

# Race Causality Discrimination – Audit Studies

Marcello Di Bello - ASU - Fall 2023 - Week #5

Over the past few weeks we discussed the manipulability or interventionist theory of causation.<sup>1</sup> Last week we asked one of the central questions of this course, can RACE<sup>2</sup> be a cause? Statistician Paul Holland thinks that the answer should be negative.<sup>3</sup> This is surprising. Obviously, race seems to play some causal role in society. The problem is, what role is it and how can we study it?

This week we discuss audit studies, a methodology for constructing randomized controlled trials in the social sciences.<sup>4</sup> If randomized controlled trials are the gold standard for studying causation in the medical and biological sciences, can audit studies become the gold standard for studying causation in the social sciences?

We will see that the answer is far from straightforward: causation in the social world—for example, gender, sex and race causation—is likely more complicated than what audit studies suggest.<sup>5</sup>

## Audit Studies

Audit studies, also known as field studies, attempt to reproduce randomized controlled trials in the social sciences. They are often used to study racial and gender discrimination.

The idea is simple, but the execution of these studies can be involved. Here is example of an audit study about racial discrimination in hiring.<sup>6</sup> A bank of resumes is created for different typologies of jobs. The bank contains several resumes divided into high- and low-quality. Resumes of the same quality should be judged by employers as equally good matches for different types of jobs.

For each job ad, a pair of high-quality and a pair of low quality resumes from the bank—so four in total—are sent to the same employer. The only relevant difference between the pairs of resumes is the name: one resume is assigned a typical White name and the other is assigned a typical African-American name. The names are randomly assigned. Callback rates are then compared across resumes associated with White and Black applicants. The callback rates tend to be higher for White than Black applicants. More specifically:

we record a callback rate close to 11 percent for White applicants with a higher-quality resume, compared to 8.5 percent for White applicants with lower-quality resumes. . . African-Americans with higher-quality resumes receive a callback 6.7 percent of the time compared to 6.2 percent for African-Americans with lower quality resumes. (p. 1000-1001)

Is this evidence of racial discrimination? It seems so:

<sup>1</sup> For example, we read about the Rubin model of causation as described in Holland (1986), *Statistics and Causal Inference*, *Journal of the American Statistical Society*, 81(396). We also read some of the philosophical literature, in particular, Woodward's manipulability theory of causation; see James Woodward (2003), *Making Things Happen: A Theory of Causal Explanation*, 2003, Oxford University Press.

<sup>2</sup> Recall Holland's convention of referring to the variable 'race' simply as RACE.

<sup>3</sup> See Holland (2003), *Race and Cause*, *Research Report*, January 2003 RR-03-03, ETS Educational Testing Services and Marcellesi (2013), *Is Race a Cause?*, *Philosophy of Science*, 80(5).

<sup>4</sup> For an overview, see Pager (2007), *The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future*, *The Annals of the American Academy of Political and Social Science*, 609, pp. 104-133.

<sup>5</sup> Hu (2022), *Causation in the Social World*. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences, in particular Chapter 1 and 2.

<sup>6</sup> See Bertrand and Mullainathan (2004), *Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*, *The American Economic Review*, 94(4), pp. 991-1013.

Because names are randomized, the White and African-American resumes we send should rank similarly on average (p. 1006)

There could be alternative explanations, but overall this seems compelling evidence of racial discrimination.<sup>7</sup>

Audit studies have also been sharply criticized. Economist James Heckman is one of the most vocal critics:

The validity of an audit study relies on its success in presenting two otherwise equally qualified job applicants who differ only by race. Given the vast number of characteristics that can influence an employer's evaluation, however, it is difficult to ensure that all such dimensions have been effectively controlled.<sup>8</sup>

### *The Audit Study Puzzle*

Audit studies mimic the randomized controlled trial even in the social sciences where observational studies are more common. The assumption is that randomization represents the gold standard for the study of causation. The implicit picture of causality, then, is that a cause *X* is what makes a difference to an effect *Y*. To provide evidence that *X* causes (i.e. makes a difference to) *Y*, researchers should show that a surgically induced change in *X* (assuming possible confounders are controlled for) makes a difference to *Y*.

Is this interventionist picture of causality adequate for modeling the causal role of social categories such as race and gender? This pair of vignettes by Lily Hu<sup>9</sup> raises a puzzle for the interventionist picture of causality (and for audit studies, as well):

AUDIT STUDY I: Skirt Interview - Two actors, one taken to be male and the other female, present identical resumes, answer interviewer questions identically, and affect the same tone, mannerisms, and general personality traits (as best as they can). The male actor also dons the same dress and wears the same facial makeup as the female actor; both actors wear skirts and facial makeup to their interviews. (p. 86)

AUDIT STUDY II Confident Men and Mild-Mannered Women - Two actors, one taken to be male and the other female, present identical resumes and answer interviewer questions identically. To avoid confounding by perception of gender nonconforming status that may be triggered by, for example, setting identical styles of dress and facial makeup across the auditors, the social scientists look to make sure that both actors display traits that they take to be gender-conforming. For example, though the male actor presents as a confident and assertive candidate, the social scientists have the female actor portray as mild-mannered and demure, as a part of the effort to maintain what they take to be gender-conforming affect. (p. 87)

These are imaginary studies. The first audit study ensures that the two interviewees are perfectly identical except for whether they are

<sup>7</sup> Here is an alternative explanation: "perhaps the average firm in our experiment aims to produce an interview pool that matches the population base rate. This rule could produce the observed differential treatment if the average firm receives a higher proportion of African-American resumes" (p. 1006). How do the authors of the study respond to this alternative explanation?

<sup>8</sup> This is a citation from page 114 of Pager (2007). How does this criticism, exactly, apply to the resume audit study?

<sup>9</sup> Hu (2022), *Causation in the Social World*. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences, Chapter 2.

taken to be male or female. This is what we would expect the perfect randomized control trial to do: have two otherwise identical units that differ only by the social category of interest, in this case sex. But oddly, something in this study is off:

One actor in a skirt portrays a gender conforming female job candidate; the other, a gender non-conforming male candidate. And this difference, we know, matters for how the interviewees will be received by the interviewer. (p. 86)

So, by controlling for all the *intrinsic* causes, the experimenter failed to control for *extrinsic* (relational) causes. Whether someone is perceived as gender conforming or not depends on the interaction between attire and gender category.

The second audit study attempts to contain the spillover effects of gender non-conforming. Thus, male and female interviewees are both portrayed as gender-conforming. But oddly, this study is also off. So many characteristics have been adjusted for to ensure the interviewees are both gender conforming. It is unclear whether the study can say anything about the causal role of sex.<sup>10</sup>

### *What the causal interventionist could say*

Perhaps, the moral here is simply that—from an interventionist account of causation—what it means to intervene on sex is unclear. Until we know what sex is, it is unclear how we should intervene on sex:

causation is a matter of what would happen under an interventionist counterfactual, where the target of manipulation, the category of sex, is a category about which there is substantial disagreement. Should sex in the social world refer only to the possession of certain sex organs? Should sex status count visible secondary sex characteristics? Should it also encompass gender stereotypes? If so, which ones? (p. 103)

The assumption here is that we can know what sex is *before* we embark in a study of its causal role. But can we really? <sup>11</sup>

The interventionist could also suggest that we disentangle the causal role of sex from the causal role of other social features, such as manners of behavior, childcare responsibilities or attire. But suppose that researchers can actually factor out the causal role of other social features and isolate the causal role of sex. Whatever the result of this study, it would fail to explain the causal role of gender:

An account that treats the causal operation of sex as independent of the causal operation of skirts and of childcare responsibilities will fail to illuminate what about *gender*, rather than skirts and the existence of dependents as such, produces the outcome (p. 112).<sup>12</sup>

<sup>10</sup> Another interesting vignette: "Suppose Daniel and Eunice apply to a position, presenting identical resumes, answering interviewer questions identically, and affecting the same tone, mannerisms, and general personality traits. Over the course of their interviews, each reveals that they are expecting a child" (p. 107). How does Hu discuss this case and what morals follow from it?

<sup>11</sup> Hu thinks that ultimately we cannot: "This brings us to what I take to be the crux of the matter: whether an account of what sex is and an account of the causal role that sex plays in the social world can be disentangled at all." (p. 104)

<sup>12</sup> If we disaggregate "sex causality" from "attire causality", the resulting account would be "unable to guide us in seeing what is socially unjust about the fact that some people wearing skirts are penalized in interview processes, whereas others are expected to wear skirts at their interviews" (p. 112).

### *Discrimination before causality*

So, then, should we do away with audit studies? Are they pointless? Not at all. Hu thinks they provide powerful evidence of discrimination:

Audit studies thereby furnish evidence for discrimination directly . . . by first nominating a standard for what constitutes fair treatment across differently raced or sexed counterparts and then interpreting a failure to meet the standard as evidence of discrimination on the basis of race or sex—a case of wrongful causal influence of race and sex. (p. 116)

So, for Hu, audit studies provide direct evidence of discrimination and indirectly evidence of the causal role of race/sex in society. The crux here is that the claim of discrimination does not rest on a causal claim. It is actually the other way around.<sup>13</sup>

On this interpretation, the standard criticism—a la Heckman—that audit studies do not account for unobservable variables is moot. Whether two resumes *should* receive the same callback rates is a normative judgment that is presupposed in the set up of the audit study itself. Once we agree with this moral judgment—and we might very well not agree—any deviation from the set moral standard is evidence of wrongful racial discrimination.

Since agreement on the normative standard is a precondition for audit studies to provide evidence of discrimination, Hu provocatively suggests that we consider the following:

Consider a case in which Jamal has *less* work experience than Greg. One might take it that under these circumstances, Jamal *still* ought not face such substantially dimmer prospects than Greg when looking for employment. . . . Of course, supporting this claim requires normative argument, but so too does . . . [the claim that] . . . identical resumes should receive equal callback outcomes—the only difference being that in that case, the normative grounds strike most as so obvious that it is fine to leave them implicit. (p. 120)

Whether two resumes are “matched” for the purpose of the audit study—resumes that should receive the same callback rates—is a moral question, and only when that question is settled, an audit study can provide evidence of wrongful discrimination.

<sup>13</sup> Hu writes, intriguingly: “an audit study need not make any claim to causation in order to substantiate a claim of discrimination, because discrimination is not an essentially causal notion” (p. 117).