

# Racial Disparities in Police Use of Deadly Force Against Unarmed Individuals Persist After Appropriately Benchmarking Shooting Data on Violent Crime Rates

Social Psychological and  
Personality Science  
1-10

© The Author(s) 2020



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/1948550620916071

[journals.sagepub.com/home/spp](https://journals.sagepub.com/home/spp)



Cody T. Ross<sup>1</sup> , Bruce Winterhalder<sup>2</sup> , and Richard McElreath<sup>1</sup>

## Abstract

Cesario et al. argue that benchmarking the relative counts of killings by police on relative crime rates, rather than relative population sizes, generates a measure of racial disparity in the use of lethal force that is unbiased by differential crime rates. Their publication, however, lacked any formal derivation showing that their benchmarking methodology has the statistical properties required to establish such a claim. We use the causal model of lethal force by police conditional on relative crime rates implicit in their analyses and prove that their benchmarking methodology does not, in general, remove the bias introduced by crime rate differences. Instead, it creates strong statistical biases that mask true racial disparities, especially in the killing of unarmed non-criminals by police. Reanalysis of their data using formally derived criminality-correcting benchmarks shows that there is strong and statistically reliable evidence of anti-Black racial disparities in the killing of unarmed Americans by police in 2015–2016.

## Keywords

racial bias, police shootings, use-of-force, inequality, causal inference

Previous studies and data sets have shown that police in the United States kill Black citizens relative to White citizens at higher rates than would be expected under a generative model in which police encounter and kill Black and White citizens in proportion to their relative population sizes (e.g., Gabrielson et al., 2014; *The Guardian*, 2016; Takagi, 1981). The usefulness of these studies for identifying unjustifiable racial disparities in police behavior, however, has been called into question (Cesario et al., 2019; Fryer, 2017; Selby et al., 2016; Tregle et al., 2019) because police primarily kill individuals—Black or White—who were armed and engaging in criminal activity at the time of the interaction (Ross, 2015; Selby et al., 2016). Underlying differences in race-specific rates of armed criminal activity, rather than—or in addition to—prejudice and/or unintended stereotype bias (Payne, 2006) by police, have therefore been cited as a possible causal driver of the elevated rates of police shootings of Black Americans. Nevertheless, anti-Black disparities in police use-of-force against unarmed individuals persist at both the nonlethal (Fryer, 2016) and lethal (Ross, 2015) level of force. Conditional on being killed by police, Black compared to White decedents are also less likely to have been armed (Charbonneau et al., 2017), so armed status alone does not appear to fully explain racial disparities in the use of force by police.

It is, however, challenging to disentangle unjustifiable racial disparities in police use-of-force from the disparities that might

emerge from justifiable responses by police to differential levels of crime—especially because police use-of-force data are observational and frequently lack impartial contextualizing information. Several approaches have been used to control for this important covariate, including direct study of police-on-police shootings (Charbonneau et al., 2017), multivariate regression analyses (e.g., Hannon, 2020; Ross, 2015; Scott et al., 2017), use of multiple crime rate—“correcting” benchmarks (e.g., Cesario et al., 2019; Miller et al., 2017; Tregle et al., 2019), and the use of encounter-conditional analyses (e.g., Fryer, 2016; Johnson et al., 2019; Worrall et al., 2018). Statistical problems with encounter-conditional approaches have been addressed already with causal models (e.g., Knox et al., 2019; Ross et al., 2018); essentially, if a subset of the police population are biased in how frequently they engage in encounters with individuals as a function of race or ethnicity—for example, with a lower threshold of suspicion leading

<sup>1</sup>Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

<sup>2</sup>University of California, Davis, CA, USA

## Corresponding Author:

Cody T. Ross, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig 04103, Germany.

Email: [cody\\_ross@eva.mpg.de](mailto:cody_ross@eva.mpg.de)

police to encounter Black individuals—then the results of encounter-conditional analyses will be confounded.

Formal theoretical analysis of the benchmarking methodology advanced by Cesario et al. (2019), however, has yet to be done. Cesario et al. argue that “benchmarking” the race-specific counts of killings by police on relative crime counts, rather than relative population sizes, generates a measure of racial disparity in the use of lethal force by police that is not statistically biased by differential crime rates. In their words, “if different groups are more or less likely to occupy those situations in which police might use deadly force, then a more appropriate benchmark as a means of testing for bias in officer decision making is the number of citizens within each race who occupy those situations during which police are likely to use deadly force” (p. 587). In other words, they aim to produce estimates of killing rates by police unique to the interaction of suspect race/ethnicity and criminal status and test for evidence of racial disparity holding constant the relative sizes of the criminal populations. Their publication, however, lacks any formal derivation showing that their benchmarking methodology has statistical properties consistent with their conceptual objectives.

In the following sections, we use the causal model of police use-of-force conditional on criminality implicit in Cesario et al. (2019) and attempt to derive an unbiased measure of police use-of-force using their benchmarking methodology. We first prove that their benchmarking methodology does not remove the bias introduced by crime rate differences but rather creates potentially stronger statistical biases that mask true racial disparities, especially in the killing of unarmed noncriminals by police. We then derive a benchmarking approach that does remove the effect of crime rate differences on estimates of racial disparity in killings by police and use this approach to reevaluate their empirical findings. Using these criminality-correcting benchmarks, we show that there is statistically reliable evidence of anti-Black racial disparities in the killing of unarmed, nonaggressing civilians by police in both 2015 and 2016.

### A Causal Model of Police Shootings as a Function of Criminality

Following the implicit generative model in Cesario et al. (2019), we can theoretically investigate the role of differential crime rates on the apparent level of racial disparities in killings by police.

Assume we have a total population of  $P_w$ , White individuals and  $P_B$ , Black individuals. And then, let us assume that with probabilities,  $\alpha_w$  and  $\alpha_B$  individuals in each respective subpopulation acquire weapons and engage in violent criminal behavior. The population will then be composed of a faction of armed criminals,  $C_w$  and  $C_B$ , and unarmed noncriminals,  $N_w$  and  $N_B$ :

$$C_w \sim \text{Binomial}(P_w, \alpha_w) \quad (1)$$

$$C_B \sim \text{Binomial}(P_B, \alpha_B) \quad (2)$$

where,

$$N_w = P_w - C_w \quad (3)$$

$$N_B = P_B - C_B \quad (4)$$

Over some interval of time, each person in the population has a probability of being encountered and killed by police. That probability varies by both race and criminal status. The parameter  $\phi$  indicates the probability of police killing an armed criminal, and the parameter  $\theta$  indicates the probability of police killing an unarmed noncriminal. Thus, in each race/ethnic group, there will be  $A$  shootings of armed criminals and  $U$  shootings of unarmed noncriminals:

$$A_w \sim \text{Binomial}(C_w, \phi_w) \quad (5)$$

$$A_B \sim \text{Binomial}(C_B, \phi_B) \quad (6)$$

and

$$U_w \sim \text{Binomial}(N_w, \theta_w) \quad (7)$$

$$U_B \sim \text{Binomial}(N_B, \theta_B) \quad (8)$$

For any realization from this generative model, anti-Black racial disparity in the overall probability of being killed by police through *either causal path* is present if the following inequality is satisfied:

$$\frac{A_B + U_B}{P_B} > \frac{A_w + U_w}{P_w} \quad (9)$$

We can investigate this expression in more detail by taking the expectations of each side:

$$\mathbb{E}\left[\frac{A_B + U_B}{P_B}\right] > \mathbb{E}\left[\frac{A_w + U_w}{P_w}\right] \quad (10a)$$

$$\mathbb{E}[A_B + U_B] \frac{1}{P_B} > \mathbb{E}[A_w + U_w] \frac{1}{P_w} \quad (10b)$$

$$(\mathbb{E}[A_B] + \mathbb{E}[U_B]) \frac{1}{P_B} > (\mathbb{E}[A_w] + \mathbb{E}[U_w]) \frac{1}{P_w} \quad (10c)$$

$$\phi_B \alpha_B + \theta_B (1 - \alpha_B) > \phi_w \alpha_w + \theta_w (1 - \alpha_w) \quad (10d)$$

$$\frac{\phi_B \alpha_B + \theta_B (1 - \alpha_B)}{\phi_w \alpha_w + \theta_w (1 - \alpha_w)} > 1 \quad (10e)$$

$$\frac{\alpha_B (\phi_B - \theta_B) + \theta_B}{\alpha_w (\phi_w - \theta_w) + \theta_w} > 1 \quad (10f)$$

Equation 10a reduces to the ratio of the probabilities of being killed by police, for Black individuals relative to White individuals, over both causal paths (i.e., engaging in criminal activity and being killed, or not engaging in criminal activity and being killed); this fact is reflected in the numerator and denominator of Equation 10e being convex combinations of the killing probability parameters for criminals,  $\phi$ , and

noncriminals,  $\theta$ , where the mixing simplex is determined by the crime rate,  $\alpha$ .

For addressing some research questions, this measure can be useful, as it shows the overall level of racial disparity in police use-of-force as it would be experienced by the Black and White communities, including through causal paths like differing levels of poverty and marginalization that might lead to differing levels of criminality. However, for evaluating the behavior of police, this measure is clearly problematic.

Differential crime rates can easily confound the measure given in Equations 10a–10f, severely limiting its usefulness—a key observation of Cesario et al. (2019). For example, assume, for illustrative purposes, that  $\phi_W = \phi_B$  and  $\theta_W = \theta_B$  and that  $\phi$ —the probability of police killing a criminal—is higher than  $\theta$ —the probability of police killing a noncriminal. In such a case, if  $\alpha_B$  were greater than  $\alpha_W$ , then the measure given in Equations 10a–10f would indicate the existence of anti-Black racial disparity in killings by police even if officers treated Black and White individuals—be they criminals or non-criminals—exactly the same. As such, Equations 10a–10f cannot be used to evaluate the appropriateness of policing behavior.

### A Bias Correction?

Cesario et al. (2019), reacting to studies using measures like that presented in Equation 10a to infer racial disparities, argue that the benchmarking of killing rates by police should be done using the count of criminals,  $\alpha P$ , not population size,  $P$ , per se. In other words, they argue that the metric of interest from Equation 9 should be written instead as:

$$\frac{A_B + U_B}{\alpha_B P_B} \stackrel{?}{>} \frac{A_W + U_W}{\alpha_W P_W} \quad (11)$$

in order to adjust for the effects of crime rate variation on *apparent* racial disparities in killings by police. However, there are problems with this approach.

Specifically, the bias-correcting benchmark used by Cesario et al. (2019) *does not actually yield unbiased estimates* except in unrealistic edge cases (discussed below). Evaluating the expectations of Equation 11 yields:

$$\mathbb{E}\left[\frac{A_B + U_B}{\alpha_B P_B}\right] \stackrel{?}{>} \mathbb{E}\left[\frac{A_W + U_W}{\alpha_W P_W}\right] \quad (12a)$$

$$(\mathbb{E}[A_B] + \mathbb{E}[U_B]) \frac{1}{\alpha_B P_B} \stackrel{?}{>} (\mathbb{E}[A_W] + \mathbb{E}[U_W]) \frac{1}{\alpha_W P_W} \quad (12b)$$

$$\phi_B + \theta_B \frac{(1 - \alpha_B)}{\alpha_B} \stackrel{?}{>} \phi_W + \theta_W \frac{(1 - \alpha_W)}{\alpha_W} \quad (12c)$$

$$\frac{\phi_B + \theta_B \frac{(1 - \alpha_B)}{\alpha_B}}{\phi_W + \theta_W \frac{(1 - \alpha_W)}{\alpha_W}} \stackrel{?}{>} 1 \quad (12d)$$

Equations 12a–12d do not, in general, yield an unbiased

estimate of either  $\frac{\phi_B}{\phi_W}$  or  $\frac{\theta_B}{\theta_W}$  and no longer reduce to the ratio of the probabilities of being killed by police for Black relative to White individuals over both causal paths. The numerator and denominator of Equation 12c are also no longer convex combinations of the killing probability parameters, making interpretation difficult. Equation 12c can yield an unbiased estimate of  $\frac{\phi_B}{\phi_W}$  only in unrealistic edge cases in which police never kill unarmed individuals of either race/ethnic group (i.e.,  $\theta_B, \theta_W = 0$ ) and/or when the population is composed purely of criminals (i.e.,  $\alpha_B, \alpha_W = 1$ ).

The validity of the Cesario et al. (2019) benchmarking methodology depends on the strong assumption that police never kill innocent, unarmed people of either race/ethnic group. While it is true that deadly force is primarily used against armed criminals who pose a threat to police and innocent bystanders (e.g., Binder & Fridell, 1984; Binder & Scharf, 1980; Nix et al., 2017; Ross, 2015; Selby et al., 2016; White, 2006), it is also the case that unarmed individuals are killed by police at rates that reflect racial disparities. Ross (2015) and Charbonneau et al. (2017), for example, show that conditional on being shot by police, a White suspect is more likely to be armed than is a Black suspect. Even unarmed noncriminals face the risk of being killed by police, and so, the relative population sizes of noncriminals cannot simply be ignored when assessing racial disparities in killings by police.

### Bias Corrections

As shown above, when analyzing data pooled from armed and unarmed suspects, it is hard to generate an estimate of racial disparity in police use-of-force that is not confounded by relative crime rates. However, if we assume that those individuals who were armed when killed by police come from the criminal subpopulation and those individuals who were unarmed come from the noncriminal subpopulation, then we can derive unbiased measures by using crime rate data as a benchmark. This property is true by assumption in our causal model, but it is unlikely to exactly hold in empirical data. Nevertheless, there is likely to be a strong correlation, whereby a citizen who is armed and killed by police is more likely to be a member of the violent criminal subpopulation than the noncriminal population.

**Criminals.** If we condition on armed status (and consider only those incidents where armed individuals are killed by police) and then benchmark on race-specific population sizes, we get the measure:

$$\mathbb{E}[A_B] \frac{1}{P_B} \stackrel{?}{>} \mathbb{E}[A_W] \frac{1}{P_W} \quad (13a)$$

$$\phi_B \alpha_B \stackrel{?}{>} \phi_W \alpha_W \quad (13b)$$

$$\frac{\phi_B \alpha_B}{\phi_W \alpha_W} \stackrel{?}{>} 1 \quad (13c)$$

The unbiased relative probability of police killing a violent criminal is given by  $\frac{\phi_B}{\phi_W}$ . So, the above measure is indeed biased exactly by  $\frac{\alpha_B}{\alpha_W}$ . Thus, applying the relative crime rate benchmark of Cesario et al. (2019)—that is, multiplying the left-hand side of Equation 13c by  $\frac{\alpha_W}{\alpha_B}$ —will yield an unbiased measure of  $\frac{\phi_B}{\phi_W}$ . For this correction to hold, however, we must first condition on the status of suspects as armed, before estimating the parameters of interest.

**Noncriminals.** On the other hand, if we consider only those incidents where unarmed individuals are killed by police and benchmark on race-specific population sizes, we get the measure:

$$\mathbb{E}[U_B] \frac{1}{P_B} \stackrel{?}{>} \mathbb{E}[U_W] \frac{1}{P_W} \quad (14a)$$

$$\theta_B(1 - \alpha_B) \stackrel{?}{>} \theta_W(1 - \alpha_W) \quad (14b)$$

$$\frac{\theta_B(1 - \alpha_B)}{\theta_W(1 - \alpha_W)} \stackrel{?}{>} 1 \quad (14c)$$

The relative probability of police killing an unarmed non-criminal is given by  $\frac{\theta_B}{\theta_W}$ . So, the above measure is biased by  $\frac{1 - \alpha_B}{1 - \alpha_W}$ . In this case, applying the relative crime rate benchmark of Cesario et al. (2019)—that is, multiplying the left-hand side of Equation 14c by  $\frac{\alpha_W}{\alpha_B}$ —will not yield an unbiased measure of  $\frac{\theta_B}{\theta_W}$ ; it will yield a *more biased* measure,  $\frac{\theta_B(1 - \alpha_B)\alpha_W}{\theta_W(1 - \alpha_W)\alpha_B}$ ; one that has no clear interpretation. Instead, to generate an unbiased estimate of the relative rate of police killing unarmed noncriminals, one needs to multiply by the ratio of *noncriminality*,  $\frac{1 - \alpha_W}{1 - \alpha_B}$ , not the ratio of criminality.

Cesario et al. (2019), however, apply the same benchmark of  $\frac{\alpha_W}{\alpha_B}$  to all of their data sets—even those consisting of police shootings of unarmed, nonaggressing civilians. This leads to incorrect estimates of the quantities they claim to identify. Using the National Crime Victimization Survey (NCVS) violent crime data from 2016 to define the  $\alpha$  parameters in the above equation, the bias introduced by the Cesario et al. (2019) methodology would result in multiplying the true anti-Black racial disparity by a scalar of approximately 0.38. This means that if crime rate differences in our theoretical model were as we find empirically, then even if we set the killing probability parameters in our causal model such that unarmed, noncriminal, Black individuals were 2.6 times more likely to be killed by police than unarmed, noncriminal, White individuals, the Cesario et al. (2019) methodology would suggest no racial disparities!

Below, we use the same data sources and statistical workflow as Cesario et al. (2019), but we break down the analysis by the status of the suspect as armed or unarmed and then apply the correct crime rate adjustment benchmark to each subpopulation. In this way, we replicate the premise of Cesario et al. (2019)—that population-level estimates of the relative risk of being killed by police can be confounded by differential rates

of criminality—but correct statistical shortcomings in their methodology.

## Method

### Data Sources

Data on the killing of civilians by police are taken from “The Counted,” an online database managed by *The Guardian* (2016). As Cesario et al. (2019) state, this database is more complete than official federal databases, as police departments underreport to the federal government (Feldman et al., 2017; Klinger et al., 2016; Nix et al., 2017; White, 2006). We use data directly from *The Guardian* (2016) in our workflow, along with the “unarmed and nonaggressing” data as presented in Cesario et al. (2019).

Following the methods described in the supplementary information of Cesario et al. (2019), we ask whether Black individuals or White individuals are more likely to be killed when benchmarking police killing data on three classes of criminal report data from 2015 to 2016: violent crimes (major), violent crimes (minor), and weapons violations. We estimate the criminal activity rates of Black and White individuals from two sources: (1) the National Incident-Based Reporting System (NIBRS; Federal Bureau of Investigation [FBI], 2016) and (2) NCVS (Bureau of Justice Statistics, 2016), as these two data sources allow for extrapolation to population-level counts. As described in Cesario et al. (2019), the NIBRS is a federal database of incidents submitted by law enforcement to the FBI (but compliance is variable and may be nonrandom), and the NCVS is a nationally representative self-report survey of criminal victimization.

The NIBRS data used herein were taken directly from Cesario et al. (2019), and the NCVS data used herein were downloaded from their official source and processed according to the methods described in Cesario et al. (2019). We used the original source data in order to take advantage of the sampling frame and weighting variables provided therein.

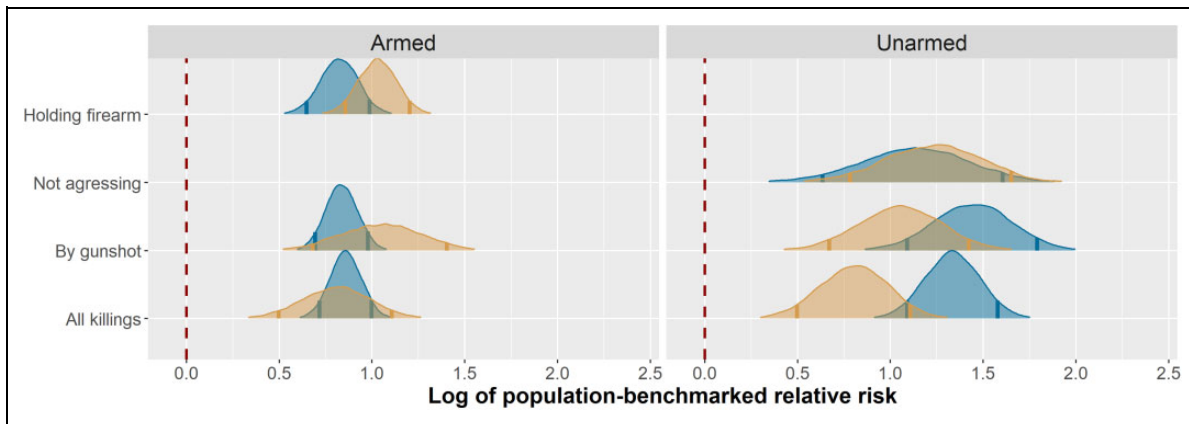
### Statistical Modeling

We use the generative stochastic models described previously as Bayesian statistical models (McElreath, 2018) coded in Stan (Carpenter et al., 2017; Stan Development Team, 2018b) and fit using R (R Core Team, 2018) and rstan (Stan Development Team, 2018a). Our complete workflow will be maintained on GitHub at <https://github.com/ctross/disparitiesandbenchmarks>.

## Analysis of Data From “The Counted”

### Population-Level Relative Risk

We first investigate the relative risk of being the victim of a police killing using standard population size benchmarks. Figure 1 plots population-level relative risk of being the armed or unarmed victim of a police killing using data from “The



**Figure 1.** Analysis of data on killings by police from “The Counted,” 2015 (blue) and 2016 (orange), using benchmarking on population size. Note. Density curves show the natural log of the posterior distributions of the relative probability (for Black individuals relative to White individuals) of being killed by police. Values greater than 0 indicate anti-Black racial disparity, and values less than 0 indicate anti-White racial disparity. Central 90% credible intervals are represented by darker vertical bars on the distributions. Across armed and unarmed subpopulations and encounter types, we find robust evidence of anti-Black racial disparities in killings by police. Panel positions without curves indicate categories not considered in our analysis.

Counted” (*The Guardian*, 2016); estimates are subdivided by year—2015 and 2016—and by encounter type. Encounter type categories include the following: (1) “All killings” which refers to all deaths by police, whether by shooting, death in custody, or other means; (2) “By gunshot” which refers to deaths caused by police gunfire; (3) “Not aggressing” which refers to deaths caused by police gunfire against unarmed and nonaggressing civilians—nonaggressing coding was done by Cesario et al. (2019); and finally (4) “Holding firearm” which refers to deaths caused by police gunfire against civilians who were themselves armed with a firearm.

Across all years, encounter types, and armed status categories, police are more likely to kill Black citizens than White citizens. However, these estimates will be affected to an unknown degree by relative crime rates (Equation 10e); to remove the effect of differential crime rates, we now apply the corrections derived in Equations 13c and 14c.

### Crime-Benchmarked Relative Risk, Armed Suspects

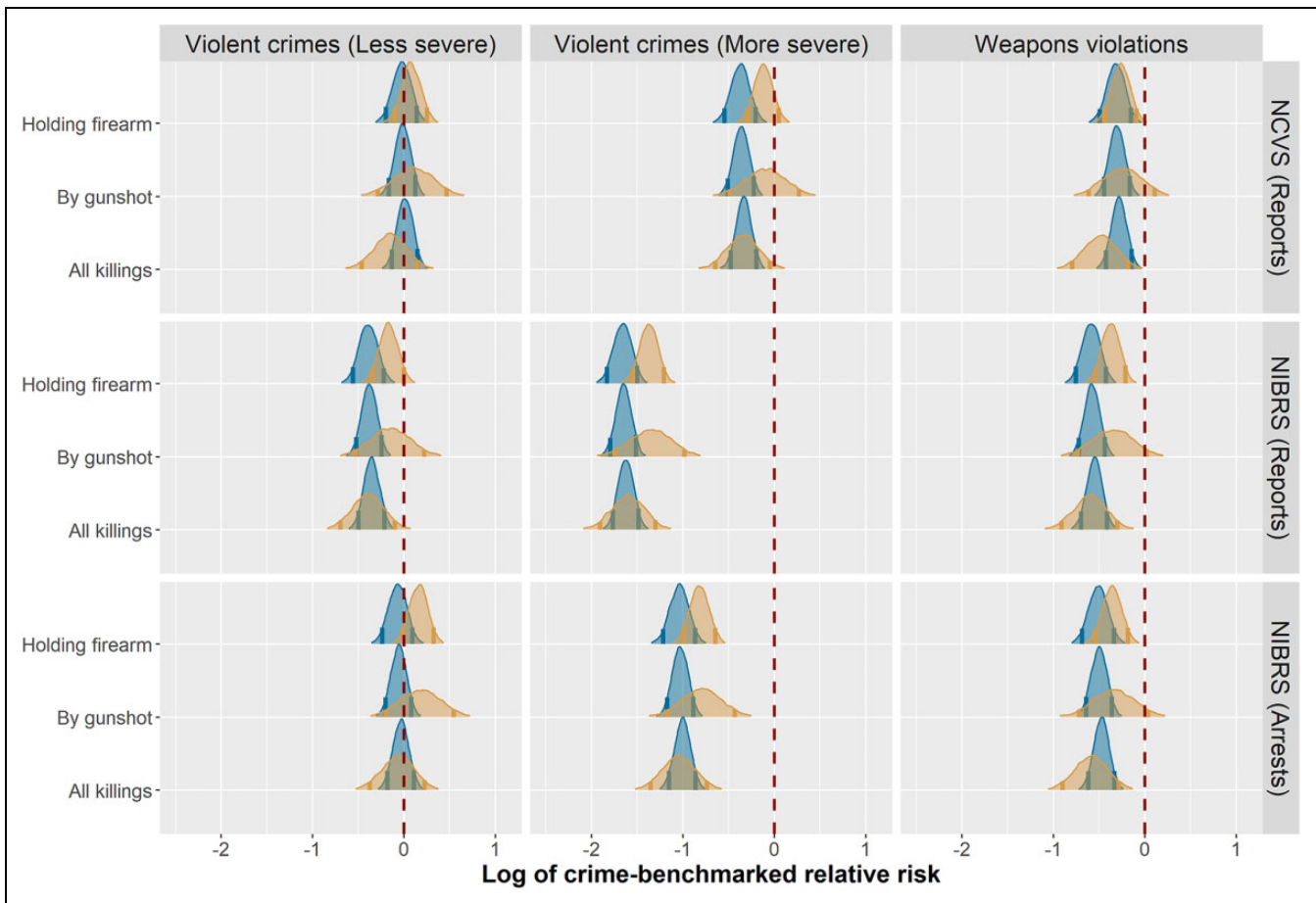
For the case of armed individuals killed by police, we apply the benchmark derived in Equation 13c to the previous estimates of racial disparities in police killings and recover an unbiased estimate of  $\frac{\theta_B}{\theta_W}$ , the relative risk of police engaging in what are normally classified as justifiable killings of armed criminals. Figure 2 plots these estimates. As expected for this subset of police killings, we recover a principle finding of Cesario et al. (2019): Racial disparities in the killing of armed suspects by police are proportional to the relative rates of violent criminality.

Our results highlight disparities between NCVS and NIBRS data sources. NCVS data suggest almost perfect proportionality between relative violent crime rates and relative police killing rates of armed suspects, whereas data from the NIBRS suggest that White suspects are killed by police at

greater rates than expected relative to their violent crime rates. This might be a true empirical pattern, or it might reflect racial disparity in the reporting of NIBRS data. Cesario et al. (2019) argued that NIBRS data were unlikely to be affected by such biases and that “NCVS [data] are uncontaminated by police bias, yet . . . yield results consistent with the . . . NIBRS data” (p. 588). We, however, find it important to acknowledge the possibility that NIBRS data are biased by reporting. We contrast the relative crime rates calculated using the NCVS and the NIBRS data sets in Figure 3. Here, we find that across almost all years and crime classifications, NIBRS data show greater racial differences in crime rates than NCVS data. NCVS data are based on a randomized sampling design, making them less likely to be biased by differential policing intensity and reporting. Benchmarking approaches based on NIBRS data may underestimate the extent of anti-Black disparities in police use-of-force.

### Crime-Benchmarked Relative Risk, Unarmed Suspects

If we apply the benchmark derived in Equation 14c to the previous estimates of racial disparities in the killing of unarmed individuals by police, we can recover an unbiased estimate of  $\frac{\theta_B}{\theta_W}$ , the relative risk of police killing unarmed individuals. Figure 4 shows that across all crime benchmarks in all years, there is substantial evidence of anti-Black racial disparities in the killing of unarmed noncriminals by police. Here, we fail to recover the principle findings of Cesario et al. (2019). Racial disparities in the killing of unarmed citizens by police do not occur in proportion to the relative rates of noncriminality; unarmed and nonaggressing Black individuals are killed in greater numbers than would be expected given the relative populations of noncriminals.



**Figure 2.** Analysis of data on killings of armed suspects by police from “The Counted,” 2015 (blue) and 2016 (orange), using crime rate benchmarking. *Note.* Density curves show the natural log of the posterior distributions of the relative probability (for Black individuals relative to White individuals) of being killed by police. Values greater than 0 indicate anti-Black racial disparity, and values less than 0 indicate anti-White racial disparity. Central 90% credible intervals are represented by darker vertical bars on the distributions. Panel positions without curves indicate categories not considered in our analysis. Across crime rate benchmarks using National Crime Victimization Survey data, we find that there is no evidence of anti-Black racial disparity in police killings of armed individuals after differential crime rates have been accounted for. National Incident-Based Reporting System (NIBRS) crime rate data actually suggest anti-White racial disparities but see issues with the NIBRS data in Figure 3.

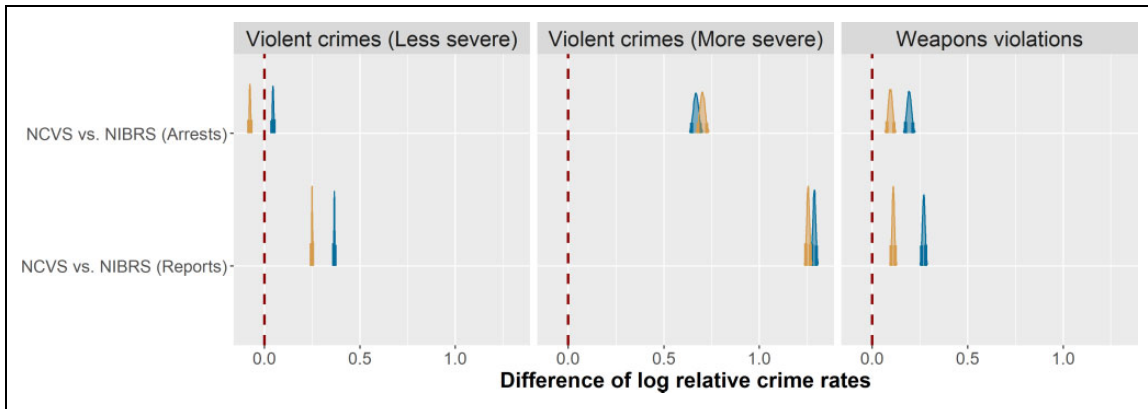
## Discussion

Writing almost a decade ago about policing, Goff and Kahn (2012) lament that it would be “shocking to think that there remained uncertainty about how to tell whether or not racial bias troubled one of our most important institutions” (pp. 177–178). They went on to address both the dearth of nationally representative data on police use-of-force and the lack of methodological paradigms for causal inference about the drivers of racial disparities in extant data. Progress is being made to address the data concerns (Garner et al., 2018; Goff et al., 2016). However, important issues concerning statistical methodology remain largely unaddressed.

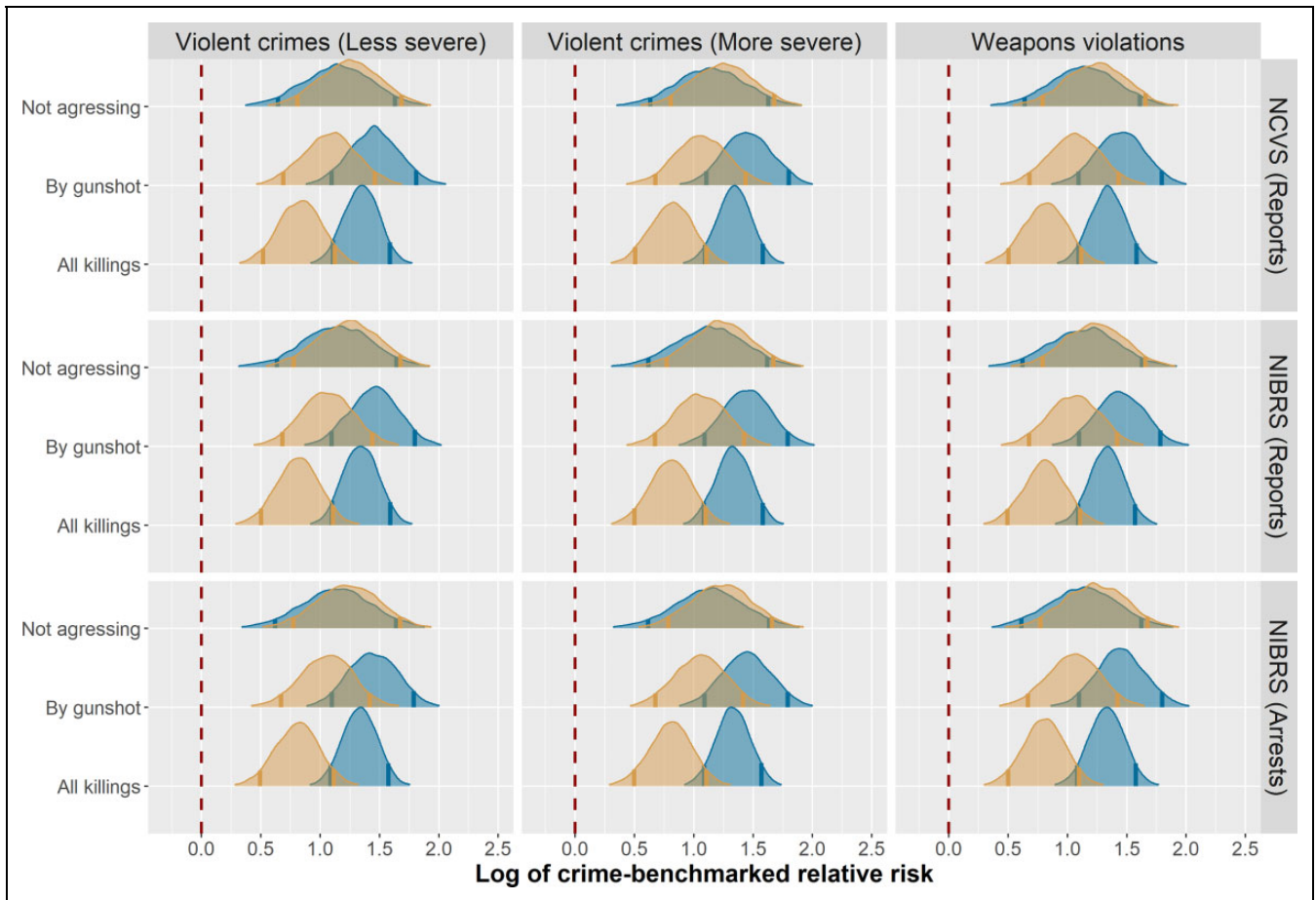
We have presented a theoretical model in which we have a direct causal understanding of the generative mechanism of police killings as a function of race and criminality—that is,

we explicitly define the probabilities for use-of-force against criminals and noncriminals as a function of race. We have shown that application of the statistical methodology advocated by Cesario et al. (2019) to data generated under this causal model can incorrectly suggest anti-White racial disparities in police use of lethal force, even when there is strong anti-Black bias hard-coded into the model. The statistical bias introduced by their methodology is conceptually, but not empirically, subtle. It can mislead not only about the magnitude but even the direction of effects, as we have shown both algebraically and through empirical analysis of the same data sets as used in their paper.

Although the principle empirical findings of Cesario et al. (2019) concerning racial disparities in the killing of unarmed citizens by police do not hold, we acknowledge that they do hold for the case of armed criminals killed by police. Our methods, however, can cover both cases reliably.



**Figure 3.** A contrast between the National Crime Victimization Survey (NCVS) and the National Incident-Based Reporting System (NIBRS) regarding crime rate differences in 2015 (blue) and 2016 (orange). *Note.* Density curves show the natural log of the posterior distributions of the ratio of relative crime rates, as NIBRS over NCVS. Values greater than 0 indicate higher Black crime rates in the NIBRS data than the NCVS data, and values less than 0 indicate the opposite. Central 90% credible intervals are represented by darker vertical bars on the distributions. We find evidence that across most years and crime classifications, NIBRS data show higher rates of crime being committed by Black individuals than is apparent from NCVS data.



**Figure 4.** Analysis of data on killings of unarmed suspects by police from “The Counted,” 2015 (blue) and 2016 (orange), using crime rate benchmarking. *Note.* Density curves show the natural log of the posterior distributions of the relative probability (for Black individuals relative to White individuals) of being killed by police. Values greater than 0 indicate anti-Black racial disparity, and values less than 0 indicate anti-White racial disparity. Central 90% credible intervals are represented by darker vertical bars on the distributions. Across all crime benchmarks in all years, we find substantial evidence of anti-Black racial disparities in killings of unarmed civilians by police.

## The Importance of Benchmarks

Empirical findings aside, Cesario et al. (2019) argue convincingly that to understand racial disparities in killings by police, we have to engage in proper benchmarking, such that we compare the relative rates of police use-of-force against criminals,  $\frac{\phi_B}{\phi_W}$ , and noncriminals,  $\frac{\theta_B}{\theta_W}$ , independently. The push to take benchmarking seriously is important for two main reasons: (1) It is important not to confound racial disparities that might arise from justifiable shootings of armed and dangerous criminals by police with racial disparities in the killing of innocent, unarmed civilians; and (2) encounter-conditional approaches that appear to mitigate the need for benchmarks are themselves typically confounded by differential encounter rates and contexts.

With respect to the first issue, researchers interested in the topic of racial disparities in police use-of-force have accounted for status as unarmed (e.g., Ross, 2015) prior to publishing relative risk estimates, meaning that the statistical bias factor that is present if one does not also benchmark on crime rates is only on the order of  $\frac{1-\alpha_B}{1-\alpha_W} \approx 1$ , but some presentations of data, for example, from *The Guardian* (2016), have not. The formal derivations provided here validate the benchmarking methods developed in Cesario et al. (2019) when applied to the relative counts of police killings of armed suspects. As in their initial analysis, we find reliable evidence that lethal force by police occurs in direct proportion to race-specific rates of violent crime. However, our formal derivations show that the benchmarking methods developed in Cesario et al. (2019) are misleading and statistically biased when applied to the relative counts of police killings of unarmed suspects. Likewise, their methods will lead to confounding when outcome data are not, or cannot be, classified into criminality categories. Our reanalysis of their data using an appropriate crime rate-correcting benchmark reveals strong and statistically reliable anti-Black racial disparity in police killings of unarmed civilians.

Second, it is important to take benchmarking seriously because, contra Johnson et al. (2019), population considerations cannot be sidestepped when estimating racial disparities in police use-of-force (see a concise proof in Knox & Mummolo, 2019). A similar proof is presented in Ross et al. (2018), who use a generative stochastic model to show that the overall racial disparity in police use-of-force can be decomposed into the product of two terms—racial disparity in use-of-force conditional on encounter and racial disparity in the frequency of encounters. So even if encounter-conditional approaches (e.g., Fryer, 2016; Johnson et al., 2019; Worrall et al., 2018) suggest no evidence of racial disparity in the use of lethal or less-than-lethal force by police conditional on encounter, the overall per capita morbidity and mortality from police use-of-force can be higher in the Black population if the Black population is subjected to higher encounter rates with police. Both recent and decade-old data show that Black individuals are more likely to be stopped by police than White individuals (Fryer, 2016; Gelman et al., 2007; Miller et al., 2017; U.S. Department of Justice, 2016), even after a variety of

statistical controls have been applied. Moreover, causal inference (e.g., Knox et al., 2019; Ross et al., 2018) approaches show that if police are biased in who they encounter, then encounter-conditional approaches can be severely confounded.

For example, consider a thought experiment in which police behavior is heterogeneous, with most officers following standard protocols and a small subset of officers engaging in additional unwarranted use of nonlethal force (like tasers) against Black individuals. Then, analysis of pooled encounter-conditional data would suggest that Black individuals are less likely to be shot rather than tased, as compared with White individuals. In other words, elevated levels of sublethal assault against innocent Black individuals by a subset of police would have the effect of diminishing the apparent severity of anti-Black racial disparities in lethal force conditional on encounter (Ross et al., 2018). Racial disparities in the frequency of taser use are consistent with such an explanation (Fryer, 2016). For this reason, anti-Black encounter bias is a confounding factor in recent encounter-conditional studies finding anti-White racial disparities (e.g., Fryer, 2016; Johnson et al., 2019; Worrall et al., 2018). For encounter-conditional analyses to be convincing, researchers would have to rule out the presence of racial bias in encounter probability (Knox & Mummolo, 2019).

Successful interventions to mitigate racial disparities in police use-of-force require that we reliably identify the drivers of such disparities. Methods leading to accurate causal inference are essential. An important step in validating our statistical tools involves applying them to simulated data sets, generated under a process that we understand explicitly—because we coded it. We should be asking ourselves: “If we knew the generative process of the data perfectly, would our statistical approach allow us to correctly detect real disparities?” In this article, we have used a simple generative model in just this way. We show that if the data generating process were such that unarmed, Black, noncriminals were more than twice as likely to be killed than unarmed, White, noncriminals (and if crime rates in our model were as we find empirically), then the statistical methodology of Cesario et al. (2019) would erroneously suggest anti-White racial disparities in the killing of unarmed noncriminals. This calls into question the validity of such methods.

As we move forward in studies of policing, new forms of data, like evidence from officer-worn cameras (Broussard et al., 2018; Willits & Makin, 2018), are becoming available. Such data raise hopes for accountability and detection of discriminatory violations of individual rights (Scheidlin, 2010) and even study of racial bias in respectfulness of language use and escalation or de-escalation of encounters (Voigt et al., 2017). However, there is also concern over how such data will be managed and protected from misuse (Ringrose, 2019). These new forms of data, however, may be able to help resolve conflicting reports about the existence of racial disparities in police behavior by recording possible disparities in both encounters and use-of-force conditional on encounter at the officer level. While such data can be powerful, they can only be appreciated



in light of statistical models that must themselves be validated on inferential targets.

## Acknowledgments

This article benefited from the feedback provided by Sheryl Gerety, the editor, and two anonymous reviewers.


## Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Cody T. Ross  <https://orcid.org/0000-0002-0067-4799>

Bruce Winterhalder  <https://orcid.org/0000-0001-6560-3302>

## References

- Binder, A., & Fridell, L. (1984). Lethal force as a police response. *Criminal Justice Abstracts*, 16, 250–280.
- Binder, A., & Scharf, P. (1980). The violent police-citizen encounter. *The Annals of the American Academy of Political and Social Science*, 452(1), 111–121.
- Broussard, M., Krishnamoorthy, B., Makin, D., & Willits, D. (2018). Extracting insights on use of force by police in encounters through topological data analysis of body-worn camera video datasets <http://www.math.wsu.edu/faculty/bkrishna/Papers/BWCVideosTDA.pdf>
- Bureau of Justice Statistics. (2016). *National Crime Victimization Survey*. <https://www.bjs.gov/index.cfm?ty=dcdetailiid=245>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32.
- Cesario, J., Johnson, D. J., & Terrill, W. (2019). Is there evidence of racial disparity in police use of deadly force? Analyses of officer-involved fatal shootings in 2015–2016. *Social Psychological and Personality Science*, 10(5), 586–595.
- Charbonneau, A., Spencer, K., & Glaser, J. (2017). Understanding racial disparities in police use of lethal force: Lessons from fatal police-on-police shootings. *Journal of Social Issues*, 73(4), 744–767.
- Federal Bureau of Investigation. (2016). *National incident-based reporting system*. <https://www.fbi.gov/services/cjis/ucr/nibrs>
- Feldman, J. M., Gruskin, S., Coull, B. A., & Krieger, N. (2017). Quantifying underreporting of law-enforcement-related deaths in United States vital statistics and news-media-based data sources: A capture–recapture analysis. *PLoS Medicine*, 14(10), e1002399.
- Fryer, R. G. (2016). An empirical analysis of racial differences in police use of force. <https://www.nber.org/papers/w22399.pdf>
- Fryer, R. G. (2017). Reconciling results on racial differences in police shootings. <https://pdfs.semanticscholar.org/90ec/2daf3f64298b7ff47190ed3a1d379005bd6e.pdf>
- Gabrielson, R., Jones, R. G., & Sagara, E. (2014). Deadly force, in Black and White: A ProPublica analysis of killings by police shows outside risk for young Black males. *ProPublica: Journalism in the Public Interest*. <https://www.propublica.org/article/deadly-force-in-black-and-white>
- Garner, J. H., Hickman, M. J., Malega, R. W., & Maxwell, C. D. (2018). Progress toward national estimates of police use of force. *PLoS One*, 13(2), e0192932.
- Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City police department’s “stop-and-frisk” policy in the context of claims of racial bias. *Journal of the American Statistical Association*, 102(479), 813–823.
- Goff, P. A., & Kahn, K. B. (2012). Racial bias in policing: Why we know less than we should. *Social Issues and Policy Review*, 6(1), 177–210.
- Goff, P. A., Lloyd, T., Geller, A., Raphael, S., & Glaser, J. (2016). *The science of justice: Race, arrests, and police use of force*. Center for Policing Equity. [https://policingequity.org/images/pdfs-doc/CPE\\_SoJ\\_Race-Arrests-UoF\\_2016-07-08-1130.pdf](https://policingequity.org/images/pdfs-doc/CPE_SoJ_Race-Arrests-UoF_2016-07-08-1130.pdf)
- The Guardian*. (2016). The Counted: People killed by police in the US. <http://www.theguardian.com/us-news/series/counted-us-police-killings>
- Hannon, L. (2020). An exploratory multilevel analysis of pedestrian frisks in Philadelphia. *Race and Justice*, 10(1), 87–113. <https://doi.org/10.1177/2153368717730106>
- Johnson, D. J., Tress, T., Burkel, N., Taylor, C., & Cesario, J. (2019). Officer characteristics and racial disparities in fatal officer-involved shootings. *Proceedings of the National Academy of Sciences*, 116(32), 15877–15882.
- Klinger, D., Rosenfeld, R., Isom, D., & Deckard, M. (2016). Race, crime, and the micro-ecology of deadly force. *Criminology & Public Policy*, 15(1), 193–222.
- Knox, D., Lowe, W., & Mummolo, J. (2019). The bias is built in: How administrative records mask racially biased policing. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3336338>
- Knox, D., & Mummolo, J. (2019). Making inferences about racial disparities in police violence. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3431132>
- McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
- Miller, T. R., Lawrence, B. A., Carlson, N. N., Hendrie, D., Randall, S., Rockett, I. R., & Spicer, R. S. (2017). Perils of police action: A cautionary tale from U.S. data sets. *Injury Prevention*, 23(1), 27–32.
- Nix, J., Campbell, B. A., Byers, E. H., & Alpert, G. P. (2017). A bird’s eye view of civilians killed by police in 2015: Further evidence of implicit bias. *Criminology & Public Policy*, 16(1), 309–340.
- Payne, B. K. (2006). Weapon bias: Split-second decisions and unintended stereotyping. *Current Directions in Psychological Science*, 15(6), 287–291.
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ringrose, K. (2019). Law enforcement’s pairing of facial recognition technology with body-worn cameras escalates privacy concerns. *Virginia Law Review*, 105, 57–66.

- Ross, C. T. (2015). A multi-level Bayesian analysis of racial bias in police shootings at the county-level in the United States, 2011–2014. *PLoS One*, *10*(11), e0141854.
- Ross, C. T., Winterhalder, B., & McElreath, R. (2018). Resolution of apparent paradoxes in the race-specific frequency of use-of-force by police. *Palgrave Communications*, *4*(1), 61.
- Scheindlin, S. A. (2010). Case of: *Floyd v. City of New York*, 739 F. Supp. 2d 376.
- Scott, K., Ma, D. S., Sadler, M. S., & Correll, J. (2017). A social scientific approach toward understanding racial disparities in police shooting: Data from the department of justice (1980–2000). *Journal of Social Issues*, *73*(4), 701–722.
- Selby, N., Singleton, B., & Flosi, E. (2016). *In context: Understanding police killings of unarmed civilians*. Contextual Press.
- Stan Development Team. (2018a). *RStan: The R interface to Stan* (R package Version 2.17.3) [Computer software]. <http://mc-stan.org/>
- Stan Development Team. (2018b). *The Stan core library* (Version 2.18.0). <http://mc-stan.org/>
- Takagi, P. (1981). *A Garrison State in “democratic” society*. Macmillan Education. ISBN 978-1-349-16588-9, pp. 205–219. [https://doi.org/10.1007/978-1-349-16588-9\\_9](https://doi.org/10.1007/978-1-349-16588-9_9)
- Tregle, B., Nix, J., & Alpert, G. P. (2019). Disparity does not mean bias: Making sense of observed racial disparities in fatal officer-involved shootings with multiple benchmarks. *Journal of Crime and Justice*, *42*(1), 18–31.
- U.S. Department of Justice. (2016). *Investigation of the Baltimore City Police Department*. <http://www.nytimes.com/interactive/2016/08/09/us/document-Baltimore-P-D-Findings-Report-FINAL.html>
- Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., Jurgens, D., Jurafsky, D., & Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences*, *114*(25), 6521–6526.
- White, M. D. (2006). Hitting the target (or not): Comparing characteristics of fatal, injurious, and noninjurious police shootings. *Police Quarterly*, *9*(3): 303–330.
- Willits, D. W., & Makin, D. A. (2018). Show me what happened: Analyzing use of force through analysis of body-worn camera footage. *Journal of Research in Crime and Delinquency*, *55*(1), 51–77.
- Worrall, J. L., Bishopp, S. A., Zinser, S. C., Wheeler, A. P., & Phillips, S. W. (2018). Exploring bias in police shooting decisions with real shoot/don't shoot cases. *Crime & Delinquency*, *64*(9), 1171–1192.

### Author Biographies

**Cody T. Ross** is a senior researcher at the Max Planck Institute for Evolutionary Anthropology in the Department of Human Behavior, Ecology, and Culture. His research focuses on the role of inequality on well-being and the emergence and maintenance of cultural institutions.

**Bruce Winterhalder** is Emeritus Professor of Anthropology and Graduate Group in Ecology at the University of California at Davis. His recent research focuses on the role of population ecology in social evolution and the economic organization of hunter-gatherer societies.

**Richard McElreath** is the director of the Department of Human Behavior, Ecology, and Culture at the Max Planck Institute for Evolutionary Anthropology. His recent research focuses on the application of formal modeling to central questions in the study of human behavior and social learning.

Handling Editor: Margo Monteith