, respectively (panel B). The callback rates for the Christian and Muslim candidates are 12.64% and 10.96%, respectively (panel A). Neither of these effects is statistically significant. However, we find significant heterogeneity in hiring bias against the Muslim candidate relative to the Christian candidate. Panel A shows that in Republican states, 17.31% of applications by the Christian candidate received interview invitations, compared with only 2.27% for the Muslim candidate ($p = 0.016$).[37] At the county level, the callback rates for the Christian and Muslim candidates in Republican counties are 22.58% and 6.25%, respectively ($p = 0.009$). In contrast, there are no significant callback biases in politically mixed or Democratic states or counties. In Democratic states, the Christian and Muslim candidates have callback rates of 11.61% and 11.74%, respectively. In Democratic counties, the Christian and Muslim candidates have callback rates of 12.13% and 12.37%, respectively. In the gay/straight manipulation, we find no significant differences in bias across any of the geographical areas.[38] Figure 1 summarizes these findings using histograms of callback rates (and 95% confidence intervals) for each manipulation in the 10 most Republican states, the 10 most Democratic states, and the remaining states.

We also test whether the biases found in panel A of Table 1 stem from one or both of the Muslim and Christian conditions. Panel C compares callbacks in each of these conditions with callbacks in the pooled gay/straight conditions, which we use as a benchmark. In Republican states, the Muslim callback rate of 2.27% is significantly lower than the pooled gay/straight callback rate of 14.77% ($p = 0.028$). In Republican counties, as well as in mixed states, the callback rate in the Christian condition is significantly *higher* than the callback rate for the pooled gay/straight condition ($p = 0.016$). This potentially, but cautiously, suggests that both the Muslim and the Christian conditions may be playing significant roles (respectively unfavorable and favorable) in generating callback bias, especially in more-Republican areas. (This result is not sensitive to which subsample is used as the benchmark. We find similar results when we separate the pooled gay/straight candidate sample into just the gay candidate sample or just the straight candidate sample. In both the gay candidate sample and the straight candidate sample, we find that the callback rate for the Muslim candidate in Republican states is significantly lower and the callback rate for the Christian candidate in Republican counties is significantly higher.)

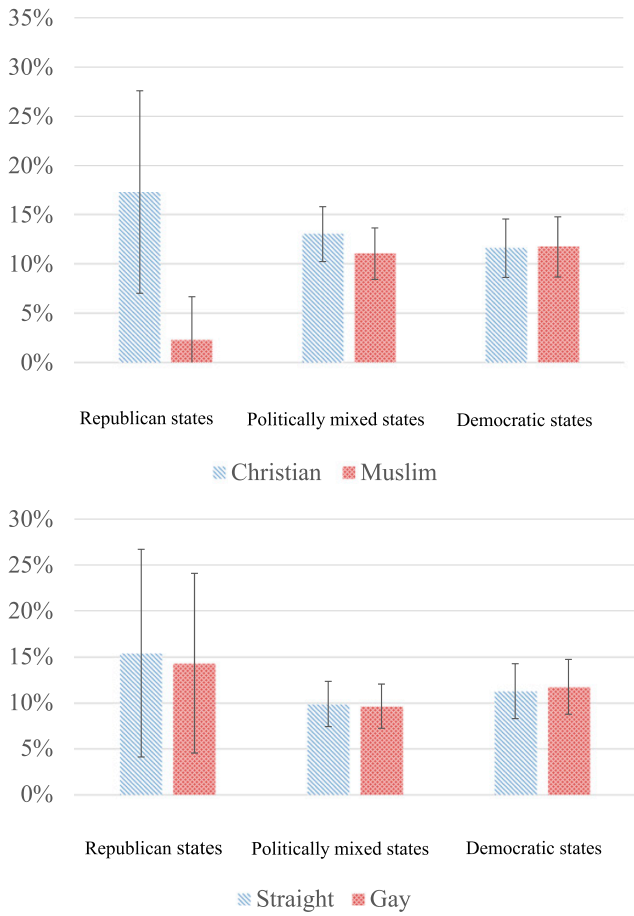
5.2.2.4. Callback Rates: Regression Analysis. The above results suggest that bias against the Muslim candidate relative to the Christian candidate might be significantly larger in Republican areas than in

Table 1. Callback Rates Across Conditions by Geographical Area

	Whole sample	State-level subsamples			County-level subsamples		
	Nationwide	Republican	Mixed	Democratic	Republican	Mixed	Democratic
Panel A							
Muslim callback rate, % (n)	10.96 (1,022)	2.27 (44)	11.05 (552)	11.74 (426)	6.25 (64)	10.00 (360)	12.37 (574)
Christian callback rate, % (n)	12.64 (1,060)	17.31 (52)	13.04 (560)	11.61 (448)	22.58 (62)	12.39 (347)	12.13 (610)
p-values (χ^2)	0.234	0.016	0.309	0.952	0.009	0.313	0.901
Panel B							
Gay callback rate, % (n)	10.69 (1,066)	14.29 (49)	9.60 (573)	11.71 (444)	9.80 (51)	9.38 (352)	11.65 (635)
Straight callback rate, % (n)	10.63 (1,025)	15.38 (39)	9.84 (559)	11.24 (427)	8.16 (49)	10.78 (334)	11.02 (608)
p-values (χ^2)	0.965	0.885	0.891	0.828	0.774	0.541	0.725
Panel C							
Pooled gay/straight callback rate, % (n)	10.66 (2,091)	14.77 (88)	9.72 (1,132)	11.48 (871)	9.00 (100)	10.06 (686)	11.34 (1,243)
χ^2 p-value for Muslim vs. gay/straight	0.804	0.028	0.395	0.892	0.525	0.976	0.527
χ^2 p-value for Christian vs. gay/straight	0.098	0.690	0.039	0.946	0.016	0.255	0.619

Note. Fisher's one- and two-sided exact tests are consistent with the results of the χ^2 tests shown in this table.

Figure 1. (Color online) Callback Rates by Political Leaning of States



Note. Error bars represent 95% confidence intervals based on tests for equality of proportions.

Democratic areas. In this subsection, we test this interaction with ordinary least squares (OLS) regression analysis of state-level data. We also present the analogous test for the gay/straight manipulation. We show in Section 5.2.3 and the e-companion that the results are robust to using county-level data, to using logit with full interaction effects following Norton et al. (2004), to various clustering options for standard error estimates, and to additional specifications and measures of political areas.

For the Muslim versus Christian pairwise manipulation, our specification is as follows:

$$Callback_i = \beta_0 + \beta_1{}^{*}M_i + \beta_2{}^{*}M_i{}^{*}PoliticallyMixedStates_i + \beta_3{}^{*}DemocraticStates_i + \mathbf{x}_i{}'\boldsymbol{\gamma} + \mathbf{z}_i{}'\boldsymbol{\delta} + \varepsilon_i, \quad (1)$$

where M_i is a dummy for firm i being assigned to the Muslim candidate condition ($M_i = 1$) compared with the Christian condition ($M_i = 0$), $\beta_0, \ldots, \beta_3$ are unknown parameters, *PoliticallyMixedStates* and *DemocraticStates* are dummy variables, $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ are vectors of, respectively, observed and unobserved regressors capturing employer i's traits, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are vectors of unknown parameters, and ε_i is an error term. $Callback_i$ equals one if employer i contacts the candidate for an interview and zero if there is no response or an explicit rejection. The regression model for the gay/straight manipulation is similarly specified.

Search rates should not differ across the randomized conditions (and we find no evidence that they do), because employers are treated by our manipulated profiles after they choose to search. However, it is likely that both search and discrimination probabilities will vary with employer characteristics, leading to an interaction between our manipulation dummy and an employer characteristic, such as political party strength in the employer's area.[39]

Table 2 presents OLS regressions for the Muslim/Christian manipulation (column 1) and the gay/straight manipulation (column 2). The dependent variable is callbacks. These are state-level regressions that include a dummy for assignment to the Muslim condition (column 1) or gay condition (column 2), dummies for the politically mixed states and Democratic states, interactions between the Muslim (or gay) dummies and the political area dummies, and a full set of control variables (see Table 2 notes for details). Standard errors are clustered at the state and county levels.[40]

The estimated callback rate for the Christian candidate in Republican states is 0.102. The estimated callback rate for the Muslim candidate in Republican states is lower compared with the Christian condition, by −0.132. This effect is statistically significant. The interaction between the Muslim condition and the Democratic states dummy is positive, 0.127, and statistically significant. The interaction between politically mixed states and the Muslim assignment dummy is also positive, 0.115, but not significant at the 5% level. Column 2 presents the analogous specification for the gay/straight manipulation. We found no significant effects in this manipulation.

The results in Table 2 are consistent with prior results that were suggestive of similar patterns but not definitive. First, they are consistent with our online pilot experiment (see Section 5.1), which showed significant bias against the Muslim candidate relative to the Christian candidate among subjects with hiring experience but no bias against the gay candidate relative to the straight candidate (see Appendix A of the e-companion). Second, in a Republican area, Gift and Gift (2014) find evidence of statistically significant discrimination against job candidates who signal Democratic Party affiliation on résumés compared with those who signal Republican Party affiliation. In a Democratic area, they find a directionally consistent bias against job candidates who signal Republican Party affiliation, but it is not significant at the 5% level. Their

Table 2. OLS Regressions Using State-Level Data in the Christian–Muslim Conditions (Column 1) and the Gay–Straight Conditions (Column 2)

	(1) Muslim vs. Christian	(2) Gay vs. straight
Muslim candidate	−0.132**	—
	(0.054)	—
Politically mixed area	−0.024	−0.039
	(0.056)	(0.054)
Democratic area	−0.025	0.024
	(0.072)	(0.072)
Muslim × Politically mixed area	0.115*	—
	(0.057)	—
Muslim × Democratic area	0.127**	—
	(0.060)	—
Gay candidate	—	−0.018
	—	(0.071)
Gay × Politically mixed area	—	0.020
	—	(0.073)
Gay × Democratic area	—	0.012
	—	(0.074)
Controls included?	Yes	Yes
Constant	0.102	0.094
	(0.067)	(0.057)
Observations	1898	1923
R^2	0.043	0.035

Notes. Dependent variable is callbacks. Standard errors (in parentheses) are clustered on the state and county where the job will be located. Dependent variable equals one if the candidate is contacted for an interview, zero otherwise. State-level geographical controls included are median age, fraction Muslim (column 1), 2012 unemployment rate, fraction foreign-born, fraction gay households (column 2), natural log of median income, fraction nonwhite, fraction college educated or more, fraction evangelical Christian, fraction urban, Facebook penetration, legal protection from religious discrimination (column 1), and legal protection from sexual orientation discrimination (column 2). Firm-level controls are as follows: dummies for women and minority owned, public firm, large firm (500 employees or more), federal contractor; and included application-level characteristics are dummies for entry-level position, references required, preferred salary required, master's degree required, one or more years of experience required, multiple fields of experience required, and nine dummies for field of employment. Continuous variables—namely, median age, fraction Muslim, 2012 unemployment rate, fraction foreign-born, fraction gay households, natural log of median income, fraction nonwhite, fraction college educated or more, fraction evangelical Protestant, fraction urban, and Facebook penetration—are demeaned. The omitted categories for dummies capturing variables with more than two categories are Republican states and jobs in information systems. The remaining variables are binary. Full regression results are available in Table A.14 in the e-companion.

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

findings are consistent with our finding of stronger discrimination in Republican areas.

Existing findings on discrimination based on sexual orientation are mixed (Tilcsik 2011, Bailey et al. 2013). We do not observe statistically significant levels of discrimination in our analysis, but our estimates are too imprecise to rule out meaningful levels of discrimination.

5.2.3. Robustness Checks. The results of Figure 1 and Table 2 are robust with respect to different choices for clustering of standard errors, different combinations of included regressors and interactions, using county-level instead of state-level data, different measures of Republican, Democratic, and politically mixed areas, and using logit instead of OLS.

In Table A.15 in the e-companion, we show that the results of Figure 1 and Table 2 are robust to substantial changes in how many control variables are included and to whether standard errors are clustered. Column 1 of Table A.15 presents regressions of a baseline specification with no included control variables (like that of Figure 1) with standard errors that have no clustering. Column 2 clusters on the state and column 3 clusters simultaneously on the state and county. Columns 4 and 5 replicate the regression of Table 2 with the full list of control variables using standard errors that have, respectively, no clustering and clustering at the state level. The effects of the Muslim candidate dummy and the Muslim × Democratic area interaction are significant at the 1% or 5% level in every column of this table.

In Table A.16 in the e-companion, we show that the results of Table 2 and Figure 1 are robust to using county-level data instead of state-level data, with

Management Science, 2020, vol. 66, no. 3, pp. 1005–1024, © 2019 INFORMS

clustering on the state and county, and both with and without state fixed effects. These county-level results address two drawbacks of the state-level analysis. First, taking all four conditions together, there are 1,745 employers in 10 Democratic states or districts (including Washington DC), 2,244 in 31 mixed states, and 184 in 10 Republican states. The county-level division between political areas is different, with sample sizes in the Democratic and Republican counties of 2,427 and 226, respectively. Second, we cannot fully control for state-level covariates with state-level definitions of political areas. County-level regressions can address this with state fixed effects. Columns 1–3 of Table A.16 show that the baseline interaction suggested in Figure 1 is significant in baseline county-level regressions with no control variables using standard errors that have, respectively, no clustering, clustering on the state, and clustering on the state and county. Columns 4–6 include state fixed effects, and the full list of control variables that were included in Table 2, with standard errors that have, respectively, no clustering, clustering on state, and clustering on state and county. Again, the effects of the Muslim dummy and the Muslim × Democratic area interaction are significant at the 1% or 5% level in all of the six columns of Table A.16.

In Table 3 we introduce a substantive change to the measure of political areas (column 1) and to the empirical specification (column 2). We have so far defined the Republican, Democratic, and politically mixed states according to fractions of voters who voted for Romney in the 2012 U.S. presidential election. Let us refer to this as the *Romney vote list*. In Table 3, we use a broader measure of Republican or conservative states and Democratic or liberal states, which we refer to as the *Combined list*. We define the Republican (Democratic) states in the *Combined list* as the union of (i) the *Romney Vote list* of Republican (Democratic) states, and (ii) the Gallup Organization's list of the 10 most politically conservative (liberal) states in 2012, which we refer to as the *Gallup list*. [41] There is no intersection of states on the *Combined* Republican and

Table 3. OLS Robustness Checks Using a Broader Measure of Republican and Democratic States, and Additional Interactions

	(1) Same as Table 2, column 1, but uses a more inclusive measure of Republican and Democratic areas	(2) Same as column 1, but with two additional interactions
Muslim candidate	–0.160***	–0.160***
	(0.056)	(0.059)
Politically mixed area	–0.064	–0.065
	(0.060)	(0.061)
Democratic or liberal area	–0.077	–0.076
	(0.074)	(0.074)
Muslim × Politically mixed area	0.150**	0.151**
	(0.060)	(0.062)
Muslim × Democratic or liberal area	0.150**	0.149**
	(0.060)	(0.062)
Median age (demeaned)	0.008	0.009
	(0.005)	(0.006)
Fraction Muslim (demeaned)	0.834	0.647
	(1.007)	(1.303)
Muslim × Fraction Muslim	—	0.404
	—	(1.668)
Muslim × Median age	—	–0.001
	—	(0.007)
Controls included?	Yes	YES
Constant	0.146**	0.146**
	(0.069)	(0.070)
Observations	1898	1898
R^2	0.044	0.044

Notes. Dependent variable is callbacks. Standard errors (in parentheses) are clustered on the state and county where the job will be located. Dependent variable equals one if the candidate is contacted for an interview, zero otherwise. Republican states are defined by the union of (i) the Gallup Organization's list of the 10 most politically conservative states and (ii) the Romney vote list of the 10 states with the highest fraction of votes for Romney in the 2012 Presidential election used in Table 2. Democratic states are defined analogously. Remaining states are "politically mixed." Included control variables in column 1 are the same as those in Table 2, column 1 (see notes to Table 2 for the full list). Column 2 is the same as column 1 except that it also includes an interaction between state-level fraction Muslim and the Muslim condition dummy, and an interaction between state-level median age and the Muslim condition dummy. Full regression results are available in columns 2 and 5 of Table A.17 in the e-companion.

$^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Combined Democratic lists. Thus, the remaining states are coded as "politically mixed." (See Table A.18 in the e-companion for complete lists of states according to the *Romney vote*, *Gallup*, and *Combined* lists.)

Column 1 of Table 3 repeats the specification of column 1 of Table 2, except that this broader *Combined list* of Republican and Democratic states is used. Recall that this specification, used in both Table 2 and Table 3, controls for a host of potentially relevant geographical background variables, including median age, fraction Muslim, unemployment rate, fraction foreign-born and fraction nonwhite, median income, education, fraction evangelical Christian, and how rural versus urban the area is. The results using the *Combined list* are very similar to the results from using the *Romney vote list* in Table 2.

Column 2 of Table 3 is identical to column 1, except that two additional interaction terms are included. These two interaction terms, after political party support, are the ones that are most consistently suggested in connection with heterogeneous bias toward Arabs or Muslims.[42] The added interactions in column 3 are interactions between (i) the Muslim candidate dummy and the demeaned state-level median age, and (ii) the Muslim candidate dummy and the demeaned state-level fraction Muslim. These additional interactions are not statistically significant, and including them has no noteworthy effects on either the effect of the Muslim condition dummy in Republican or conservative areas, estimated at the mean of state-level fraction Muslim and the mean of state-level median age, or the interactions between the Muslim condition dummy and the dummies for politically mixed and Democratic areas.[43]

Table A.17 in the e-companion summarizes robustness analysis using three alternative measures of political areas. This table replicates the specification with full controls in Table 2 using the *Gallup list* (column 1) and *Combined list* (column 2). The rest of the table uses the *Combined list* for additional robustness checks. Columns 3 and 4 show the smallest and largest interactions with the Democratic states that we could find by deleting each of the 50 states (including Washington, DC), one regression at a time. Column 5 is the same as column 2 of Table 3 (but with all of the control variable estimates presented). The effects of the Muslim dummy, the Muslim × politically mixed area interaction, and the Muslim × Democratic area interaction are all significant at the 1% or 5% level in all five columns of Table A.17. Finally, because the dependent variable (callbacks) in our analysis is binary, we present logit analysis in Table A.19 and Figure A.3 in the e-companion. We present marginal effects for full interactions, which have been derived by Norton et al. (2004) for a single interaction. To specify a single interaction for the logit analysis, we pool two of our political area dummies (politically mixed areas and Democratic areas) and interact that single pooled non-Republican dummy with the Muslim candidate condition dummy. The logit analysis thus tests for a significant difference in Muslim versus Christian callback bias between Republican areas and non-Republican areas. Column 1 of Table A.19 presents these results for a baseline logit specification with no included control variables. In this special case, with only two dummy variables and an interaction, the Norton et al. (2004) full interaction effects do not vary with the dependent variable. These results replicate OLS results. The marginal effects of the Muslim candidate dummy (−0.143) and its interaction with non-Republican states (0.133) are roughly similar in size to OLS marginal effects and are significant at the 5% level. Table A.19, column 2 presents the full marginal effects for a logit specification that includes all of the control variables that were included in the OLS regressions of Table 2. The marginal effect for the Muslim candidate dummy (−0.132) is again similar in size to the OLS marginal effect and is significant at the 5% level. With the addition of control variables to the model, the interaction effects now vary with the predicted callback probability. The marginal effects range in size from 0.007 to 0.341. The z-values range from 0.819 to 2.515. Figure A.3 in the e-companion plots the marginal effects of these full interactions on the predicted callback probability.

5.3. Limitations

A number of limitations are inherent to the experiment's design. First, a null effect in the gay/straight manipulation may not necessarily be interpreted as an absence of discrimination, because a number of potential explanations may exist, including the size of the sample. Second, heterogeneity in the assignment effect according to firm characteristics cannot be separated into firms' differences in search rates and differences in the treatment effect. Third, the results stem from candidates with certain characteristics, applying to certain types of jobs, with certain types of employers; they may not apply to other types of candidates, positions, or organizations. Fourth, we measure employers' traits with state- or county-level data; we do not capture the traits of the actual employers or human resource decision makers (such as his or her political identification). Thus, as noted in the Introduction, our results focus on geographical *areas* with varying aggregate voting outcomes. Although we cannot rule out the influence of other uncontrolled sources of cross-sectional heterogeneity across political areas, our various tests suggest that the results are robust to the consideration of multiple controls and additional interactions. Fifth, as we noted, unlike tightly controlled laboratory experiments, our manipulation operates through multiple subtle differences across profiles; on the upside, the manipulation increases the realism and ecological validity of our experimental design. Furthermore, discrimination based on information posted online may stem partly from the signal that the posted information sends about a candidate's judgment in choosing to post that information. However, the disclosures we investigated in this paper took place in personal profiles that employers must actively seek, rather than in résumés in which candidates intentionally reveal potentially sensitive information to employers, and the consequences of the fact that people use online social networks to disclose this information in this way, is part of what we sought to understand. Our experimental approach—namely, designing profiles to replicate information found in real social media profiles of individuals demographically comparable to our candidates (see Section 4.2)—increased the realism and ecological validity of our experimental design and decreased the chances that employers might interpret the profiles as "over-sharing" beyond the levels that we observe in real online social network profiles. Finally, as noted in Section 5.2, sample power calculations were made harder by the absence of reliable prior literature on employers' search rates. Although the sample size we ultimately achieved may still be under-powered, it represents the entirety of employers who met the multiple submission criteria listed in Section 5.2 during an extended period of study.

6. Conclusion

We set out to answer the question, do hiring decision makers seek online information that should not be used in the hiring process, and are they affected by what they find? We focused on the effects of religious affiliation and sexual orientation. We used randomized experiments to answer those questions, capturing discrimination both in behaviors of real employers and attitudes of survey subjects. Our results were broadly consistent across the online pilot and the field experiment. They suggest that although hiring discrimination via internet searches and social media may not be widespread for the companies and jobs we considered in this experiment, revealing certain traits online may have a significant effect on the hiring behavior of self-selected employers who search for the candidates online.

The findings provide further evidence of a phenomenon increasingly studied in the information systems literature: the impact of online information technology services on offline outcomes and behaviors. They also highlight the novel tensions arising between regulatory interventions designed in a nondigital world (that attempt to protect personal information) and technological innovations that bypass those protections (by making said information otherwise available).

This work suggests various directions for future research. There is a tension between a potentially efficiency-enhancing effect of online information on labor market search and a potential for that same information to increase labor market discrimination. To the extent that online markets facilitate not just matching but also discrimination, people in disadvantaged categories may face a dilemma: if they censor their online activity, they may be able to protect their privacy, but a limited online presence might, in itself, signal that there is something to hide, or that something is missing. Imagine a charismatic Muslim candidate with both a strong social network in his religious community and many professional contacts. If he were to reduce his online presence to hide his friendships in his religious community, at least three problems arise. The first is a "lemons" effect: people may assume that those with a restricted online presence belong to less preferred categories. The second is what we may term an "online network exclusion effect": by protecting his privacy, the candidate is deprived of the opportunity to present professionally relevant positive traits about himself. The third effect is a bundled traits effect: candidates may have some traits that employers find less desirable and some traits that are more desirable, and online mechanisms sometimes bundle together information about undesirable and desirable traits. For the charismatic Muslim candidate described above, information about personal traits that are potentially less desirable to employers may be bundled with information about traits that are more desirable to employers, including the strength of his professionally relevant social networks. This presents additional challenges for disadvantaged candidates who may wish to protect information about traits that are disadvantageous in the labor market while promoting their positively viewed traits. This dilemma may motivate design of mechanisms that facilitate communication only about job-relevant characteristics. Thus, a number of multifaceted questions for future research arise, which could be asked descriptively using the methodology in this paper: Does the amount of online self-disclosure affect labor market success? (For instance, in an age of self-disclosure, is the absence of social media presence a neutral, positive, or even negative signal? Do employers pay attention to the size of a candidate's personal social network?) Will market mechanisms develop to solve some of these problems? It may also be interesting to ask these questions normatively: How should online self-disclosure affect labor market success? How should markets be structured to address these problems?

Some employers openly admit to using social media in the hiring process (De La Llama et al. 2012). At the same time, legal professionals warn firms about the risks of doing so. Both the EEOC and state legislators have become interested in the role of the Internet and social media in the job search process. However, companies like HireRight.com and Rapleaf .com have offered related services—arguably shielding firms from the risks associated with doing the search themselves. In sum, the rise of the internet and social media has increased the amount of personal and professional information available to employers about a job candidate. This manuscript offers a first insight as well as a methodological path toward understanding the extent to which this new information channel may reduce labor market frictions or lead to more discrimination. Our results also point to the importance of future research to identify underlying mechanisms that might help explain why Republican areas in our study showed significantly more bias than other areas.

Acknowledgments

The authors thank the editors and the anonymous reviewers for thoughtful feedback. In addition, they thank a large team of research assistants for outstanding work, especially (but not limited to) Chris Canning (head research assistant), Jessica Dougan, Jeffrey Flagg, Oliver Haimson, Steven Mollura, Seth Monteith, Nicole Navolio, Merav Pishoto, Nikolas Smart, and Ben Wasser. The authors thank Idris Adjerid, Laura Brandimarte, Brian Kovak, Amelia Haviland, Eyal Peer, Lamar Pierce, Sasha Romanosky, Sonam Samat, Fred Stutzman, Curtis Taylor, Lowell Taylor, and Liad Wagman for helpful comments; they also thank attendees at conferences and workshops, including the European Association of Research in Industrial Economics, Northwestern University Workshop on Privacy, International Economic Science Association Meetings, the National Bureau of Economic Research's Economics of Information Technology and Digitization Workshop, the Economic Science Association Conference, the International Economic Science Association Conference, and the Bay Area Behavioral and Experimental Economics Workshop. The authors also thank participants in numerous seminars and lectures, including those at Boston University; Cornell University; New York University; University of California, Berkeley; University of Chicago; and University of Minnesota. The authors are also very grateful to the Ponemon Institute, CareerBuilder, and Socialbakers for providing data and to Ralph Gross and Tom Mosser for technical help.

Endnotes

[1] Various surveys suggest that U.S. employers search job candidates online, but the reported frequency of searches varies considerably across surveys and over the years. See, for instance, "Survey shows 48% of employers conduct social media background checks," *EmployeeScreenIQ*, accessed June 1, 2017, http://www.employeescreen.com/iqblog/social-media-2/48-of-employers-conduct-social-media-background-checks/; "Employers using social networks for screening applicants," *Wikibin*, accessed June 1, 2017, http://wikibin.org/articles/employers-using-social-networks-for-screening-applicants.html; "Ponemon Institute/Littler Mendelson Study," *International Association of Privacy Professionals*, accessed June 1, 2017, https://www.privacyassociation.org/publications/2007_12_ponemon_institute_littler_mendelson_study; and, more recently, "Number of employers using social media to screen candidates at all-time high, finds latest CareerBuilder study," *CareerBuilder*, accessed March 29, 2018, http://press.careerbuilder.com/2017-06-15-Number-of-Employers-Using-Social-Media-to-Screen-Candidates-at-All-Time-High-Finds-Latest-CareerBuilder-Study, reporting an increase in the number of U.S. employers using social media to screen candidates.

[2] See Theodore Claypoole, "EEOC regulations spotlight social media," Womble Carlyle, last modified May 24, 2011, accessed June 1, 2017, http://www.wcsr.com/Insights/Alerts/2011/May/EEOC-Regulations-Spotlight-Social-Media. On the potential risks of using social media in hiring for firms, see Harpe (2009).

[3] For instance, Texas S.B. 118 aims to prohibit an employer from "requiring or requesting access to the personal accounts of employees and job applicants through electronic communication devices; establishing an unlawful employment practice."

[4] Different types of personal information enjoy different levels of protection across U.S. states. Some personal traits cannot even be inquired about in interviews, whereas others cannot be used in the hiring decision. Some are protected across all states and others only in some states. For simplicity, we refer to information about these traits collectively as "protected" information but investigate state-level differences in the degree of their protection through our empirical analysis.

[5] For instance, an overwhelming majority of Facebook users in a sample of North American college students surveyed by Acquisti et al. (2014) used real first and last names on their profiles.

[6] Numerous examples exist of techniques through which seemingly pseudonymous online profiles can be re-identified across a variety of platforms and scenarios. See, for instance, Narayanan and Shmatikov (2009).

[7] Consider the gap between stated and actual privacy settings of online social network users reported by Acquisti and Gross (2006). Similar results have been subsequently found by Madejski et al. (2012).

[8] For instance, using data collected for this study (see Section 4.2), we estimate that, in 2011, 42% of all profiles members of the Facebook network of a major North American college shared "likes" publicly (where a like could be an interest, book, movie, music, or a television program). Data reported in Acquisti et al. (2014) also show that a majority of Facebook members in a North American city used facial images in their primary profile photos (which are public by default). Facebook data analyzed by Johnson et al. (2012) indicates that approximately 54% of Facebook members made available to strangers at least some of their profile information.

[9] For instance, Jernigan and Mistree (2009) show that a Facebook member's sexual orientation may be inferable from knowledge of her friends network.

[10] See also Drydakis (2009) and Hebl et al. (2002).

[11] The design was institutional review board approved. Like other résumé studies, this experiment involved deception (employers were not informed about the fictional nature of the candidates).

[12] Both the Christian and Muslim profiles are presented as straight. Neither the gay or straight profile provides information about the candidate's religious affiliation.

[13] Under Title VII of the Civil Rights Act of 1964, companies with 15 or more employees may not inquire about a candidate's religious beliefs.

[14] In several, but not all, U.S. states, employers cannot ask job candidates questions regarding their sexual orientations. See U.S. Equal

Employment Opportunity Commission, "Laws enforced by EEOC," accessed June 1, 2017, http://www.eeoc.gov/laws/statutes/.

[15] We also created candidates' accounts on Google+ and Google Sites. These profiles were identical across conditions and did not include professional and candidate-specific information.

[16] Many traditional résumé studies use multiple applications to each employer, but for an exception see Ahmed et al. (2013), who sent one application per employer.

[17] Our candidates were purported to be born in 1982. Mike was the most popular name for male newborns in the 1980s according to U.S. Social Security data (see Social Security Administration, 2015, "Top names of the 1980s," Last modified February, accessed June 1, 2017, http://www.ssa.gov/oact/babynames/decades/names1980s.html) and was assigned to the gay/straight conditions. Adam was the 22nd most popular name and was assigned to the Christian/Muslim conditions, being both a Christian and a Muslim name.

[18] See Pew Research Center (2007), "Muslim Americans: Middle class and mostly mainstream," last modified July 2011, accessed June 1, 2017, http://pewresearch.org/pubs/483/muslim-americans.

[19] The experiment was not started until search engines had indexed the social media profiles. After the experiment began, and on a weekly basis throughout the experiment's duration, we monitored the search results for the candidates' names. We conducted searches through proxy servers so that they appeared as coming from foreign IP addresses, which enabled us to avoid contaminating the actual employer search data that we hoped to gather from services such as Google Keyword Tool trends (see Section 5.2.1 and Appendix B of the e-companion).

[20] See Appendix A of the e-companion.

[21] Research assistants blind to the experimental conditions added up to two technical skills (for instance, Java) or certifications (for instance, CISSP—Certified Information Systems Security Professional—certification) to a résumé if the job description required those skills or certifications as preconditions. This fine-tuning always occurred *before* a candidate's name was randomly added to the résumé and therefore before the candidate was assigned to a job application.

[22] We prepared 10 versions of the same résumés, focusing on slightly different sets of expertise: web development, software development, quality assurance, project or product management, medical/healthcare information, information systems, information security, business intelligence, business development, and analytics. A sample résumé is presented in Table A.2 in the e-companion.

[23] We also created websites, email accounts, and PN presence for some of the companies reported in the candidates' résumés, as well as other workers at those companies and potential letter-writers, to have a complete, believable background should anyone search. The candidates were also registered as alumni from the institution that granted their degrees. As noted, we took down these materials after the completion of the experiment.

[24] We created 41 connected accounts. In 2012, 40% of LinkedIn users had between 51 and 200 connections. See Wayne Breitbarth (2012), "LinkedIn infographic: Want to know what others are doing?" *Power Formula*, last modified March 4, accessed June 1, 2017, http://www.powerformula.net/2347/linkedin-infographic-want-to-know-what-others-are-doing/. Our ability to create a large number of additional connections for our candidates faced technical barriers, such as the need to provide mobile phone numbers to the social network provider to validate newly created profiles.

[25] Johnson et al. (2012) report that only 14.2% of the users of a popular social network whose profile information they mined had a public wall (that is, visible status updates and comments). That noted, we reiterate that the profiles used in the experiment contained equivalent amounts of information across conditions, so that any difference in callback rates could not be merely imputed to a higher amount of disclosure by one type of candidate over another. Furthermore, we note that several fields (such as name, gender, primary photo, profile photo, and networks) are mandatorily public on the network we used for the experiment.

[26] The Muslim candidate's social network profile also presented him as speaking Arabic and English.

[27] Also in the case of the SN, we attempted to create a number of friends that would match the median number of connections of U.S.-based SN profiles at the time the experiment started (approximately 100). However, we were only able to create 49 friends, owing to technical barriers such as (as noted earlier) the need to provide unique mobile phone numbers to create new profiles on the network. As noted, we prevented cross-linkages of the candidates via their common networks of friends by appropriately selecting privacy and visibility settings for the social network profiles.

[28] See Johnson et al. (2012). Although we made the list of friends private, we did not delete the friend profiles because (1) some of those friends' comments appeared on the candidates' profile photos, adding realism to the profiles; (2) the presence of a network of friends decreased the probability that the candidates' profiles could be identified as fake and deactivated by the network (none of the candidate profiles was deactivated during the duration of the experiment).

[29] Technical limitations inherent in designing an experiment via online social networks precluded us from the possibility of randomizing the entire content of the candidates' profiles by creating thousands of different profiles for each candidate. On the other hand, our approach achieves ecological realism by relying on existing profile data.

[30] Table A.4 in the e-companion lists the search terms used to find different type of jobs.

[31] As noted, we could not tolerate multiple candidates' applications being sent to the same firm, because employers may then have been able to find applications and social media profiles with identical photos and overlapping information from different candidates. We also took great care to avoid cross-linkage of the candidates by the same employer through other means. For instance, employers could not navigate from one of the candidates to the other using their (shared) friends, and profile images were subtly modified with random noise (invisible to the human eye) to avoid being listed as similar images by Google's reverse image search functions.

[32] Small firms were excluded from the experiment per agreement with our institution's institutional review board. Note that these firms may not be constrained by EEOC laws, so their exclusion actually may yield more conservative estimates.

[33] See website for ArcGIS: https://www.arcgis.com/features/.

[34] See website for the United States Census Bureau, American Community Survey: https://www.census.gov/programs-surveys/acs/.

[35] See Michael Erwin (2013), "More employers finding reasons not to hire candidates on social media, finds CareerBuilder survey," *CareerBuilder*, accessed June 1, 2017, http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?sd=6%2f26%2f2013&id=pr766&ed=12%2f31%2f2013.

[36] We did not apply to job openings that had been posted for more than 30 days, because some job openings may have been filled or were otherwise inactive. We expect these cases to be uncorrelated with the condition assignments.

[37] This difference is consistent with our estimates of search rates. Consider an employer search rate of 30% (see our search rate estimates in Section 5.2.1). Suppose there are three types of employers: type 1 rejects the application and does not search the candidate online; type 2 considers the candidate further and searches online; and type 3 considers the candidate further but does not search online.

Suppose that 65% of the employers are type 1, 30% are type 2, and 5% are type 3. In this example, the overall search rate is just the prevalence of type 1 employers (i.e., 30%). If we assume that the callback rate by type 3 employers (who do not search) equals the average of the callback rates for the two candidates by type 2 employers (who do search), then we have two equations with two unknowns: $P1 \times 0 + P2 \times M + P3(M + C)/2 = 0.02$ and $P1 \times 0 + P2C + P3(M + C)/2 = 0.17$, where P1, P2, and P3 are the probabilities of employers being types 1, 2, and 3; and M and C are the callback rates for the Muslim and Christian candidates among employers who search. Sample-wide callback rates of 2% for the Muslim candidate and 17% for the Christian candidate can be explained by callback rates (among employers who search) of 2% for the Muslim candidate and 52% for the Christian candidate, and a callback rate when employers consider the candidate but do not search of 27%. It is worth noting that if, in Section 5.2.1, we underestimated actual employers' search rates, or if employers in Republican states have a higher than average search rate, then this example implies a smaller callback bias for employers who search.

[38] We also test for any effect of the application channel (email application versus web application) or the date of the job opening. We do not find any evidence of impact of those variables on callback rates.

[39] For instance, screening and hiring procedures differ significantly across firms, with some firms doing all steps of the hiring process in-house and others outsourcing screening to external firms (including emerging services specializing in online screenings, such as HireRight.com) or mathematical algorithms. Hiring procedures used will influence a given employer's experimental treatment probability. We are agnostic regarding the specific hiring procedures used within a given firm. However, we assume this heterogeneity in screening procedures to be similarly distributed across experimental conditions.

[40] Note that including county-level clusters reduces our sample size. We could not obtain local addresses for some cases in which we did know the state where the job would be located; in addition, there were a handful of states with only one county represented in our data set. In extensive unreported robustness checking, we find that the results are not sensitive to this choice. See Section 5.2.3 as well as Tables A.14–A.19 and Figure A.3 in the e-companion for selected robustness checks.

[41] The Gallup Organization regularly produces lists of the 10 most conservative and 10 most liberal states. The 2012 list that we use was constructed from survey measures of self-reported political party identification in a sample of 321,233 respondents surveyed by Gallup Daily tracking over the course of the year, at the rate of 1,000 respondents per day (Saad 2013).

[42] Our prior Mturk survey suggests a stronger bias among Republicans and older people, and surveys by the Arab American Institute (2012) show stronger bias among Republicans, older people, and those who report not knowing anyone who is Arab or Muslim. In addition to the three interactions suggested by prior results and reported above (with political party support, median age, and fraction Muslim), a fourth and a fifth interaction are suggested by literatures on in-group cooperation and networking in markets. In unreported results (available from the authors on request), we include interactions between the Muslim candidate dummy and (iv) how rural the area is and (v) how much geographical mobility the area has. The political interaction remains significant at the 5% level even after including all five interactions in the same regression, with county-level data and state fixed effects.

[43] When we use the *Romney vote* list instead of the *Combined* list to estimate this specification, there are no noteworthy differences in the results. The Muslim candidate dummy is again statistically significant in Republican states, and the only interactions with the Muslim dummy that are statistically significant are the ones with politically mixed and Democratic states. Put differently, adding additional interactions to the equations estimated in Table 2 has no noteworthy effect on the results.

References

Acquisti A, Gross R (2006) Imagined communities: Awareness, information sharing, and privacy on the Facebook. *Proc. Privacy Enhancing Tech.* 4258:36–58.

Acquisti A, Gross R, Stutzman F (2014) Face recognition and privacy in the age of augmented reality. *J. Privacy Confidentiality* 6(2):1–20.

Acquisti A, Taylor CR, Wagman L (2016) The economics of privacy. *J. Econom. Literature* 54(2):442–492.

Ahmed AM, Hammarstedt M (2009) Detecting discrimination against homosexuals: Evidence from a field experiment on the internet. *Economica* 76(303):588–597.

Ahmed AM, Andersson L, Hammarstedt M (2013) Are gay men and lesbians discriminated against in the hiring process? *Southern Econom. J.* 79(3):565–585.

Arab American Institute (2012) The American divide: How we view Arabs and Muslims. Accessed November 29, 2018, http://b.3cdn.net/aai/4502fc68043380af12_oum6b1i7z.pdf.

Aslund O, Skans ON (2012) Do anonymous job application procedures level the playing field? *ILR Rev.* 65(1):82–107.

Bailey J, Wallace M, Wright B (2013) Are gay men and lesbians discriminated against when applying for jobs? A four-city, internet-based field experiment. *J. Homosexuality* 60(6):873–894.

Banerjee A, Bertrand M, Datta S, Mullainathan S (2009) Labor market discrimination in Delhi: Evidence from a field experiment. *J. Comparative Econom.* 37(1):14–27.

Bertrand M, Mullainathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *Amer. Econom. Rev.* 94(4):991–1013.

Bohnert D, Ross WH (2010) The influence of social networking web sites on the evaluation of job candidates. *Cyberpsych. Behav. Soc. Networking* 13(3):341–347.

Carlsson M, Rooth D-O (2007) Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Econom.* 14(4):716–729.

De La Llama VA, Trueba I, Voges C, Barreto C, Park DJ (2012) At Face(book) value: Uses of Facebook in hiring processes and the role of identity in social networks. *Internat. J. Work Innovation* 1(1): 114–136.

Drydakis N (2009) Sexual orientation discrimination in the labour market. *Labour Econom.* 16(4):364–372.

Ewens M, Tomlin B, Wang LC (2014) Statistical discrimination or prejudice? A large sample field experiment. *Rev. Econom. Statist.* 96(1):119–134.

Garg R, Telang R (2017) To be or not to be linked: Online social networks and job search by unemployed workforce. *Management Sci.* 64(8):3926–3941.

Gift K, Gift T (2014) Does politics influence hiring? Evidence from a randomized experiment. *Political Behav.* 37(3):653–675.

Goldfarb A, Tucker CE (2011) Privacy regulation and online advertising. *Management Sci.* 57(1):57–71.

Goldin C, Rouse C (2000) Orchestrating impartiality: The impact of 'blind' auditions on female musicians. *Amer. Econom. Rev.* 90(4): 715–741.

Harpe LD (2009) Social networks and employment law: Are you putting your organization at risk? Accessed June 23, 2014, http://www.peoplefluent.com/sites/default/files/resources/files/Social_Networks_Employment_Law.pdf.

Hebl MR, Foster JB, Mannix LM, Dovidio JF (2002) Formal and interpersonal discrimination: A field study of bias toward homosexual applicants. *Personality Soc. Psych. Bull.* 28(6):815–825.

Jernigan C, Mistree BFT (2009) Gaydar: Facebook friendships expose sexual orientation. *First Monday* 14(10):https://doi.org/10.5210/fm.v14i10.2611.

Johnson M, Egelman S, Bellovin SM (2012) Facebook and privacy: It's complicated. *Proc. 8th Sympos. Usable Privacy Security* (ACM, New York), Article 9.

Kluemper DH, Rosen PA, Mossholder KW (2012) Social networking websites, personality ratings, and the organizational context: More than meets the eye? *J. Appl. Soc. Psych.* 42(5):1143–1172.

Kroft K, Pope DG (2014) Does online search crowd out traditional search and improve matching efficiency? Evidence from Craigslist. *J. Labor Econom.* 32(2):259–303.

Leip D (2013) Election 2012. Atlas of U.S. presidential elections. Accessed November 29, 2018, http://uselectionatlas.org/2012.php.

Madejski M, Johnson M, Bellovin SM (2012) A study of privacy settings errors in an online social network. *2012 IEEE Internat. Conf. Pervasive Comput. Comm. Workshops* (IEEE, Piscataway, NJ), 340–345.

Manant M, Pajak S, Soulié N (2014) Do recruiters 'like' it? Online social networks and privacy in hiring: A pseudo-randomized experiment. Working paper, Université Paris-Sud, Paris.

Miller AR, Tucker C (2009) Privacy protection and technology diffusion: The case of electronic medical records. *Management Sci.* 55(7):1077–1093.

Mobius MM, Rosenblat TS (2006) Why beauty matters. *Amer. Econom. Rev.* 96(1):222–235.

Narayanan A, Shmatikov V (2009) De-anonymizing social networks. *Proc. 30th IEEE Sympos. Security Privacy* (IEEE, Piscataway, NJ), 173–187.

Neumark D (2012) Detecting discrimination in audit and correspondence studies. *J. Human Resources* 47(4):1128–1157.

Norton EC, Wang H, Ai C (2004) Computing interaction effects and standard errors in logit and probit models. *Stata J.* 4(2):154–167.

Pew Research Center (2012) Two-thirds of Democrats now support gay marriage. Accessed November 29, 2018, http://www.pewforum.org/2012/07/31/2012-opinions-on-for-gay-marriage-unchanged-after-obamas-announcement/.

Pew Research Center (2013) In gay marriage debate, both supporters and opponents see legal recognition as 'inevitable.' Accessed November 29, 2018, http://www.people-press.org/2013/06/06/in-gay-marriage-debate-both-supporters-and-opponents-see-legal-recognition-as-inevitable/.

Posner RA (1981) The economics of privacy. *Amer. Econom. Rev.* 71(2):405–409.

Saad L (2013) In the U.S., blue states outnumber red states, 20 to 12. Accessed February 9, 2016, http://www.gallup.com/poll/160175/blue-states-outnumber-red-states.aspx.

Sanders SD (2011) Privacy is dead: The birth of social media background checks. *Southern Univ. Law Rev.* 39:243–264.

Spence M (1973) Job market signaling. *Quart. J. Econom.* 87(3):355–374.

Sprague R (2011) Invasion of the social networks: Blurring the line between personal life and the employment relationship. *Univ. Louisville Law Rev.* 50(1):1–34.

Sprague R (2012) Facebook meets the NLRB: Employee online communications and unfair labor practices. *Univ. Pennsylvania J. Bus. Law* 14(4):957–1011.

Stigler GJ (1962) Information in the labor market. *J. Political Econom.* 70(5):94–105.

Stigler GJ (1980) An introduction to privacy in economics and politics. *J. Legal Stud.* 9(4):623–644.

Stutzman F, Gross R, Acquisti A (2012) Silent listeners: The evolution of privacy and disclosure on Facebook. *J. Privacy Confidentiality* 4(2):7–41.

Taylor CR, Yildirim H (2011) Subjective performance and the value of blind evaluation. *Rev. Econom. Stud.* 78(2):762–794.

Tilcsik A (2011) Pride and prejudice: Employment discrimination against openly gay men in the United States. *Amer. J. Sociol.* 117(2):586–626.

Vigdor JL (2008) Measuring immigrant assimilation in the United States. Accessed February 9, 2016, http://www.manhattan-institute.org/pdf/cr_53.pdf.

Weichselbaumer D (2003) Sexual orientation discrimination in hiring. *Labour Econom.* 10(6):629–642.

Wright BRE, Wallace M, Bailey J, Hyde A (2013) Religious affiliation and hiring discrimination in New England: A field experiment. *Res. Soc. Stratification Mobility* 34:111–126.