

Abstract - JH:

"Does the gender or race of a potential customer affect response rates when requesting catering orders from U.S. states that historically supported slavery? It is well known that the U.S. fought a civil war over the right to own slaves roughly 160 years ago, and it is widely assumed in popular culture that there remains significant divergence in the experiences of men and women of color in historically slave-owning states. In an audit study employing blocking, we contacted restaurants and caterers sampled randomly across the blocks of states that composed the “Union” and the “Confederacy”, evenly distributing the identity of the messenger by race, and then by gender, with a standardized vague message stating the number of people to be fed with a moderate budget. We found that, in historically slave-owning states compared to other states, catering requests from profiles identifying as men and women of color received response rates at a rate that was 7.5 percentage points lower. This observed difference was statistically significant (e.g. 95% confidence interval: [-9.8, -5.2] percentage points; $p < 0.02$). Additionally, their proposed catering budgets were 15 percentage points less likely to be accepted, a finding also statistically significant (e.g., 95% Confidence Interval: [-17.5, -12.5] percentage points; $p < 0.01$). These results further elucidate the potential contrast in attitudes that exist in the U.S. today. However, we caution that further research is needed to better understand whether response rates may have been impacted by local economic conditions, education, or technology.“

Theory - JH:

The race and gender of the potential customer (the treatment) is expected to cause a change in the measured behavior (response rate and budget acceptance) of the restaurant/caterer. This is grounded primarily in the economic theory of discrimination and the literature on everyday bias in service contexts. An individual employer who discriminates is expected to act as if they incur non-pecuniary, psychic costs, when interacting with or employing members of the disfavored group. In the context of our experiment, the restaurant owner may view the transaction with a customer of color as carrying an imputed cost, which is factored into the transaction’s net cost. The purported higher net cost makes the transaction less desirable, leading to a lower probability of response, or higher likelihood of rejecting the moderate budget.

By blocking on geography, Becker’s (1957) analysis suggests that regional differences in market discrimination are significant, and the persistence of these regional “tastes” could explain the significantly lower response rates observed in the former Confederacy, lending weight to the idea that the treatment’s effect is rooted in established regional biases. The email-based request for a catering order also falls squarely within the concept of “everyday” or “paper cut” discrimination explored by Block et al. (2021). This type of discrimination manifests frequently in simple interactions, such as electronic communications, and is characterized by the differential treatment of individuals based solely on their racial or gender identity.

Concept Under Investigation - JH:

We operationalize discrimination as the difference in response and budget acceptance rates between otherwise identical requests differing only in race/gender signaled by the name. In our field experiment, we measure discrimination in transactional and professional contexts. The most relevant studies are those focusing on market discrimination, subtle everyday bias, and discrimination within the service industry.

Becker (1957) conceptualized discrimination economically as a “taste for” it, where he derived a Market Discrimination Coefficient (MDC) by indirectly using aggregate economic data on wages and income differentials blocked by group and region. Block (2021) focused on everyday treatment, through large-scale field correspondence experiments. They sent randomized email requests and measured the differential response rate to an email requesting voluntary help (taking a survey) based on whether the sender was ostensibly Black or White - while also measuring response quality (survey completion rates). They found White senders received a response, on average, 15.5% higher than the odds for a Black Sender (95% CI: [1.095, 1.219]).

ROC United (2015) examined occupational segregation and discrimination within the food service sector, and studied the implicit bias among employers and consumers, in a matched pair audit test by sending equally qualified applications of different races to fine-dining restaurants to track hiring likelihood. This research implies the existence of bias within the restaurant industry, from a labor perspective. Finally, Bertrand (2004) in a field experiment that sent fictional resumes with names designed to be perceived as either White or Black, measured the callback rate for interviews and segmented their findings as driven by “taste-based” or “statistical” discrimination. The found White names received 50 percent more callbacks for interviews, with a gap that widened with increasing resume quality. For African Americans, having a higher quality resume had a much smaller and statistically insignificant effect on call backs.

We, therefore, leverage the frameworks in the literature to block on geography and randomize identity signals, employing random assignment of sender identities and stratification by region to causally identify the effects of race and gender on service response.

Treatment Variants - DS:

The control variant that our subjects will experience are emails requesting price quotes for catering orders from names that are normally associated with whiteness in social studies (e.g. Emily, Hannah, John). For the treatment variant, subjects will experience emails requesting price quotes for catering orders from names that are normally associated with African American names (e.g. Lakisha, Jamal, Andre). *Ceteris paribus*, the email body remains consistent while only the name of the sender will vary. Moreover, treatments’ email address and phone number may vary (given a set of masked emails and numbers we will rotate through as needed).

Reason About Mechanisms - DS:

Conventional wisdom and experiments in other domains tell us that minority groups are systematically discriminated against, either consciously or unconsciously. This is specifically why we would expect our treatment to lower the response rate and to increase the quoted prices for minority groups. Even though there will be no profile pictures attached to the emails, we anticipate that minorities will be treated more harshly and have a tougher time receiving quotes just because of their name. The mechanism that we are testing is the conscious or subconscious bias that individuals have against minority groups in the United States. In order to explicitly test this mechanism, we could instead consider adding profile pictures to the emails or signatures and instead use the same name for all emails; however, the treatment effect that we think we will uncover is that this mechanism is pervasive even without the visual queue such as a profile picture.

Rough Experiment Design - DS:

In our rough experiment design, we are planning for a between-subject design where the people that we recruit will see either treatment or control because they will receive one of the email templates that we create and not the others. We won’t have any measurement before or after treatment since our measured outcome is a response to the treatment, as opposed to change in effects before or after treatment.

Recruitment - BM:

Restaurants and caterers are to be identified using public online directories and contacted via their website, or email. Sampling will be stratified by region (Union/Confederacy) to ensure balanced representation.

Appendix - JH:

Supplementary Outstanding Items Needing Clarification / TODO - JH:

- Specify how many restaurants/caterers will be contacted, per group and region, to ensure stat power.
- Clarify bounced/undeliverable message handling
- Describe coding scheme for data collection and handling (e.g. logging and statistically comparing response quality, response rates, and budget acceptance)
- Consider pilot test design, resources (e.g. cost of burner addresses, emails, or online hosting), and tooling (e.g. yellow pages, automated data extraction and/or delivery, spreadsheets)
- Detail what will be done if an effect is not detected. Will we check for group imbalance, data quality, measurement strategy, etc.?

Supplementary References - JH:

- [Acharya, A., Blackwell, M., & Sen, M. \(2016\). The political legacy of American slavery. The Journal of Politics.](#)
- [Becker, G. S. \(1957\). The Economics of Discrimination. University of Chicago Press](#)
- [Bertrand, M., & Mullainathan, S. \(2004\). Are Emily and Greg more employable than Lakish and Jamal? A field experiment on labor market discrimination. American Economic Review.](#)
- [Block, Crabtree, Holbein, & Monson \(2021\). Are Americans less likely to reply to emails from Black people relative to White people? Proceedings of the National Academy of Sciences, 118\(52\), e2110347118.](#)
- [Restaurant Opportunities Centers \(ROC\) United, Benner, C., & Food Labor Research Center, University of California, Berkeley. \(2015\). Racial and gender occupational segregation in the restaurant industry: Ending Jim Crow in America's restaurants. University of California, Berkeley Labor Center.](#)