INSTITUTE for REPLICATION

No. 129
I4R DISCUSSION PAPER SERIES

Contested Killings Replication: A Comment on Morris and Shoub (2023)

Jacques Courbe
Joshua Ferrer
Graham Straus

May 2024

I4R Discussion Paper Series

I4R DP No. 129

Contested Killings Replication: A Comment on Morris and Shoub (2023)

Jacques Courbe[1], **Joshua Ferrer**[1], **Graham Straus**[1]

[1]*University of California, Los Angeles/USA*

May 2024

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

Contested Killings Replication: A comment on Morris and Shoub (2023)

Jacques Courbe,* *UCLA*
Joshua Ferrer,† *UCLA*
Graham Straus,‡ *UCLA*

February 28th, 2024

Abstract

Morris and Shoub (2024) study whether fatal police shootings mobilize voter participation in presidential elections. They use a discontinuity-in-time design to causally estimate the effect of a police killing on turnout, comparing the voter participation of communities near a killing before and after election day. Morris and Shoub (2024) find that police killings spurred increased turnout, especially in Black communities, where the killing trended on Google, where the community was plurality Black, and where the victim's race was Black. They find that the local average treatment effect on participation within a quarter-mile radius of a police killing is upwards of 7 percentage points and statistically significant at the 95% level of confidence. We encounter difficulties when attempting to reproduce the analysis, but are able to replicate the main results using similar data. In fact, we find the effect of a proximate police killing on participation to be upwards of 8 percentage points.

KEYWORDS: Voter turnout, criminal justice, race and ethnicity politics, police violence, discontinuity-in-time

JEL CODES: P00.

*Ph.D. Student, Department of Political Science. E-mail: jcourbe@g.ucla.edu
†Corresponding Author. Ph.D. Candidate, Department of Political Science. Email: joshuaferrer@ucla.edu
‡Ph.D. Student, Department of Political Science.gpstraus@gmail.com

1 Introduction

Morris and Shoub (2024) combine granular demographics data from the U.S. Census Bureau's 5-year American Community Surveys, census-block-level turnout data from L2, and data on geolocated police shootings from the *Washington Post's* Fatal Force Database and the Mapping Police Violence. They employ a discontinuity-in-time design, comparing the effects of police shootings on turnout that occur directly before and directly after a presidential election. The authors reason that this produces a causally valid analysis, since police shootings should occur randomly in a short period of time and only police shootings that occur before an election can affect that election's voter turnout. The running variable employed is time, and they examine the effect size of police shootings as distance increases between a fatal police shooting and a census block. The "treatment" is a fatal police shooting occurring near a neighborhood immediately preceding a presidential election, and the control is a fatal police shooting occurring immediately after an election. The outcome of interest is the difference in voter turnout in census block groups between the treatment and control conditions.

The paper tested the effects of a fatal police shooting on voter turnout for populations where fatal police shootings tend to occur. The main claims are that: the discontinuity-in-time design produces causally valid inferential results, the effect of a fatal police shooting on voter turnout is upwards of 7 percentage point boost in the immediate vicinity of where a shooting has occurred (less than 0.1 miles) with a 95% confidence interval of +/- 4 percentage points, and that the effect is concentrated in places where the killing trended on the internet, where the census block is plurality Black, and where the victim was Black. The paper also studied the effects of police killings on support for a police abolition referendum in Minneapolis, MN.

Our replication focuses on the effects of a fatal police killing on voter turnout. We first attempt to reproduce the findings using the data and code repository archived by the authors in the American Political Science Review Dataverse (hosted by the Harvard Dataverse).[1] We were grateful that authors deposited their code and data to ease the replication process. However, we found these files incomplete, some data missing, many references to directories incorrect and/or missing, and some packages used unavailable. This hindered our ability to easily reproduce Morris and Shoub (2024)'s findings. Despite these obstacles, we were able to succesfully reproduce the main results using their code and data provided, with some necessary corrections and data additions.

We also sought to conceptually replicate the main findings by conducting our own analysis of the L2 voter file and creating original code. We did not have the exact same L2 data snapshots as the authors did, but we did manage to successfully produce the main findings. In fact, we observe slightly larger effects than those presented in Morris and Shoub (2024), with the effect size of an officer-involved fatality within a quarter-mile radius on increased voter turnout to be upwards of 8 percentage points.

We discuss our reproduction and replication of Morris and Shoub (2024) in more detail below, including the issues we ran into and a comparison of the results they arrived at, what we calculated in the reproduction, and what we calculated in the replication.

2 Reproducibility

2.1 L2 data

We have access to the L2 voter files for each state from 2014 - 2020, similar to the authors. It is worth noting that this is proprietary, privately-owned data that is expensive and only possessed by a handful of academic institutions. This makes it impossible for most scholars to reproduce or replicate the authors' study. The UCLA L2 voterfile is stacked so as to contain all demographics and vote history

[1] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DMGOXD

data together. We similarly aggregate by state, election date, and fips code down to the census block group level of geography. We count registrants (people in the file) and voters at this level of geography, just as the authors do. We have nearly exactly the same voting data. Differences are due to slightly different file dates in some cases. The UCLA L2 voterfile contains snapshots at least 90 days after a given election. There is a trade-off in using a voterfile snapshot too close to the election because counties are still processing and recording votes and reporting to L2. Too far from the election means that voters in a state may be purged from the file and therefore miscalculated. For example, in the Colorado 2018 election case the authors use the L2 file form August 2019. We use a file from May 2019. The authors source Indiana 2020 election results from a January 2021 file. We use a file from July 2021. These differences end up being negligible as we reproduce their results almost exactly, but it is worth noting as voterfile best practices become more well established. The authors note the file snapshots they use in the appendix following best practices (Kim and Fraga 2022).

2.2 Reproducing the Main Findings

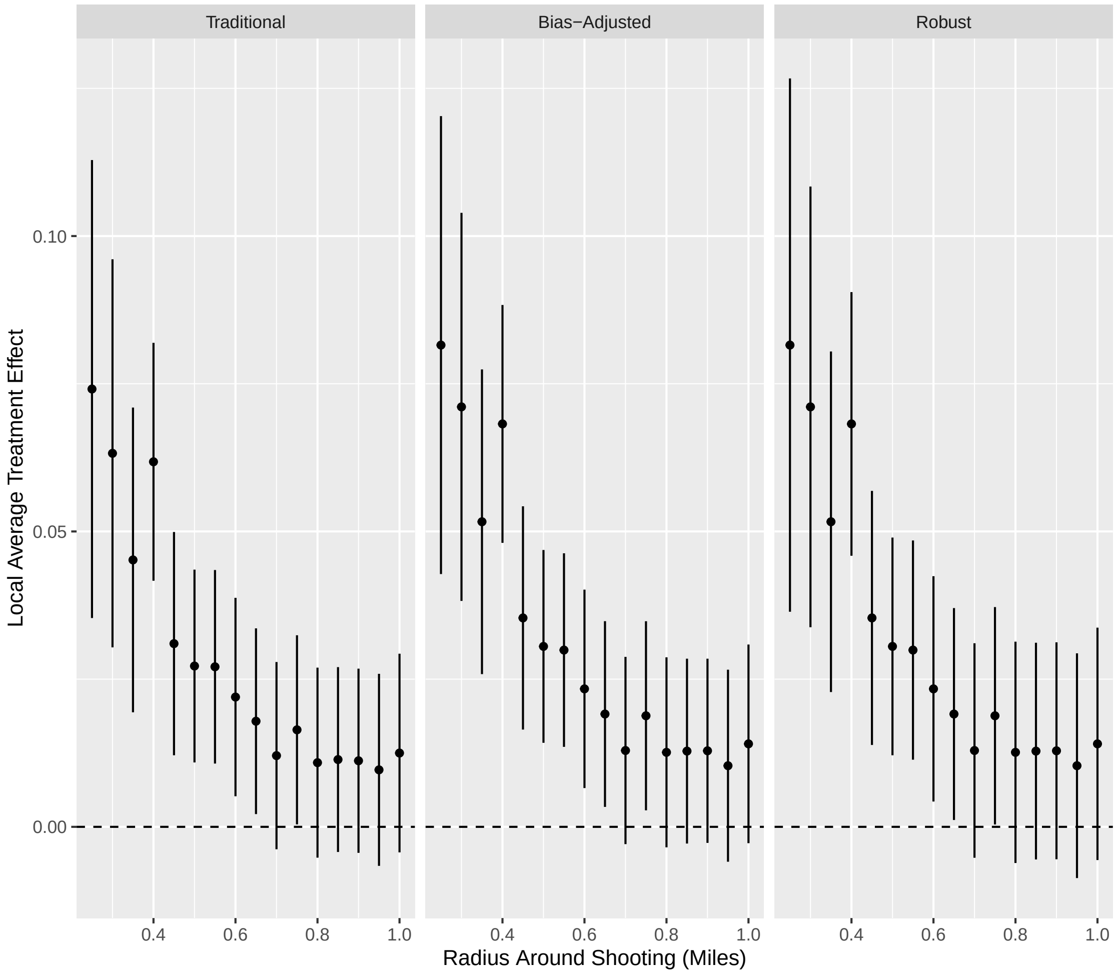
In terms of reproducibility, we were able to exactly reproduce Figure 3, which contains the main results of the paper. However, in order to reproduce the results, we made modifications to the replication code published on the APSR's Dataverse repository. The code included numerous incorrect references to directories in the replication archive, which we fixed manually in order to load the datasets necessary to perform the analysis. The codebase also made use of three R packages (rgeos, maptools, and rgdal) that CRAN no longer hosts, because they were removed at the request of the maintainer. Given that "code rot" is an issue known to affect the long-term computational reproducibility of social science papers (Peer et al. 2021), and the existence of dependency management systems like Docker and the packrat package in R, it seems reasonable to expect that top journals would take greater care to ensure the reproducibility of articles they elect to publish. Our attempt to computationally reproduce this paper's results took place less than a year after its initial publication.

The data repository files included the authors' code to draw large amounts of census data that the authors used in their analysis. This code worked as intended, but the exclusion of the raw data was not specified in the repository and running the provided tidycensus code significantly increased the computational time required to reproduce the author's main findings. The authors make efforts to state what data is and is not included in the dataverse, but we encourage them to be clearer with this documentation state the reason why any data used in the analysis is omitted from the dataverse.

3 Replication

3.1 Replicating Main Results

Figure 1: Morris and Shoub (2024) Figure 3 Replication



Using the provided tidycensus code and CVAP raw data, we went through the same process as detailed by the authors to replicate Figure 3, the main regression discontinuity in time results. We used 180 days of shootings on either side of the election cutoff in 2020 and 2016 generals sourcing the previous period from 2014 and 2018 midterms just as the authors do. We calculated entropy balancing weights in the same manner the authors do.

Despite utilizing L2 data with different dates, independently downloading census data, and creating our own code for the analysis, we successfully produced results that almost exactly mirror Morris and Shoub (2024).

We compare the point estimates of our replication, presented in Table 1 with those presented by Morris and Shoub (2024) in Table B1 of SI-B (Table 2 below). The results largely align. To take one example, Morris and Shoub (2024) report the following results of an RD-Robust estimator at a 0.25-mile threshold: 308 effective sample size, .45-mile bandwidth, 0.069 LATE, 0.011 p-value, a confidence interval of [0.016, 0.122], and a standard error of 0.027. In comparison, for this same model we report a 281 effective sample size, .41-mile bandwidth, 0.082 LATE, 0.000 p-value, a confidence interval of [0.026, 0.127], and a standard error of 0.23. For every model run, we return LATEs that are slighlty higher than those reported in Morris and Shoub (2024) and standard errors and p-values that are slightly smaller.

4 Conclusion

Overall, our impression is that the authors made commendable efforts to ensure computational reproducibility of their results. In a one-day replication session, we were able to successfully arrive at similar conclusions to their main findings. However, we encountered a series of difficulties in this exercise: outdated packages, missing data and file directories, and L2 data that can not be exactly replicated.

One additional extension worth pursuing is using the L2 data and CVAP estimates produced by the ACS to study whether these effects are specific to Black voters. Morris and Shoub (2024) conduct a heterogeneity test showing that the positive effects of officer-involved fatal shootings on voter turnout concentrate in plurality Black areas. However, this analysis relies on an ecological inference–that the effect is driven by Black voters in these census blocks. We propose an extension using the L2 voter file data to study whether the effect is indeed driven by Black voters, or by voters who happen to live in plurality-Black neighborhoods.

References

Kim, S.-y. S. and Fraga, B.: 2022, When Do Voter Files Accurately Measure Turnout? How Transitory Voter File Snapshots Impact Research and Representation.

Morris, K. T. and Shoub, K.: 2024, Contested killings: The mobilizing effects of community contact with police violence, *American Political Science Review* **118**(1), 458–474.

Peer, L., Orr, L. V. and Coppock, A.: 2021, Active maintenance: A proposal for the long-term computational reproducibility of scientific results, *PS: Political Science* **54**(3), 462–466.

5 Tables

Table 1: Replication of Morris and Shoub 2023 Table B.1

Miles	n	bw	Type	LATE	pv	ci	se
0.25	281	41	Traditional	0.074	0.000	[0.035, 0.113]	0.020
0.25	281	41	Bias-Adjusted	0.082	0.000	[0.043, 0.12]	0.020
0.25	281	41	Robust	0.082	0.000	[0.036, 0.127]	0.023
0.30	453	44	Traditional	0.063	0.000	[0.03, 0.096]	0.017
0.30	453	44	Bias-Adjusted	0.071	0.000	[0.038, 0.104]	0.017
0.30	453	44	Robust	0.071	0.000	[0.034, 0.108]	0.019
0.35	654	51	Traditional	0.045	0.001	[0.019, 0.071]	0.013
0.35	654	51	Bias-Adjusted	0.052	0.000	[0.026, 0.077]	0.013
0.35	654	51	Robust	0.052	0.000	[0.023, 0.08]	0.015
0.40	657	37	Traditional	0.062	0.000	[0.042, 0.082]	0.010
0.40	657	37	Bias-Adjusted	0.068	0.000	[0.048, 0.088]	0.010
0.40	657	37	Robust	0.068	0.000	[0.046, 0.091]	0.011
0.45	1027	45	Traditional	0.031	0.001	[0.012, 0.05]	0.010
0.45	1027	45	Bias-Adjusted	0.035	0.000	[0.016, 0.054]	0.010
0.45	1027	45	Robust	0.035	0.001	[0.014, 0.057]	0.011
0.50	1756	64	Traditional	0.027	0.001	[0.011, 0.044]	0.008
0.50	1756	64	Bias-Adjusted	0.031	0.000	[0.014, 0.047]	0.008
0.50	1756	64	Robust	0.031	0.001	[0.012, 0.049]	0.009
0.55	2162	66	Traditional	0.027	0.001	[0.011, 0.043]	0.008
0.55	2162	66	Bias-Adjusted	0.030	0.000	[0.014, 0.046]	0.008
0.55	2162	66	Robust	0.030	0.002	[0.011, 0.048]	0.009
0.60	2192	58	Traditional	0.022	0.010	[0.005, 0.039]	0.009
0.60	2192	58	Bias-Adjusted	0.023	0.006	[0.007, 0.04]	0.009
0.60	2192	58	Robust	0.023	0.016	[0.004, 0.042]	0.010

Table 1: Replication of Morris and Shoub 2023 Table B.1 *(continued)*

Miles	n	bw	Type	LATE	pv	ci	se
0.65	2263	54	Traditional	0.018	0.026	[0.002, 0.034]	0.008
0.65	2263	54	Bias-Adjusted	0.019	0.017	[0.003, 0.035]	0.008
0.65	2263	54	Robust	0.019	0.037	[0.001, 0.037]	0.009
0.70	2808	57	Traditional	0.012	0.136	[-0.004, 0.028]	0.008
0.70	2808	57	Bias-Adjusted	0.013	0.110	[-0.003, 0.029]	0.008
0.70	2808	57	Robust	0.013	0.163	[-0.005, 0.031]	0.009
0.75	3669	65	Traditional	0.016	0.044	[0, 0.032]	0.008
0.75	3669	65	Bias-Adjusted	0.019	0.021	[0.003, 0.035]	0.008
0.75	3669	65	Robust	0.019	0.045	[0, 0.037]	0.009
0.80	4326	67	Traditional	0.011	0.185	[-0.005, 0.027]	0.008
0.80	4326	67	Bias-Adjusted	0.013	0.124	[-0.003, 0.029]	0.008
0.80	4326	67	Robust	0.013	0.187	[-0.006, 0.031]	0.010
0.85	5037	70	Traditional	0.011	0.154	[-0.004, 0.027]	0.008
0.85	5037	70	Bias-Adjusted	0.013	0.108	[-0.003, 0.028]	0.008
0.85	5037	70	Robust	0.013	0.171	[-0.006, 0.031]	0.009
0.90	5442	67	Traditional	0.011	0.160	[-0.004, 0.027]	0.008
0.90	5442	67	Bias-Adjusted	0.013	0.106	[-0.003, 0.028]	0.008
0.90	5442	67	Robust	0.013	0.169	[-0.005, 0.031]	0.009
0.95	5212	59	Traditional	0.010	0.245	[-0.007, 0.026]	0.008
0.95	5212	59	Bias-Adjusted	0.010	0.212	[-0.006, 0.027]	0.008
0.95	5212	59	Robust	0.010	0.286	[-0.009, 0.029]	0.010
1.00	5701	57	Traditional	0.012	0.145	[-0.004, 0.029]	0.009
1.00	5701	57	Bias-Adjusted	0.014	0.102	[-0.003, 0.031]	0.009
1.00	5701	57	Robust	0.014	0.162	[-0.006, 0.034]	0.010

Table 2: Morris and Shoub 2023 Table B.1

Miles	n	bw	Type	LATE	pv	ci	se
0.25	308	45	Traditional	0.064	0.006	[0.018, 0.109]	0.023
0.25	308	45	Bias-Adjusted	0.069	0.003	[0.024, 0.114]	0.023
0.25	308	45	Robust	0.069	0.011	[0.016, 0.122]	0.027
0.30	528	52	Traditional	0.052	0.008	[0.014, 0.091]	0.020
0.30	528	52	Bias-Adjusted	0.058	0.003	[0.019, 0.097]	0.020
0.30	528	52	Robust	0.058	0.012	[0.013, 0.103]	0.023
0.35	798	57	Traditional	0.035	0.027	[0.004, 0.066]	0.016
0.35	798	57	Bias-Adjusted	0.041	0.011	[0.009, 0.072]	0.016
0.35	798	57	Robust	0.041	0.027	[0.005, 0.077]	0.018
0.40	702	37	Traditional	0.048	0.009	[0.012, 0.085]	0.018
0.40	702	37	Bias-Adjusted	0.055	0.003	[0.019, 0.091]	0.018
0.40	702	37	Robust	0.055	0.009	[0.013, 0.096]	0.021
0.45	1,069	46	Traditional	0.024	0.073	[-0.002, 0.05]	0.013
0.45	1,069	46	Bias-Adjusted	0.027	0.041	[0.001, 0.053]	0.013
0.45	1,069	46	Robust	0.027	0.084	[-0.004, 0.058]	0.016
0.50	1,363	50	Traditional	0.020	0.051	[0, 0.039]	0.010
0.50	1,363	50	Bias-Adjusted	0.022	0.028	[0.002, 0.042]	0.010
0.50	1,363	50	Robust	0.022	0.060	[-0.001, 0.045]	0.012
0.55	2,298	69	Traditional	0.017	0.057	[-0.001, 0.034]	0.009
0.55	2,298	69	Bias Adjusted	0.019	0.034	[0.001, 0.036]	0.009
0.55	2,298	69	Robust	0.019	0.073	[-0.002, 0.039]	0.010
0.60	2,556	67	Traditional	0.016	0.073	[-0.001, 0.033]	0.009
0.60	2,556	67	Bias-Adjusted	0.018	0.039	[0.001, 0.036]	0.009
0.60	2,556	67	Robust	0.018	0.078	[-0.002, 0.039]	0.010
0.65	2,608	59	Traditional	0.010	0.259	[-0.008, 0.029]	0.009
0.65	2,608	59	Bias-Adjusted	0.010	0.262	[-0.008, 0.029]	0.009

Table 2: Morris and Shoub 2023 Table B.1 *(continued)*

Miles	n	bw	Type	LATE	pv	ci	se
0.65	2,608	59	Robust	0.010	0.346	[-0.011, 0.032]	0.011
0.70	3,489	69	Traditional	0.008	0.349	[-0.009, 0.025]	0.009
0.70	3,489	69	Bias-Adjusted	0.009	0.284	[-0.008, 0.026]	0.009
0.70	3,489	69	Robust	0.009	0.366	[-0.011, 0.029]	0.010
0.75	3,926	69	Traditional	0.010	0.274	[-0.008, 0.027]	0.009
0.75	3,926	69	Bias-Adjusted	0.012	0.191	[-0.006, 0.029]	0.009
0.75	3,926	69	Robust	0.012	0.271	[-0.009, 0.033]	0.011
0.80	4,515	71	Traditional	0.003	0.722	[-0.016, 0.023]	0.010
0.80	4,515	71	Bias-Adjusted	0.004	0.666	[-0.015, 0.024]	0.010
0.80	4,515	71	Robust	0.004	0.722	[-0.019, 0.028]	0.012
0.85	5,031	71	Traditional	0.005	0.584	[-0.014, 0.024]	0.010
0.85	5,031	71	Bias-Adjusted	0.006	0.549	[-0.013, 0.025]	0.010
0.85	5,031	71	Robust	0.006	0.622	[-0.017, 0.029]	0.012
0.90	5,604	71	Traditional	0.005	0.650	[-0.016, 0.026]	0.011
0.90	5,604	71	Bias-Adjusted	0.005	0.644	[-0.016, 0.026]	0.011
0.90	5,604	71	Robust	0.005	0.707	[-0.021, 0.031]	0.013
0.95	6,102	70	Traditional	0.003	0.790	[-0.018, 0.023]	0.010
0.95	6,102	70	Bias-Adjusted	0.003	0.772	[-0.017, 0.024]	0.010
0.95	6,102	70	Robust	0.003	0.813	[-0.022, 0.028]	0.013
1.00	6,799	71	Traditional	0.005	0.630	[-0.016, 0.026]	0.011
1.00	6,799	71	Bias-Adjusted	0.006	0.590	[-0.015, 0.027]	0.011
1.00	6,799	71	Robust	0.006	0.661	[-0.02, 0.031]	0.013