Correction Appendix to "Racial Bias in Bail Decisions"

David Arnold[*]   Will Dobbie[†]   Crystal S. Yang[‡]

September 2020

This document makes precise the formal definition of racial bias in our article "Racial Bias in Bail Decisions" published in the *Quarterly Journal of Economics* in November 2018. Our paper defines a judge as racially biased if her decisions cannot be solely explained by accurate statistical discrimination. Therefore, a judge is racially biased if she perceives a higher threshold of release for black defendants than white defendants at the margin, or under an alternative model, if she overestimates the cost of release for black defendants relative to white defendants at the margin. We refer to this verbal definition repeatedly throughout the paper. However, our formal definition of bias was insufficiently precise as to our intended definition of bias, which we realized in light of a working paper by Canay, Mogstad, and Mountjoy (2020).

To make our intended definition of bias clear, we make the following amendments to the published paper, where page numbers refer to the published version:

(1) On p. 1893, at the end of the paragraph beginning "The perceived benefit of release for defendant i..." we add the following definitions: "Let the non-race characteristics of the marginal defendant for judge j and race r be denoted $\mathbf{V}_{i,r}^*$. We correspondingly define $t_r^{j*} = t_r^j(\mathbf{V}_{i,r}^*)$."

(2) On p. 1893, Definition 1 should be: "DEFINITION 1. Following Becker (1957, 1993), we define judge j as racially biased against black defendants if $t_W^{j*} > t_B^{j*}$. Thus, for racially biased judges, there is a higher perceived benefit of releasing white defendants than black defendants at the margin."

(3) On p. 1894, at the end of the sentence beginning "Given this decision rule...," $t_r^j(\mathbf{V}_i)$ should be $t_r^j(\mathbf{V}_{i,r}^*)$ and at the end of the sentence beginning "We simply our notation...," α_r^j should be $\alpha_r^j = \mathbb{E}[\alpha_i^j | \mathbf{V}_i = \mathbf{V}_{i,r}^*, r_i = r]$.

(4) On p. 1895, Definition 2 should be: "DEFINITION 2. We define judge j as making racially biased prediction errors in risk against black defendants if $\tau_W^j(\mathbf{V}_i = \mathbf{V}_{i,W}^*) > \tau_B^j(\mathbf{V}_i = \mathbf{V}_{i,B}^*)$. Thus, judges making racially biased prediction errors systematically overestimate the true cost of release for black defendants relative to white defendants at the margin."

(5) In Equations (4), (5), (6), (8) and any discussion of these equations, t_r^j should be t_r^{j*}.

---
[*]UC San Diego. Email: daarnold@ucsd.edu

[†]Harvard Kennedy School and NBER. Email: will_dobbie@hks.harvard.edu

[‡]Harvard Law School and NBER. Email: cyang@law.harvard.edu

(6) On p. 1922, in the sentence beginning with "Bail judges could, for example, harbor..." the phrase "observably similar" should be struck.