

Reply to Knox and Mummolo: Critique of Johnson et al. (2019)

David J. Johnson

University of Maryland at College Park
Michigan State University

Joseph Cesario

Michigan State University

Recently we published a report showing that officer race did not predict the race of a civilian shot and, additionally, there was no evidence of anti-Black racial disparities among those fatally shot by police (Johnson, Tress, Burkel, Taylor, & Cesario, 2019). In response, Knox and Mummolo (2019) produced a critique of this work centering around two main issues: (1) The informational value of the analysis by Johnson et al. (2019); (2) The misleading nature of the quantity calculated by Johnson et al. (2019).

In what follows we address each of these points, arguing that point 1 is largely an issue of debate on which reasonable people will disagree, and showing that point 2 holds only for implausible states of the world, i.e., it is unlikely to apply in most cases given actual crime rates across different racial groups. Thus the original findings, as described in that manuscript, largely stand unchanged.

A Point of Clarification

Before addressing the critique, it is worth clarifying a common misunderstanding of our paper. In Johnson et al. (2019), we calculate the probability that a civilian is Black, Hispanic, or White *given* that a person has been fatally shot and some covariates. That is, the dataset contains only information about individuals fatally shot by police, and the race of the individual is predicted by a set of variables. Thus, we compute $\Pr(\text{race}|\text{shot}, X)$ where X is a set of covariates including officer race.

One criticism is that this probability is uninformative, and we should have calculated $\Pr(\text{shot}|\text{race}, X)$, that is, the probability that a person is shot *given* their race and some covariates. In some online commentary, this criticism has morphed into stating that we not only computed the wrong probability but that we did not understand the difference between the two—i.e., we had intended to compute $\Pr(\text{shot}|\text{race})$ but inadvertently computed $\Pr(\text{race}|\text{shot})$.

This misconception is incorrect. As we state in our intro-

duction, $\Pr(\text{shot}|\text{race}, X)$ is an important quantity. In fact, we have computed this quantity in our past research (Cesario, Johnson, & Terrill, 2019), which we discuss in more detail below. However, the goal of Johnson et al. (2019) was *not* to compute $\Pr(\text{shot}|\text{race}, X)$. It was instead to provide information about the nature of fatal police shootings from a different angle, by predicting the race of the person shot from a set of predictor variables. Indeed, this is exactly what we said we were going to do in the introduction to the paper.

Again, the approaches ask different (but related) questions. The key difference is that our approach $\Pr(\text{race}|\text{shot}, X)$ tests for racial disparities among those fatally shot. The approach $\Pr(\text{shot}|\text{race}, X)$ advocated by Knox and Mummolo (2019) tests for racial disparities among some population of individuals who encounter police. We acknowledge that the latter question can identify disparities not apparent when examining fatal officer-involved shootings alone. But the crux of the issue is that the latter quantity is difficult to estimate without a debate about *which* benchmark is the most appropriate. For this reason, we instead estimated a well-defined quantity that addresses one component of officer use of force, racial disparities among those fatally shot by police.

Part of the reason for this misconception is likely due to the language used in our significance statement. We wrote “White officers are not more likely to shoot minority civilians than non-White officers.” We should have written this sentence more carefully as it currently refers to estimating $\Pr(\text{shot}|\text{race}, X)$. What we should have written was a sentence about what we did estimate— $\Pr(\text{race}|\text{shot}, X)$: As the proportion of White officers in a FOIS increased, a person fatally shot was not more likely to be of a racial minority. This was our mistake, and we appreciate the feedback on this point. To be clear, this issue in no way invalidates our findings with regards to $\Pr(\text{race}|\text{shot}, X)$. It is, however, important to clearly state that our approach does not make statements about $\Pr(\text{shot}|\text{race}, X)$.

Critique 1: Informational Value of $\Pr(\text{race}|\text{shot}, X)$

The first critique of Johnson et al. (2019) centers around the idea that the *only* quantity of interest in understanding racial bias in fatal police shootings is $\Pr(\text{shot}|\text{race}, X)$, that is, the probability of being shot given a civilian’s race and some covariates. Thus, according to Knox and Mummolo (2019), $\Pr(\text{race}|\text{shot}, X)$ does not provide any useful information in

David Johnson is postdoctoral researcher in the Lab for Applied Social Science Research at the University of Maryland. Joseph Cesario is associate professor of psychology at Michigan State University. Both the original Johnson et al. (2019) manuscript and the Knox and Mummolo (2019) critique are available at the second author’s website.

understanding the topic of race and police shootings. As they state:

Anti-black bias in police shootings is defined as $\Pr(\text{shot}|\text{civilian black}, X) > \Pr(\text{shot}|\text{civilian white}, X)$ — all else equal, black civilians have a greater chance of being shot by police than white civilians.

Although $\Pr(\text{shot}|\text{race}, X)$ is an important quantity to estimate (and to be clear, *an approach we have taken in our past research*), it is not the only estimate that can tell us about race and fatal police shootings. Useful information on the topic of deadly force can be gained by also computing $\Pr(\text{race}|\text{shot}, X)$. This approach can answer questions such as: "Does the race of the police officer predict the race of the person fatally shot?" or "Among those fatally shot, is a person more likely to be Black than White when controlling for crime rates?" The answers to these questions tell us about cases where police officers used force that resulted in civilian fatalities. Do these questions tell us $\Pr(\text{shot}|\text{race})$? No, obviously not. But from our view more information is better. If we study racial disparities from a variety of different approaches and get consistent answers, this increases confidence in our findings.

We strongly suspect this is a point on which reasonable people will simply disagree. If one thinks that the question of racial disparity is answered *by and only by* $\Pr(\text{shot}|\text{race}, X)$, then no other quantity is worth knowing. Of course, if $\Pr(\text{race}|\text{shot}, X)$ is a misleading quantity, then computing it and trying to gain any information at all is in error. This is Knox and Mummolo (2019)'s second critique.

Critique 2: $\Pr(\text{race}|\text{shot})$ is a Misleading Quantity

The major critique by Knox and Mummolo (2019) centers around the issue that $\Pr(\text{race}|\text{shot})$ is a misleading quantity. They present two arguments in support of this position: One is the transformation of $\Pr(\text{race}|\text{shot})$ to $\Pr(\text{shot}|\text{race})$ via Bayes' rule, and the other is a thought experiment. We address each in turn.

Transformation of $\Pr(\text{race}|\text{shot})$ to $\Pr(\text{shot}|\text{race})$

Using Bayes' rule, Knox and Mummolo (2019) show that:

$$\Pr(\text{shot}|\text{civilian race}, X) = \Pr(\text{civilian race}|\text{shot}, X) \frac{\Pr(\text{shot}|X)}{\Pr(\text{civilian race}|X)} \quad (1)$$

They explain that $\Pr(\text{race}|\text{shot})$ is only equivalent to $\Pr(\text{shot}|\text{race})$ —the only quantity of interest for Knox and Mummolo (2019)—under the assumption that officers encounter equal numbers of black and white civilians in a given setting. We do not dispute this. However, we also did not make this claim, that $\Pr(\text{race}|\text{shot}) = \Pr(\text{shot}|\text{race})$. Rather,

this misunderstanding appears to be partially due to the issues we discussed in our opening clarification.

That being said, let us operate under the assumption that the quantity we should estimate is in fact $\Pr(\text{shot}|\text{race})$, as Knox and Mummolo (2019) argue. If $\Pr(\text{race}|\text{shot})$ leads to a different conclusion than $\Pr(\text{shot}|\text{race})$ we would classify it as a Type S (sign) error (Gelman & Carlin, 2014). It is illustrative to examine the real-world circumstances necessary to show $\Pr(\text{race}|\text{shot})$ yields an estimate in the opposite direction—a misleading quantity. In other words, what are the real-world circumstances required to 1) show a *lack* of anti-Black disparity in the overall number of individuals fatally shot by police while 2) showing an anti-Black bias in the probability of being fatally shot by police?

One way to answer this question is to examine how much estimates of police exposure to situations where fatal shootings typically occur— $\Pr(W)$ and $\Pr(B)$ —need to deviate from equality to create significant anti-Black bias, given our estimates of $\Pr(\text{race}|\text{shot})$. We can use known benchmarks of police exposure to examine whether this degree of disparity is plausible.

Looking at the raw numbers in our dataset (ignoring covariates for simplicity), 27% of people fatally shot (245/917) were Black, compared to 55% who were White (501/917). Thus, a person fatally shot was half as likely to be Black than White (or, equivalently, a person fatally shot was 2.0 times more likely to be White than Black). That is, $\Pr(B|S) / \Pr(W|S) = 0.5$. To convert that to the likelihood that a person shot is Black vs. White we apply Bayes' rule:

$$\frac{\Pr(S|B)}{\Pr(S|W)} = \frac{\Pr(B|S)}{\Pr(W|S)} \cdot \frac{\Pr(W)}{\Pr(B)} \quad (2)$$

Where $\Pr(W) / \Pr(B)$ is a constant, such that a value of 1 indicates that Whites have equal exposure compared to Blacks to police encounters where fatal force is likely to be used.

Given the values from our dataset, to see evidence of anti-Black bias, White individuals would have to be more than twice as likely to encounter police in situations where fatal force is likely to be used, $[\Pr(B|S) / \Pr(W|S)] \cdot [\Pr(W) / \Pr(B)] = 0.5 \cdot 2.0 = 1.0$. An odds ratio of 2.0 (i.e., a Black person is twice as likely to be fatally shot than a White person) would require White individuals to be four times as likely to encounter police in situations where fatal force is likely to be used.

The issue is that at a number of reasonable benchmarks for police exposure suggest Black individuals may actually be *more* likely to encounter police than White individuals in situations where fatal force is typically used. Research on real-world policing behavior indicates fatal shootings are strongly tied to situations where violent crime is committed (Klinger, Rosenfeld, Isom, & Deckard, 2016). This is reflected in the fact that between 85% - 90% of fatal police shootings involve

armed civilians (Johnson et al., 2019; Nix, Campbell, Byers, & Alpert, 2017). While we do not know the exact rate at which police encounter Black and White individuals in the context of violent crime, we can estimate these rates from several different proxies.

For example, consider 2015 Center for Disease Control counts of homicide by firearm (3070 White / 7515 Black = 0.4). Plugging this value into Equation 2 yields a value of $0.5 \cdot 0.4 = 0.2$, which indicates that a Black person would be five times less likely to be fatally shot than a White person, given exposure rates as defined by homicide counts. We could also reference a more conservative benchmark: victim reports of the race of offenders who committed violent crime against them as reported in the 2015 National Crime Victimization Survey (261 White, 156 Black, = 1.7). Plugging this value into Equation 2 yields a value of $0.5 \cdot 1.7 = 0.85$, which indicates that a Black person would be 1.2 times less likely to be fatally shot than a White person. Note that neither of these benchmarks are generated from police data, which sidesteps the issue of circularity in whether the benchmarks themselves are subject to bias from police.

Although in both cases our estimate of $\Pr(S|B) / \Pr(S|W)$ differs in magnitude from the estimate of $\Pr(B|S) / \Pr(W|S)$, neither discrepancy results in a Type S error where $\Pr(S|B) / \Pr(S|W)$ identifies anti-Black disparity and $\Pr(B|S) / \Pr(W|S)$ anti-White disparity, or vice versa. However, this example analysis demonstrates that claims about anti-Black disparity in the likelihood of being shot depend strongly on the benchmark for exposure used and it is difficult to identify the most appropriate benchmark at the national level.

We have tackled this issue in the past. Rather than try to identify one single benchmark for exposure to police in violent crime situations, we came up with 14 different proxies for exposure, some of which were generated from police data and some independently (Cesario et al., 2019). We used these proxies to analyze racial disparities (i.e., $\Pr(S|B) / \Pr(S|W)$) in three types of fatal shootings across two years (all shootings, unarmed and nonaggressing shootings, and shootings where a civilian was reaching for an object). Using a multi-verse analysis, we found only one significant anti-Black disparity in 144 possible tests.

Thus, when using benchmarks for police exposure, we again find no significant evidence of anti-Black disparity in the likelihood of being fatally shot by police. These results are consistent with the findings that a person fatally shot by police is not more likely to be Black than White, even when controlling for relevant covariates.

Thought Experiment

As additional evidence for the misleading nature of $\Pr(\text{race}|\text{shot})$, Knox and Mummolo (2019) present a thought experiment:

Imagine police encounter 100 civilians—10 black and 90 white—in identical circumstances. Due to anti-black bias, they shoot five black civilians (50%), and nine white civilians (10%). The approach in Johnson et al. would show a much higher chance the victim is white, conditional on being shot ($9/14 = .64$), than black ($5/14 = .36$), and erroneously conclude no anti-black bias.

It is true that *if* police encountered and killed Black and White civilians at the proportions described in this thought experiment, then $\Pr(\text{race}|\text{shot})$ would produce a misleading quantity. The question is not whether some arbitrary world can be created in which $\Pr(\text{race}|\text{shot})$ is misleading; the question is whether *given the actual data* $\Pr(\text{race}|\text{shot})$ is misleading. The relevance of the thought experiment is dependent on the contact patterns listed being a reasonable approximation of the real world, namely, that police encounter Black and White civilians on the order of 10, 90 and shoot on the order of 5, 9, respectively.

Thus we can ask, do the encounter rates provided by the Knox and Mummolo (2019) thought experiment resemble actual encounter rates? The problem is that the thought experiment does not resemble actual police-civilian contacts as they pertain to situations in which deadly force decisions are relevant. As discussed previously, using firearm homicide cases from 2015, Black civilians were exposed to police in situations where fatal force is relevant 1.7 times more than Whites, rather than Whites being exposed to police 9.0 times more than Blacks as in Knox and Mummolo (2019)'s thought experiment. Even when using the more conservative benchmark of victim reports of the race of offenders who committed violent crime against them from 2015, White civilians were only exposed to police 1.7 times more than Black civilians, still less than the 2.0 times greater exposure needed to show evidence of anti-Black racial disparity in the likelihood of being shot.

Given that 261 Blacks and 526 Whites were shot by the police in our 2015 database, our estimate of $\Pr(\text{race}|\text{shot})$ reveals no anti-black disparity and is in the same direction as the estimate of $\Pr(\text{shot}|\text{race})$. The values provided by the thought experiment are simply not close to actual patterns of police contact. While the thought experiment reveals there are possible situations in which our calculation would mask anti-Black disparity, we question whether those situations are common.

If, in fact, the point of the thought experiment is simply to show that $\Pr(\text{race}|\text{shot})$ is only equivalent to $\Pr(\text{shot}|\text{race})$ when exposure rates are equal, again, we do not dispute this and did not argue this point in our paper. However, the patterns of exposure that lead to a Type S error as described in the thought experiment simply do not match real world benchmarks.

A Point of Agreement

One point where we agree with Knox and Mummolo (2019) is that current analyses on police use of force are limited by the lack of comprehensive and complete national databases on police interactions with the public, both those where force is used *and* those where force is not used. In 2019, the Federal Bureau of Investigation launched the National Use-of-Force Data Collection, which aims to provide information about civilians, officers and circumstances surrounding shootings and other types of force. Although a step in the right direction, participation is voluntary, meaning departments may opt out of the program. Second, this database will still not provide information about how often police encounter civilians of different races, which is needed to calculate exposure rates.

If such data were made available, this would increase the ease at which other approaches to studying use of force could be applied. Given that $\Pr(\text{shot}|\text{race})$ is dependent on estimates of police exposure, better information about exposure rates would provide more definitive tests of this question.

References

- 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 Personality Science*, *10*, 586–595.
- Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651.
- 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., & Mummolo, J. (2019). Making inferences about racial disparities in police violence.
- 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.