The Impact of Affirmative Action Litigation on Police Killings of Civilians*

Robynn Cox Jamein P. Cunningham Alberto Ortega

March 11, 2024

Abstract

Although research has shown that court-ordered hiring quotas increase the number of minority police officers in litigated cities, there has been little insight into how workforce diversity, or lack thereof, may impact police violence against civilians. Using an event-study framework, we find that the threat of affirmative action litigation reduces police killings of non-White civilians in the long-run. In addition, we find evidence of lower arrest rates for non-White civilians and more diverse police departments 25 years after litigation. Our results highlight the vital role that federal interventions have in addressing police behavior and the use of lethal force.

Keywords: affirmative action, excessive use of force, police shootings, race
JEL *Classification*:: I28, J15, J78, K42

*__Cox__: School of Public Policy at the University of California at Riverside. robynn.cox@ucr.edu. **Cunningham**: The Law School & LBJ School of Public Affairs at the University of Texas at Austin. jamein.cunningham@law.utexas.edu. **Ortega**: O'Neill School of Public and Environmental Affairs, Indiana University. alorte@iu.edu. We are grateful for funding from the Washington Center for Equitable Growth and the PSI Center for Workplace Diversity and Inclusion at the University of Memphis. We would also like to thank participants in the NEA Sessions at the ASSA Annual Meetings, Duke Empirical Criminal Law Roundtable, APPAM Fall Research Conference, SEA Fall Meetings, ALEA Annual Meetings, The Federal Reserve Bank of Minneapolis OIGI Day-Ahead Fall Research Conference, Transatlantic Workshop on the Economics of Crime, and the NBER SI Crime Workshop; seminar participants at Harvard University, Indiana University, BCTR at Cornell University, the University of Pennsylvania, RAND Corp, the Virtual Law and Economics Workshop, the University of Memphis, Yale's Economic History Workshop, University of Rutgers–Camden, Queen's University, Duke University, Emory University, University of Pittsburgh, Brookings Institute, Syracuse University, University of Chicago (Harris), University of Texas (LBJ) Dartmouth College, Stanford University, NYU, Wesleyan University, and our colleagues Andrew Goodman-Bacon, Nic Duquette, Jose Joaquin Lopez, Anthony Yezer, David Abrams, John Pfaff, Mallika Thomas, Patrick Bayer, and Patrick Mason for comments and advice. Finally, we thank Justin McCrary for making the litigation data available and Mohsen Naghavi and the GBD Collaborators for making their data programs available for measurement error in police killings. Any errors or omissions are ours alone.

1 Introduction

Questions regarding the efficacy of policing, particularly as it pertains to the disproportionate use of excessive force towards minorities, are of critical concern in the United States. Police violence is a leading cause of death for young men of color following accidents, suicide, other homicides, heart disease, and cancer (Edwards et al., 2019). Black men and women are 2.5 and 1.4 times more likely, respectively, to die because of police violence than their White counterparts (Edwards et al., 2018). After more than fifty years of protests in response to police violence in Black communities, the conversation has shifted from police reform to abolition. While most Americans agree that improvements are necessary to address gross inequities in policing (Shannon, 2020), there is considerable disagreement about how to transform the police. Abolitionists recommend defunding the police and redirecting those funds to social programs devoted to easing the burdens of poverty and disenfranchisement. Reformists argue for greater incorporation of social programs and initiatives within the current structure of policing, which would likely increase funding for public safety (Bell, 2016).[1]

Historically, police reform has centered around three themes: 1) diversity training, 2) transparency and oversight, and 3) more diversity among police officers.[2] As early as 1968, the Kerner Commission[3] recognized the contentious relationship between the Black community and mostly White police departments as a primary cause of the civil unrest in the 1960s and explicitly advised local municipalities to "recruit more Negroes into the regular police" as a potential solution (Commission and on Civil Disorders, 1968). Following the Kerner Commission, the National Advisory Commission on Criminal Justice Standards and Goals (1973) advised that "... Affirmative action should be taken to achieve a proportion of minority group employees in a [police] agency that is an approximate proportion of their numbers in the population."

Beginning in 1969, municipal police departments experienced one of the most aggressive affirmative action programs across the country in the form of court-ordered hiring quotas. The quotas increased the number of Black officers (and women, to a lesser extent) hired by police

[1] For most cities, spending on public safety as a share of total expenditures has remained approximately 11 to 14 percent since the 1960s. See Appendix Figure A3.

[2] For example, the President's Task Force on 21st Century Policing recommended that police departments emphasize diversity in the workplace (Commission and on Civil Disorders, 1968; on 21st Century Policing, 2015).

[3] The Kerner Commission Report was published in February 1968, two months before the assassination of Dr. Martin Luther King, Jr. and the Passage of the Civil Rights Act of 1968.

departments (McCrary, 2007; Miller and Segal, 2012). Although police agencies became more diverse as a result of affirmative action hiring, there is little evidence that diversity has impacted police productivity (McCrary, 2007; Garner et al., 2019). Nonetheless, prior research has found that police performance may vary by race and gender. Empirical evidence shows that in the presence of a diverse police force, Blacks and Hispanics are less likely to be arrested and are subject to fewer stops (Donohue III and Levitt, 2001; Close and Mason, 2007). Although minority police officers conduct a smaller percentage of vehicle searches, they also have greater rates of success, suggesting that minority cops may be able to identify non-White criminals more effectively (Close and Mason, 2007).

Moreover, Hoekstra and Sloan (2020) find that White officers use force 60 percent more than Black officers and use gun force twice as often, particularly when dispatched to neighborhoods with a large fraction of Black residents. Female police officers are similarly less likely to engage in excessive use of force (McElvain and Kposowa, 2008; Rabe-Hemp, 2008; Smith, 2003). Police departments with affirmative action plans are also associated with lower crime victimization rates for Black and White residents (Harvey and Mattia, 2022), with a much larger decrease for Black residents. There is also evidence that affirmative action plans decrease intimate partner homicides for men and women, increase reports of violence against women, and decrease the rate of non-lethal domestic violence (Miller and Segal, 2019).

Despite the potential benefits of affirmative action policies to lower racial disparities in policing outcomes, few studies have rigorously analyzed their effect on racial disparities in the most severe and arguably the most costly use of force–police killings of civilians. This is especially important as affirmative action litigation and police diversity was not a crime prevention strategy but a riot prevention policy, with the goal of reducing police violence, improving police-community relations, and reducing the likelihood of future violent demonstrations in Black communities. Crime prevention or police productivity was a second-order concern. This study will fill that gap by examining the effect of the threat of affirmative action on excessive use of force, as measured by deaths due to legal intervention.[4] Following McCrary (2007), we analyze the threat of court-ordered affirmative action plans between 1969 and 2000. Although Title VII of the Civil Rights Act of 1964 includes provisions for remediating grievances related to discrimination, a series of exec-

[4]We use police killings, deaths due to legal intervention, and policing-related civilian fatalities interchangeably.

utive orders and amendments to the Civil Rights Act expanded the scope of enforcement and the criteria for affirmative action implementation. After the passage of the Civil Rights Act of 1968, there was a sharp increase in the number of civil rights cases brought through private litigation to U.S. District Courts (Farhang, 2010).[5]

Our analysis takes advantage of variation in the location and timing of when a city within a county is threatened with affirmative action litigation. We implement a Difference-in-Difference (DiD) research design within an event-study framework to test if the trends in deaths due to legal intervention change in response to litigation. The data on deaths due to legal intervention comes from the CDC's Vital Statistics, which provide information on cause of death and the decedent's county of residence. The event-study design allows us to test the common trends assumption and visualize the effect of police killings of civilians over time. We confirm previous research findings by providing suggestive evidence that the threat of court-ordered affirmative action quotas increases the employment of Black police officers (McCrary, 2007; Miller and Segal, 2012). Our results are in line with Miller and Segal (2012), who find that litigated departments increase their Black employment shares, but at a lower rate than those facing court-ordered affirmative action.

The potential effect of affirmative action litigation on police killings of non-White civilians is ambiguous. On the one hand, the threat of affirmative action could change the morale and behavior of currently employed police officers. Historically, police departments have been predominantly White, male, and resistant to change. Black professional police organizations seeking changes in employment opportunities initiated most of the lawsuits associated with affirmative action litigation (McCrary, 2007). Therefore, the threat of litigation could have caused feelings of animus towards minority groups due to the prospect of a changing work environment. Consequently, threat of an affirmative action lawsuit and/or expected changes in workforce composition may anger a White male-dominated profession (Hidalgo, 2019), leading to more killings of non-White civilians with no change in killings of White civilians. Moreover, Devi and Fryer Jr (2020) show that police officers can become less engaged when police departments are involved in a federal investigation. Similarly, Ba and Rivera (2019) demonstrate that public or known investigations of police departments alter community-police relations. Thus, federal intervention into local

[5]Although the initial Civil Rights Act was passed in 1968, it was Executive Order 11478 of 1969 that prohibited the federal government from considering race in hiring, and the 1972 amendments to the Civil Rights Act that extended non-discriminatory practices to state and local governments.

police departments could change police behavior, even without changing the racial composition of law enforcement.

On the other hand, such a threat may also lead to the preemptive hiring of more female officers and more racial and ethnic minorities to win lawsuits, leading to greater diversity. To the extent that police departments become more representative of their constituents, we would expect the hostility between police departments and communities of color to decrease, leading to a decrease in negative interactions between the police and the community and fewer non-White deaths. Research suggests that racial stereotypes are used to determine the presence of threats and whether to shoot when decisions need to be made very quickly (Correll et al., 2011, 2002). Although it is not clear whether racial and ethnic minorities will hold fewer racial stereotypes than White officers, exposure to racially diverse peers may lead to improved perception of, or more favorable interactions with, members of minority groups (Carrell et al., 2019). Thus, we would expect police killings of non-Whites by White officers to decrease after their departments grow more diverse. Moreover, a reduction in both Black and White police killings may result if racial and ethnic minorities are more likely to view racial and ethnic neighborhoods within a less threatening context (Correll et al., 2011). Relatedly, if minority police officers are better at identifying and deescalating violent situations among non-White criminals, we would also expect police killings of non-White civilians to decline. Lastly, if minority police officers are less likely to use force in general, it is reasonable to expect police killings of White civilians to decrease as well (Hoekstra and Sloan, 2020).

Our findings indicate that the threat of affirmative action results in fewer deaths of non-White civilians from legal intervention in the long-run. By the year 2000, litigated counties averted roughly 50 non-White deaths at the hands of law enforcement per year. We consider the changing composition of police departments as a potential mechanism and provide evidence that affirmative action litigation increases the racial and gender representation of police departments in the long-run, consistent with previous research (McCrary, 2007; Miller and Segal, 2012). We also find lower property crime arrest rates for non-White civilians and little impact on violent crime arrests, suggesting changes in the frequency of police-civilian contact as another possible mechanism. There is also suggestive evidence that litigation itself, independent of racial composition, may contribute to the decline in non-white police killings. In addition, we find suggestive ev-

idence of short- and long-run decreases in police killings of White civilians, albeit this result is not robust to reasonable specification checks. Although our findings indicate long-run declines in non-White police killings, we also find that the threat of affirmative action leads to a relative increase in non-White killings in the initial year of litigation.

We run numerous specification check to test the validity of our results: included but not limited to, Poisson, ordinary least squares, weighted least squares, and Callaway and Sant'Anna (2020) estimators. The historical data on police killings of civilians from the Vital Statistics data suffers from under-reporting, and the manner of reporting may be problematic for causal inference. We acknowledge this shortcoming and take several steps to assure that our findings are not completely driven by the reporting of police-related fatalities. First, we show that recordings of police killings from the Vital Statistics mirror other non-governmental sources for the years when both are available. We also check to see if there are systematic changes in the reporting of suicides and homicides. Lastly, we use estimates from Collaborators et al. (2021) to account for under-reporting in the Vital Statistics data. We then attempt to estimate how much measurement error has to be induced by litigation to generate our results and we find that the differential in under-reporting between the control and treatment counties would have to more than double for our estimates to go to zero. Or simply stated, treated counties, would have had to record police killings of non-White civilians fairly accurately prior to litigation. There is no reason to believe this to be true (Sherman and Langworthy, 1979; Fyfe, 2002; Loftin et al., 2003). Yet, we still cannot rule out the possibility that the treatment induced under-reporting of civilian deaths.

The remainder of this paper is organized as follows. Section 2 describes our data and methodology. Sections 3 and 4 present our primary findings, along with a series of robustness checks. Section 5 explores potential mechanisms, and Section 6 contextualizes our results. Section 7 concludes the paper.

2 Data and Methodology

2.1 Primary Data Sources

We use a variety of data sources to conduct our analysis. We collect data on civilian deaths involving law enforcement for the period 1960-2012 from the National Vital Statistics System (NVSS)

Multiple-Cause of Death files (US Department of Health and Human Services, 2007). The vital statistics classify deaths by cause, age, race, and county where the death occurred. We use the Vital Statistics data to create race-specific mortality rates for deaths due to legal intervention, as well as race-specific aggregates or counts of deaths due to legal intervention.[6] The Vital Statistics data is particularly useful because it captures deaths caused by law enforcement beyond police shootings. Unfortunately, Vital Statistics data is dependent on local municipalities' reporting of police involvement and therefore grossly under-counts the number of deaths due to legal intervention (Fyfe, 2002; Loftin et al., 2003; Sherman and Langworthy, 1979). In general, government-collected data on police killings accounts for roughly 50 percent of the police-involved deaths in other non-governmental data sources on police killings (Barber et al., 2016; Feldman et al., 2017). Although a herculean effort in data collection has recently increased the number of recorded deaths, there is still a debate about the nature in which data on police killings is collected (Fryer, 2018). Recent data collected by Fatal Encounters and The Washington Post began long after the 1960s and is therefore not helpful for this analysis. Although we are unable to use more recent data sources, Appendix Figure A1 plots police killings in the NVSS and Fatal Encounters data from 2000-2016, respectively, and shows that the two closely track each other over this period. This is also true when we compare both series by treatment group status.[7]

One way to proceed with the Vital Statistics data is to assume that measurement error in the dependent variable is captured by the error term and will not bias our estimates.[8] This approach assumes that the recording of deaths due to policing over time is exogenous to the treatment. This is a strong assumption; policy parameters may be associated with the treatment, changing how local municipalities record deaths due to policing. However, it is unclear if this would lead to an increase or decrease in reporting. Nonetheless, we acknowledge the shortcomings of the Vital Statistics data, and readers should interpret our results considering these caveats.

Information on the litigated departments comes from McCrary (2007). The litigated designa-

[6] We exclude deaths due to legal execution from our calculation of deaths due to legal intervention.

[7] Appendix Figure A2 shows Fatal Encounters and Vital Statistics time series by treatment status. Appendix Figure B18 shows how the two series are correlated by state and how our main estimate change when we drop states where the two series do not track closely over time.

[8] This is true for crime data in general. Myers (1980) shows that between 1970 and 1974, only 1/3 to 2/3 of all crimes were reported. In general, roughly 50 percent of all crimes are reported in the UCR. See Boggess and Bound (1997). Also, Chalfin and McCrary (2018) find measurement error in police employment data, which is more problematic because police are typically an independent variable, and measurement error, in this case, will lead to biased estimates.

tion arises from a series of class-action lawsuits filed across the country, beginning in 1969. This movement led to one of the most aggressive implementations of affirmative action, resulting in a substantial number of court-ordered racial hiring quotas. The dates of litigation in our data vary over time and cover the period from 1969-2000.

2.2 Descriptive Analysis

In Figure 1, we use Illinois's Cook County, where Chicago is located, to illustrate how police killings of civilians change in response to the threat of affirmative action litigation. From the figure, we see that Chicago experienced a racial uprising in 1965.[9] The aftermath of this event resulted in an increase in non-White police-related deaths (consistent with Cunningham and Gillezeau (2019)).[10] In 1970, a class-action suit was filed against the Chicago Police Department (CPD) for discriminatory practices in police hiring. The number of non-White deaths declined in the year the lawsuit was filed and continued to substantially decline the following year before the hiring quota was imposed. Once the quota was imposed, non-White police killings declined further and remained substantially lower than their pre-litigation rates. The Cook County example provides suggestive evidence that behavioral responses to federal interventions in policing–in addition to the composition of the police department–are important in understanding the dynamics of police outcomes in general and police killings of civilians in particular.

To investigate changes in police killings due to court-ordered quotas, we merge litigation data and Vital Statistics data with county demographics data provided by the Surveillance, Epidemiology, and End Results and the County and City Data Books consolidated files from the ICPSR. We remove counties that report having no non-White residents in any year between 1960 and 2012. The final sample consists of 2,985 counties, of which 75 counties are treated. Treatment status is denoted by having at least one city in the county in which a discrimination suit was filed. The timing of the treatment is observed by the year the first city within a county enters into litigation. Once a city is treated, the treatment status does not change over time. Thus, we do not change the treatment status (1) if another city in the county is treated or (2) if the treated municipality does

[9] Data on racial uprisings is from Carter (1986). For this source, a particular incident has to meet several criteria to be classified as an uprising, so the data does not capture the entire universe of possible uprisings in the 1960s.

[10] Racial uprisings peaked in 1968 with the assassination of Martin Luther King, Jr. Following the uprisings of the Long Hot Summer of 1967, the Kerner Commission was established to determine the cause of civil unrest. The commission published its report in 1968–almost two months prior to King's assassination–and found that hostile community-police relations were a major contributor to the start of the uprisings.

not implement an affirmative action plan. This highlights a limitation of our study: the litigation data is at the police department level, but our outcome measures are only available at the county level. Thus, if a county contains a litigated department, we treat the entire county as litigated. Although not ideal, from Figure 1, we see that the relationship between litigation, the share of new Black hires, and police killings can persist at the county level. However, we cannot rule out the possibility that other police departments in the county also change their behavior in response to the treatment. Figure 2 presents a map that contains the counties with a litigated city over the time period in our sample. As one would expect, many of the litigated counties are in the South; however, we do see a non-trivial number in the Midwest and Northeast. Between 1969 and 2000, 75 counties are treated (i.e., litigated), amounting to about 2.5 percent of the counties in our sample but covering 32 percent of the U.S. population (See Table 1). Specifically, over 46 percent of the non-White population in 1960 resided in these locations. By 1977, roughly 40 percent of the non-White population lived in a county with at least one city involved in a dispute over minority police employment.

The relationship between federal involvement in local policing and police killings is also evident during the 1970s. In Figure 3a, we see that in the U.S., the number of non-White deaths per 100,000 non-White residents at the hand of police decreased drastically after the Civil Rights Act of 1968, coinciding with subsequent increases in the number of court-ordered racial hiring quotas. There appears to be relatively little change in the killing of White civilians over time. While Figure 1 provides anecdotal evidence that federal involvement influences police behavior, we see stark changes in police killings by treatment status in Figure 3b. Locations that were eventually treated reported much higher non-White policing-related civilian fatalities prior to the Civil Rights Act of 1968. Both treated and control counties reported rising deaths due to policing prior to 1968; however, the increase is more pronounced in the treated group. After 1968, there is a decline in police killings resulting in a convergence in police killings between the treatment and control groups. By 1977, the gap had significantly narrowed between treated locations and the control group; the two groups reported the same rate of non-White police killings by the 2000s.

This difference in police killings between the treated and control groups becomes even more evident when we directly compare their relative changes over time. Panel (c) of Figure 3 normalizes per capita non-White deaths due to legal intervention to zero in 1968. This also allows for a

more direct comparison between the treated and control groups. It is clear that police killings rose faster in treated locations before 1968 compared to counties in the control group. Although this may have to do with a law-and-order response to racial uprisings, it is highly unlikely since many of the locations where uprisings occurred experienced their first racial uprising in 1968 (Cunningham and Gillezeau, 2019). Moreover, counties in both the treated and control groups experienced a decrease in police killings of civilians after 1968. This is important because it highlights that the control group captures trends in police killings and can plausibly serve as a control group. Lastly, there is a significant drop in police killings of non-White civilians in the treatment group compared to the control group. The decline in police killings becomes starker in panel (d) of Figure 3, which plots the normalized difference in non-White deaths between the two groups. This is consistent with Devi and Fryer Jr (2020) and Ba and Rivera (2019), who find that public or federal intervention influences police behavior. Although they suggest that public or known federal interventions change police performance and decrease productivity, we present suggestive evidence that federal intervention may reduce police killings of non-White civilians in the long-run.

Table 2 presents county-level summary statistics for the variables used in our analysis for the year 1960. Column (1) shows the mean for all counties in the sample. Columns (2) and (3) provide mean characteristics for the treated (i.e., litigated) and non-treated counties, respectively. This basic comparison reveals significant cross-sectional differences between the treatment and control groups, shown by the p-values reported in column (4). Litigated counties are more populous, more educated, and have higher income, and a larger share of their population is non-White. Also, the treated counties are denser and more likely to have experienced an uprising during the 1960s. In fact, roughly 84 percent (63 of the 75) of treated counties experienced at least one violent protest between 1964 and 1971, compared to only about 7 percent (199 out of 2910) counties in the control group.[11] In addition, treated locations have more police killings. This is as expected, given the fact that the treated counties are larger and denser. Moreover, the treated counties tend to be highly urbanized; at least 47 percent of each treated county's population resides in urban areas.

[11] Both McCrary (2007) and Miller and Segal (2012) highlight the importance of the 1960 uprisings in providing the initial impetus for implementing affirmative action programs.

2.3 Event-Study Framework

Our identification strategy relies on the evolution of police killings of civilians prior to a city in a county experiencing litigation; we do not identify causal effects based on cross-sectional differences between litigated and non-litigated counties but on cross-sectional comparison of changes in trends. Given that we cannot observe the counterfactual, causality relies on deaths due to legal intervention evolving similarly in litigated and non-litigated counties before the threat of affirmative action occurs. Thus, it would be reasonable to assume that a common trend, pre-treatment, would have persisted in the absence of litigation. Put another way, we require that the timing of the first litigation be exogenous to pre-existing trends in police use of deadly force and that similar trends would exist in the absence of treatment. If this holds, county fixed effects will account for key cross-sectional differences that are time-invariant, and non-litigated counties will capture trends in police killings over time as well as provide a counterfactual for how police killings of civilians are expected to evolve in the absence of litigation.

A preview of the identification strategy appears in Figure 3, which suggests that treated locations would see an immediate decrease in police killings of civilians due to federal involvement. However, many locations were treated well after 1968, so it is plausible that we will not uncover any impact of litigation on deaths due to legal intervention in the short-run. Also, Figure 3 shows that at least before 1968, treated and non-treated locations have different pre-trends. Lastly, the figure shows that in the absence of treatment, police killings of civilians would decrease over time, capturing a general trend of lower deaths due to legal intervention. However, it is possible that the control group is not the ideal comparison group. We proceed by using the entire sample, but supplement our analysis by restricting the sample to 1) counties with a relatively large population, 2) highly urbanized counties, and 3) counties that experienced at least one racial uprising during the 1960s. It is important to note that Figure 3 only exploits variation in location, not in timing.

To check for parallel trends, we run several tests to examine the influence of pre-existing trends on the timing and location of litigation. Panel (a) of Figure 4 plot the pre-period growth rates in non-White deaths per 100,000 non-White civilians prior to 1969 against the year of treatment. This provides a simple test of whether changes in non-White deaths in the 1960s are correlated with the timing of treatment. It is clear that there is no pattern associated with the timing

of treatment and pre-period growth rates. Counties treated relatively early experienced large increases in non-White deaths, while a significant number of counties experienced a decline in police killings prior to 1969.

We test for differences in pre-period trends in non-White deaths between the treatment and control groups in panel (b) of Figure 4, which plots estimates of the average differences in non-White deaths between treated and non-treated counties prior to 1969. We regress non-White deaths per 100,000 non-White residents on treatment status, year fixed effects, and treatment-by-year effects. The reference year is 1968, thus comparing the difference in non-White deaths to 1968 (analogous to panel (d) of Figure 3). The triangle markers plot the coefficient for the treatment-by-year effects. Although the pre-trend coefficients are negative, the point estimates are statistically indistinguishable from zero. Neither Figure 4a nor 4b shows a distinct difference in the pre-period growth rates. The lack of a statistical difference provides suggestive evidence that non-White deaths due to police intervention evolved similarly in the litigated and non-litigated counties prior to 1969. An additional test of the common trends assumption is embedded in our analysis.

Our main specification employs the following difference-in-differences (DiD) event-study framework:

$$y_{ct} = \alpha_c + \gamma_{r(c),t} + \sum_{j=2}^{9} \pi_j D_c \mathbb{1}\{t - t_c^* = -j\} + \sum_{j=0}^{26} +\phi_j D_c \mathbb{1}\{t - t_c^* = j\} + v_{ct} \tag{1}$$

where y_{ct} is the number of non-White civilian deaths due to legal intervention in county c in year t. The term α_c represents county fixed effects, while $\gamma_{r(c),t}$ are region-by-year fixed effects.[12] D_c is an indicator variable equal to one if the county was ever threatened with litigation. $\mathbb{1}\{t - t_c^* = -j\}$ is an indicator variable equal to one if the observation year is $-j$ years from the date of litigation; $\mathbb{1}\{t - t_c^* = j\}$ is equal to one if the observation year is j years after litigation. We omit $\mathbb{1}\{t - t_c^* = -1\}$ due to collinearity and as a reference year for our analysis. Lastly, t_c^* is the year of litigation (threat) for county c.

Our coefficients of interest, π_j and ϕ_j, capture how the relationship between our measure of

[12] We do not consider state-by-year fixed effects in our main specification, given that many states only have one treated location that is often the largest city or county in the state. Nonetheless, in Appendix Figure B3, we consider an OLS specification that includes state-by-year fixed effects and find our results are robust to this specification.

litigation and police killings varies over time, both before and after litigation. A key assumption of the DiD model is "parallel trends," where trends in an outcome should be common or parallel prior to policy implementation. Given the inability to observe the counties' counterfactual trends, we assume that pre-policy trends persist in the absence of litigation. A test of this is embedded in the event-study framework. Specifically, the difference in pre-policy trends is captured by π_j. Therefore, the common trends assumption is valid if π_j is statistically insignificant and close to zero. The remaining coefficient, ϕ_j, allows us to examine any dynamics resulting from treatment post-litigation. We group event-times for nine years before (i.e., $\pi_{-7} D_c \mathbb{1}\{t - t_c^* \leq 9\}$) and twenty-six years after (i.e., $\phi_{26} D_c \mathbb{1}\{t - t_c^* \geq 26\}$) the policy for the long sample (places treated prior to 1987); we group them for thirteen years after the policy for the full sample.[13] Our primary analysis will focus on the long sample, six years before litigation to twenty-five years afterward.

We consider a variety of specifications to estimate equation (1). Our dependent variable, the number of deaths due to police, takes on non-negative integers with a significant number of zeros. Due to the nature of the dependent variable, ordinary least squares (OLS) can produce biased estimates with the wrong sign or direction. Therefore, we proceed by estimating equation (1) using a Poisson estimator.[14] In addition to the Poisson estimator, we estimate the impact of the threat of litigation on the number of deaths using an ordinary least squares estimator, and we estimate the effect on deaths per 100,000 civilians using weighted least squares (WLS).[15] Police encounters that result in death are rare occurrences. Changes in mortality rates, consequently, will be larger for counties with smaller populations due to the fact that most counties have very few (or zero) deaths. Therefore, we use the 1960 population to give more weight to larger counties that experience wider variations in police killings of civilians.[16] We also contend with the possibility that the estimates obtained from the standard two-way fixed effects (TWFE) difference-in-differences (DiD) model may be biased due to treatment effect heterogeneity since litigation varies

[13] Since the last treated unit experienced litigation in 2000, the maximum number of post-treatment years in a balanced panel for this timing group is 12 years (i.e., until the last year of our sample period, 2012). Although the event-study design can estimate treatment effects for an unbalanced panel, due to composition bias, it can lead to misleading estimates. Specifically, the average treatment effect can change over time because of the composition of timing groups are changing. This is even more problematic if there is heterogeneity in the treatment effects across timing groups or heterogeneity over time.

[14] In our Poisson model, we use population (by demographic group) to account for differences in exposure related to county size and demographic make-up.

[15] We examine both raw and an age-adjusted mortality rates.

[16] Population weights also correct for heteroskedasticity related to county size in the error term. The population in 1960 will be used as weights.

over time.[17] Thus we also present results using the Callaway and Sant'Anna (2020) (CS) estimator.

We summarize event-study estimates with joint treatment effects using the following equation:

$$y_{ct} = \alpha_c + \gamma_{r(c),t} + \sum_{\tau} \widetilde{\pi}_j D_c \mathbb{1}(t - t_c^* \in \tau) + \sum_{\omega} \widetilde{\phi}_j D_c \mathbb{1}(t - t_c^* \in \omega) + v_{ct} \tag{2}$$

where τ accounts for the pre-period event-years $j \leq -9$ and $-8 \leq j \leq -2$, while ω accounts for the short-run ($0 \leq j \leq 7$), medium-run ($8 \leq j \leq 15$), and long-run ($16 \leq j \leq 25$) event-years, as well as $j \geq 26$. This specification allows for testing the joint significance of pre-period trends and will also be used to summarize our robustness checks results.

3 Results

We start by describing the event-study estimates, which allow us to analyze pre-trends further while also examining the subsequent dynamic effects post-litigation. We plot the pre-and-post-treatment effects from equation (1) with a solid line and circle markers for the long sample. The 95-percent confidence intervals are shown with dashed lines and circle markers, and the gray shaded area identifies the 90-percent confidence intervals. Confidence intervals are constructed from robust standard errors, clustered at the county level to address over-dispersion. We also plot the results from the full sample with a solid red line. We interpret our effects as an "intent-to-treat" estimate because we only know when the threat to litigate occurred (we are unable to identify if or when a city in a treated county implemented an affirmative action hiring program), and we are unable to distinguish between the actions of the treated city in the county versus non-treated cities within the same county.

3.1 Primary Findings: Non-White Deaths

Figure 5 graphs the pre- and post-treatment effects of litigation on the number of non-White civilian deaths due to legal intervention from our Poisson model. As indicated in Figure 4, we see no pre-trend difference in police killings of non-White civilians. In the year of treatment, the post-treatment effect for event-year 0 is positive but imprecisely estimated in all of our specifications. A possible explanation for this increase may be that the threat of affirmative action led to an initial

[17] Standard DiD techniques may have biased parameter estimates if the treatment effects change over time (Goodman-Bacon, 2021), which may be true even for dynamic specifications (Sun and Abraham, 2020). Appendix Figure B13 plots the 2x2 difference-in-difference estimates and their associated weights from the conventional two-way fixed effect model (Goodman-Bacon, 2021).

backlash effect aimed at non-Whites, consistent with the literature on conflict theory and racial threat (Jacobs and O'brien, 1998; Smith, 2003). However, given the overlap between affirmative action litigation and social unrest, we cannot rule out other factors contributing to this increase during this time. Afterward, post-treatment effects decrease and become negative by event-year 2 for all specifications. From Figure 5, we see that the post-treatment effects are negative and statistically significant starting in event-year 5 and decrease slightly over time. On average, 1.457 non-White civilians were killed by police in a treated county in the year prior to the threat of litigation. According to the point-estimate for event-year 11, non-White deaths decreased by 44 percent, which amounts to 45 fewer deaths in 70 treated counties.[18] Our effect sizes are comparable to those in this literature. For instance, our results are not far from Campbell (2021) who finds that Black Lives Matter Protests lead to a 15 percent decrease or approximately 200 fewer deaths four years after treatment. Conversely, we only find a seven percent decrease in non-White deaths after four years. Masera (2021) finds that police militarization increased overall police killings by 8.4 percent or within seven years or an additional 64 deaths a year. These are more deaths (in magnitude) than what our results suggests are averted in any given year from Affirmative Action litigation.

We summarize the joint pre- and post-treatment effects in Table 3, which presents estimates from equation (2). Column 1 estimates equation (2) using OLS for the number of deaths due to policing; column 2 presents joint effects from the Poisson model; column 3 reflects the WLS model where the dependent variable is a mortality rate;[19] column 4 reports estimates using the Callaway and Sant'Anna (2020) estimator. For non-White deaths, the pre-treatment effects are statistically insignificant in all four models. This provides evidence that the pre-treatment effects are not jointly statistically significant. In the short-run, post-treatment effects are negative and not statistically significant in all three models. The joint effect masks the initial increase in non-White deaths discovered in Figure 5. The estimates are also negative and statistically or marginally statistically significant in the medium and long-run. If we interpret the long-run coefficients, the

[18] The coefficient in event-year 11 is -0.582, which implies, on average, a 44% difference between the number of deaths in event-year -1 and event-year 11. There are 70 treated counties in the long sample, resulting in an estimated 45 fewer non-White deaths.

[19] Appendix Figures B2a and B2b display event-study estimates from our OLS and WLS models, respectively. Mortality rates are adjusted by age. See Appendix Figure B4 for unadjusted mortality rates. WLS models use non-White population in 1960 as weights. OLS models include the log of non-White population as an independent variable.

threat of affirmative action litigation is associated with a 40 percent decrease in non-White deaths in the Poisson model. The WLS model suggests a 37 percent decrease in non-White deaths, while the OLS model indicates a 46 percent decrease. In Appendix Table C2, we limit the sample to counties that report at least one non-White death over the sample period to check the validity of our estimated joint effects. This restriction produces post-treatment effects that are similar in magnitude.[20] In column (4) of Table 3 we present the he CS estimator results, and find that our estimates are very similar to our OLS findings; we see a similar statistically significant long-run decrease in non-White deaths due to police.[21]

Our results provide evidence that after litigation, police killings of non-White civilians gradually decrease in the long-run. Figure 6 plots the number of non-White deaths prevented in litigated counties over time. We estimate the number of deaths averted by first transforming the pre- and post-treatment effects from the Poisson specification, which is $(\exp^{\widehat{(\pi;\phi)}_j} - 1) \times 1.457$, and then maps event years to actual calendar years. Using this approach we estimate that in the year 2000, litigated counties averted nearly 50 non-White deaths due to legal intervention.[22] The gradual decrease in police violence slightly mirrors McCrary (2007), who finds a gradual increase in diversity five years after litigation, which coincides with the decline in police killings we see five years after treatment in Figure 5. The key difference is the reduction in police killings is persistent but levels out 11 years after treatment, while McCrary (2007) finds a steady and continual increase in diversity up to 25 years after treatment. Although we later examine whether changes in the racial composition of police departments may be an important mechanism in lowering police killings of non-White civilians, we also consider whether the threat of federal intervention in and of itself may have had an effect police killings of non-White civilians.[23]

3.2 Primary Findings: White Deaths

Appendix Figure B1 and Appendix Figures B6a and B6b plot event-study estimates for White deaths at the hands of police. In general, results are similar across the OLS, WLS, and Poisson spec-

[20] We find similar results when we limit the sample to counties that report at least one non-White death prior to 1969. See Appendix Table C3.

[21] See Appendix Figure B15 for the corresponding event-study figure.

[22] Our results suggest that the threat of affirmative action prevented over 800 non-White deaths between 1970 and 2000.

[23] It is important to note, we do not observe when a municipality implements an affirmative action program. As shown in Figure 1, deaths can decrease prior to a significant shift in the racial composition of a department. Relatedly, it is possible that the threat of federal intervention caused other departments in the county to change their behavior, without causing the treated department to do so.

ifications; however, the common trends assumption is likely violated for White deaths. Estimates plotted in Appendix Figures B6a and B6b provide evidence that the common trends assumption does not hold across specifications.[24] Although we cannot imply causality for all of our models, White deaths immediately decreased after treatment. Post-treatment effects for White civilians are negative and statistically significant, or marginally statistically significant, in certain event-years. When using the point estimate for event-year 0, White deaths decreased by approximately 43 percent (36 fewer White deaths). The decrease in police killings in Figure B1 of White civilians runs counter to Cunningham and Gillezeau (2019), who find that police killings of both White and non-White civilians move in the same direction after a violent protest. However, these results should be interpreted with caution, given the violation of the pre-trends assumption across our specifications.[25] It is plausible that White deaths increased in the year before treatment but afterward reverted to pre-ligation trends.[26] Therefore, we mainly focus on non-White civilian deaths when discussing our findings.

Given that we see a suggestive initial increase in non-White police killings in Figure 5 and an immediate decrease in White police killings post-litigation, we can rule out the possibility that officers indiscriminately increased force in response to the threat of litigation. The initial increase in racial disparities in police violence could indicate a distinct change in police behavior in response to federal intervention. Appendix Figure B16 plots pre-and post-treatment effects from a triple difference model and shows an initial statistically significant increase in racial disparities in police killings at event-year zero. After the initial increase in racial disparities of police killings, post-treatment effects decrease and eventually become negative but are not statistically significant. The initial disparity is counter to the finding of Fryer (2019), who finds no racial differences in police use of force.

[24] The point estimates for event-years -3 and -4 are statistically or marginally statistically significant, depending on the model. In the Poisson model, event-year -5 is marginally statistically significant.

[25] Table C1 presents the joint treatment effect estimates for White deaths.

[26] The 1972 Multiple Caused of Death files contains a 50 percent sample of all deaths. Following Bailey and Goodman-Bacon (2015), we multiply the death count by two in that year. We check whether this data manipulation influences our Poisson specification results in Appendix Figure B24 and Appendix Table C4 and find similar results across alternative specifications.

4 Robustness Checks

Our results suggest that in the year of litigation, racial disparities in police killings increase, while several years after the threat of litigation, non-White deaths decrease. This result holds across multiple specification checks. However, it is clear from Table 2 that the treatment group differs drastically from the control group. Although our empirical strategy can identify a causal relationship despite key cross-sectional differences, it is reasonable to be concerned about the interpretation of our results given the control group. Figure 7 present joint effects for a series of robustness checks using the Poisson model for non-White deaths.[27] The estimated joint effects are denoted by a circle marker, while a bold horizontal line indicates a 95-percent confidence interval. The columns of each figure display pre-treatment effects (2–8 years before treatment), short-run effects (0–7 years after treatment), medium-run effects (8–15 years after treatment), and long-run effects (16–25 years after treatment). The first row presents joint effects from columns 2 of Table 3; the remaining rows report joint effects from a series of robustness checks.

4.1 Treatment Status

The current empirical strategy takes advantage of variations in the timing and location of affirmative action litigation. In row 2, we restrict the sample solely to treated counties. Restricting the sample in this manner only exploits variation in timing for estimating treatment effects. For non-White deaths, pre-treatment effects are statistically indistinguishable from zero, and the short-run effects are negative and close to zero. Post-treatment effects in the medium- and long-run have the same sign as those from the main specification. However, they are larger, suggesting that the control group helps capture changes in police killings of non-White civilians over time. [28]

4.2 County Characteristics

The smallest non-treated county in our sample had 57,000 residents in 1960, while only four treated counties had less than 100,000 residents. We first attempt to account for key cross-sectional difference by employing a semi-parametric reweighing design (Abadie, 2005; Cunningham and

[27] See Appendix Figure B7 for White deaths.

[28] In Appendix Figure B12, we examine the impact of the threat of litigation on nearby (contiguous) counties. It is possible that, agencies in neighboring counties change behavior, to prevent or avoid future litigation. However, when we assign the treatment status to contiguous untreated counties and compare them to non-contiguous untreated counties, we find little evidence of changes in police killings of civilians.

Gillezeau, 2019). We use propensity scores, $p(x)$ to reweigh counties in the control group by their inverse propensity scores, $\frac{p(x)}{1-p(x)}$, to balance the distribution of covariates across groups (DiNardo et al., 1996). As shown in Appendix Table C5, the reweighing scheme results in a control group that mirrors the treatment group across multiple dimensions, and as shown in row 3 of Figure 7, this specification produce similar results as Figure 5.[29] We further explore how the composition of counties in the control group influence our results. In row 4, we restrict the sample to counties with a population greater than 100,000 residents and in row 5 we limit the sample to counties where the proportion of residents residing in urban areas was higher than the median share of urbanization in 1960.[30] In both sub-samples, the inclusion of only large urban counties produces pre-treatment effects near zero for non-White deaths and post-treatment effects are similar in magnitude and statistical significance to our original specification for non-Whites.

Litigation is highly correlated with uprisings. McCrary (2007) finds that uprisings are associated with the timing of treatment. Such uprisings may lead to concerns of mean reversion after the exogenous shock of riots. To rule this out, row 6 of Figure 7 restricts the sample to counties that experienced at least one racial uprising in the 1960s. This limits the sample to 63 treated counties and 199 non-treated counties. Once again, our estimates for non-White deaths are robust to this restriction and are of similar size to the models restricting the sample to highly urbanized counties or counties with a population greater than 100,000. This is not surprising since the restrictions are highly correlated. Bigger counties tend to be urbanized and have experienced uprisings in the 1960s.[31]

In Rows 7 and 8, we split the treatment group into counties where the share of the Black population in 1960 is below and above the median share for the treatment group;[32] removing the above-median group in row 7 produces joint treatment effects that are estimated less precisely, but similar to those in row 1. Alternatively, row 8 displays the results for removing the below-median group. Both, the medium and long-run effects are statistically significant in this specification. The

[29] Event-study estimates using this alternative re-weighting procedure are shown in Figure B5. We also report estimates from a synthetic difference-in-differences design proposed by Arkhangelsky et al. (2021) in Table C9, and find that our results for non-White police killings are robust to this model.

[30] In row 4 and 5 we have 71 and 75 treated counties (212 and 853 counties in the control group) respectively.

[31] In these robustness checks, the common trends assumption is violated for White deaths. Figure B7 shows that the large negative post-treatment effects are driven by locations where the common trends assumption does not hold.

[32] Appendix Figures B35 and B36 show that White population growth slowed dramatically in the decades following litigation, increasing the number of non-White civilians relative to White civilians in treated counties compared to non-treated counties.

medium-run effect is similar in size to row 1, but the long-run effect is smaller than our baseline estimate. These results show that locations with a large Black population do not drive the decrease in non-White deaths.[33]

Lastly, we test for heterogeneous effects by restricting the treatment group to 1) counties treated prior to 1980 and 2) counties treated after 1980. Under both scenarios, the control group includes counties that were never treated. It is reasonable to assume that the increase immediately after treatment occurred in the aftermath of riots, where tensions between the police and non-White community were palpable. Appendix Figure B14 plots event-study estimates for counties treated prior to 1980 and counties treated after 1980. The later-treated group produces estimates that are noisy and volatile, but the pattern of a long-run decrease in police killings exists for both groups. However, in earlier treated counties, we see the pattern emerge from our baseline model of an initial increase in non-White deaths followed by a persistent decrease over time. Therefore, we cannot rule out that the initial rise in police killings of non-Whites in the year of treatment is related to uprisings that occurred around the same time, as the later-treated group was treated well after the 1960s.[34]

Together, these findings highlight the importance of having non-litigated counties as the comparison group: post-treatment effects are smaller when including all non-litigated counties. Specifically, including smaller counties helps capture national trends in police killings. Many of the smaller counties are not treated but are policed by majority-White police departments. Feigenberg and Miller (2018) find that heterogeneity in a county's racial composition is associated with a more punitive criminal justice system; relatedly, Cunningham et al. (2020) also show that heterogeneity in a county's racial composition is associated with more police killings of non-White civilians. Appendix Figure B10 shows that the long-run decrease in non-White deaths due to the threat of affirmative action is driven by treated locations outside of the South. When restricting our analysis to southern counties, statistically significant increases in non-White killings are evi-

[33]This analysis stratifies the treatment group by the county's non-White population. Likely more important is the percentage of the Black population residing in treated cities within a county. Appendix Figure B8 shows that our results are driven by treated locations where a significant proportion of non-White residents reside in treated cities. Treated counties with more than one treated city, or counties where the treated cities comprise a significant proportion of the county's population, provide similar results to counties with one treated city, or where the treated locations comprise a relatively smaller proportion of the county's population. Also see Appendix Figure B9.

[34]In Appendix Figure B25, we also rule out the possibility that Pattern and Practice Investigations and federal consent decrees in the 1990s and 2000s drive our results.

dent two years after litigation and then again seven years after.[35] The regional heterogeneity of our findings highlights the importance of using region-by-year fixed effects to capture the cultural differences and heterogeneity in racial composition that contribute to criminal legal system outcomes.

4.3 Threat to Internal Validity

Police killings of civilians are severely under-reported in government agency data, including our primary source.[36] According to Collaborators et al. (2021), roughly 50 percent of police killings are not reported and under-reporting is more likely to occur for non-White decedents. In our case, measurement error in the dependent variable would lead to imprecise estimates by increasing the variance and the likelihood that we are unable to uncover a statistically significant relationship. However, if treatment changes reporting behavior, our results would be biased. For example, if the threat of litigation increases oversight and decreases under-reporting, then our estimates understate the impact of federal intervention on police killings of civilians. However, if federal intervention increases under-reporting, then our estimates overstate the impact.

Unfortunately, we are unable to test directly for changes in reporting behavior. However, to observe changes in reporting behavior indirectly, we examine the impact of the threat of affirmative action on suicide deaths. Suicides serve as a legitimate test of changes in reporting behavior, as there are no direct linkages between affirmative action litigation and suicides. However, both police killings of civilians and suicides depend on local inter-agency cooperation for reporting in the Vital Statistics and both are under-reported. Appendix Figure B27 displays pre- and post-treatment effects for suicides per 100,000 residents by racial group. In both panels, we see no systematic change in suicides after litigation. Non-White suicide post-treatment effects are statistically insignificant and are both positive and negative.

If police killings are under-reported, where are the missing deaths? According to Feldman et al. (2017), in 2015, 86 percent of non-accounted police killings were recorded as homicides/assaults. Therefore, if local officials are changing reporting behavior, it could be captured by changes in homicide victimization. Panel (a) of Appendix Figure B28 report treatment effects

[35] Pre-treatment effects violate the parallel trends assumption, so the results in Appendix Figure B10 should be interpreted with caution.

[36] We reproduce Figure 5 using data on police killings of civilians from the FBI's Supplemental Homicide Report (SHR) and we find similar size long-run effects. See Appendix Figure B19.

for the impact of the threat of affirmative action litigation on non-White homicides per 100,000 non-White residents.[37] According to Appendix Figure B28, non-White homicides decrease four years after treatment; the point estimates are statistically significant in event-years 4 and 5.[38] However, eventually the post-treatment effects begin to increase and eventually become positive. The post-treatment effect twenty years post-litigation is positive and statistically significant. In general, there is suggestive evidence that litigation may reduce non-White homicides in the short-run, while welfare gains dissipate in the long-run. However, there is little evidence that the decrease in police killings of civilians is reflected in higher homicide victimization rates. Nonetheless, we acknowledge that diversifying the police may lead to changes in policing and criminal behavior that lower homicides (Miller and Segal, 2014). It is possible that the short-run decrease in homicides is an outcome directly connected to the treatment. However, given that police killings are relatively rare compared to homicides it is highly unlikely that we would uncover changes in reporting by examining homicides. Panel (b) of Appendix Figure B28 and Figure B29 show how combining homicides and police killings has little to no effect on the point estimates relative to solely focusing on homicides.

Together these results show no systematic change or pattern with regards to reporting suicides or homicides and provide additional evidence that we are capturing changes in police behavior in response to the threat of litigation and not changes in reporting behavior due to interagency cooperation or the lack thereof. It is important to note that although we find little evidence of changes in reporting of suicides, we cannot rule out changes in cooperation with regards to reporting police-related homicides. The incentive to misreport police related civilian deaths differs from that of suicides as a large number of police killings may yield substantial repercussions.[39] This is evident as the number of police killings in the Vital Statistics data remains severely undercounted Collaborators et al. (2021). In addition, the rarity of police killings relative to the number

[37] To capture pre-period differential trends in homicides, the analysis includes urban-by-year fixed effects in equation (1). Urban-by-year fixed effects are constructed by interacting urban status indicator variables with year indicator variables. Urban status is defined as the following: percent of the population residing in urban areas $(\mu) : 0, 0 < \mu < 25, 25 \leq \mu < 50, 50 \leq \mu < 75, 75 \leq \mu \leq 100$. Urban-by-year fixed effects capture unobserved heterogeneity that varies across time and urbanicity. Not including urban-by-year fixed effects results in statistically significant pre-treatment effects for non-White homicides.

[38] Figure B28 plots the estimates for white homicides.

[39] Conversely, any evidence of misreporting police killings reported by a police department may also lead to scrutiny, which may incentivize accurate reporting. However, recent data suggest that even if this were the case, it is not enough to outweigh the factors that contribute to underreporting.

of homicides may result in only a slight variation in the measurement of homicides, plausibly masking any impact from under-reporting.

We address the potential issues from under-reporting in the vital statistics data in a few ways. First, we use estimates of the bias in reporting of police killings by state and race from (Collaborators et al., 2021). Using their corrected estimates of police killings from 1980 to 2019, we predict police killings of civilians by race and state for 1960 to 1979. Appendix Figure B17 plots the complete time series of police killings of civilians from 1960 to 2016. We assume that each county's contribution to the state estimate of police killings of civilians by race is constant over time. Therefore, our new measure of police killings at the county-level is dependent on year-to-year variation in the Vital Statistics data, the estimated bias in police killings by state and race, and the county's contribution to the state level number of police killings in the Vital Statistics data. We do this to avoid recording deaths in counties that never report a police killing between 1960 and 2016. Appendix Figure B32 replicates Figure 1 but includes the updated measure of police killings of non-White civilians. In general, the new estimate mirrors the original, with police killings approximately twice as large as before. Appendix Figure B20 plot pre- and post-treatment effects using the new measure of police killings. Interestingly, panel (a) shows a much larger decrease in police killings of non-White civilians, while panel (b) report similar estimates for White deaths. For non-White deaths, the long run estimates are more than 60 percent larger than the original treatment effects. Although accounting for under-reporting results in larger treatment effects, the new measure of police killings and the estimates are contingent on the treatment being exogenous to reporting.

Next, we explore the degree of measurement error that has to exists to reproduce our results. We first run a placebo test by randomly assigning treatment status and treatment year to 60 counties where 1960 uprisings occurred but were not exposed to litigation. We then repeat the assignment 250 times and report the joint long-run effects. Panel (a) of Figure B21 plot event study effects from the first iteration and panel (b) plot long-run joint effects compared to our OLS estimate. In general, we are unable to obtain an estimate close to our long-run joint effects.[40]

[40] In Appendix Figure B22 we also consider simulations that add measurement error from a uniform distribution to our non-White police killing outcomes using the 60 placebo riot counties. This figure shows that it would take introducing measurement error equivalent to 200 percent of the standard deviation to obtain our true long-run treatment effect. Additionally, in Figure B23 we use the measurement error from comparing NVSS to Fatal Encounters data in overlapping years. We find that fewer than five percent of our iterations result in a long-run estimate comparable to or larger (in magnitude) than our main result.

To further address the measurement error concern, we take a county's difference between the NVSS non-White police killing number and the adjusted value using the correction from Collaborators et al. (2021), and iteratively add a percent of this difference to the NVSS measure of non-White police killings for treatment counties. Panel (c) of Figure B21 presents the results from this analysis. It would take a correction of roughly 46 percent to drive our results to zero. Put another way, litigation would have to cause under-reporting to increase by 46 percent. The analysis in panel (c) assumes that the treated and control groups report accurately before treatment. However, we know under-reporting is an issue for both treated and control units. Therefore, we attempt to examine what happens when we vary the differential measurement error between the treatment and control group. We take a ratio of the average error in police killings and the average number of deaths using the correction from Collaborators et al. (2021) in treated (or litigated) counties and divide by the same ratio for the control counties.[41] Interestingly, the control group within our sample exhibits a higher rate of under-reporting for non-white deaths resulting from police encounters in contrast to litigated counties. Specifically, the difference in measurement error between the treatment group and control group is 10 percent. Panel (d) of Figure B21 shows what occurs when we vary differential measurement error. For example, 100 signifies that the initial differential in misreporting was doubled the current rate, transitioning from 20 percent before treatment to 10 percent differential in reporting after treatment. Panel (d) reveals that litigation has to change relative reporting by more than 400 percent to push our estimates to zero in the long run. Another interpretation is that the litigation in treated counties would have had to induce more than four times the difference in relative measurement error, transitioning from 53 percent before litigation to 10 percent after treatment. To attribute our findings entirely to changes in reporting, panel (d) implies that treated locations reported comparatively accurately relative to the control group prior to litigation.

Together, these exercises suggest that it takes a considerable amount of measurement error

[41] We estimate relative differential reporting with the following equation:

$$differential = 1 - \frac{\frac{error_t}{CorrectedDeaths_t}}{\frac{error_c}{CorrectedDeaths_c}}$$

where $error$ refers to the average difference between Collaborators et al. (2021) and NVSS measure of police-related deaths for the treatment group, t, or the control group, c; and $CorrectedDeaths$ refers to average number of police-related deaths using estimates from Collaborators et al. (2021).

induced after treatment to reproduce our results. However, it remains plausible that litigation changed reporting in such a way that we overstate the impact of litigation on police killings of civilians. Additionally, we acknowledge the potential for litigation to improve reporting. Assuming equal levels of measurement error in reporting prior to the intervention, we are likely underestimating the influence of litigation on police violence. Nonetheless, these exercises highlight the importance of reforming how data is collected for police activities and the use of force. Currently, there does not exist a governmental source that accurately captures the degree to which police officers engage in lethal and non-lethal use of force nationally. Thus, making it difficult to conduct in-depth policy analysis or provide serious solutions regarding discriminatory policing, use of force, and police-community relations.

5 Possible Mechanisms

Now, we explore possible mechanisms that may explain the non-trivial decrease in police killings of civilians in response to the threat of litigation.

5.1 Employment Effects

The threat of litigation and the implementation of affirmative action can change the racial and gender composition of police departments (McCrary, 2007; Miller and Segal, 2012). Employment effects are important because, as previously mentioned, a police department's demographic composition is a potential mechanism through which non-White police killings decrease over time in counties with at least one litigated police department. Prior research suggests that Black and female officers are less likely to use force (Hoekstra and Sloan, 2020; McElvain and Kposowa, 2008; Rabe-Hemp, 2008; Smith, 2003). Miller and Segal (2012) find that litigated departments increase their Black employment shares, but at a lower rate than departments with court-ordered affirmative action. They also find that litigation leads to increases in the hiring of female officers.

Nonetheless, our litigation data is limited because we do not have information on dispositions of lawsuits, nor do we have data on the imposition of court-ordered mandates. Moreover, we do not have information on the year-to-year racial and gender composition of police departments. However, we can use police employment data from the Law Enforcement Management and Administrative Statistics (LEMAS) and the Law Enforcement Officers Killed and Assaulted (LEOKA)

files to examine the compositional effects of police departments over our analysis period. LEMAS is obtained from the Bureau of Justice Statistics (BJS), and the LEOKA data comes from the Uniform Crime Reporting (UCR) program. While LEOKA data are available after 1971, the UCR only contains police employee data by gender. It does not have information by race. LEMAS data, on the other hand, provides employment information by gender and race, but the data collection is infrequent and did not begin until 1987. Appendix Figure B30 plots the share of Black and female officers over time using the LEMAS data. Both groups have gained representation since 1987, with the female and Black share being very similar. Using these two data sources will allow us to get a clearer picture of changes in the police force composition over our sample period.

We estimate multiple models to understand the effect of litigation on police composition. We start with the LEMAS data and regress the county share of Black and female sworn officers, respectively, on three dummies that indicate whether a police department was treated in the 1970s, 1980s, or 1990s along with region and year fixed-effects. We report these estimates in Appendix Table C6, where the reference group is counties that were never treated. These estimates indicate that counties treated earlier (in the 1970s and 1980s) have on average significantly larger shares of Black and female officers relative to never treated counties. There is no statistically significant difference in the share of Black officers between counties treated in the 1990s and those never treated.

Next, using our main specification (see equation (1)), we use the LEOKA employee data to calculate changes in the share of sworn female officers in a county (as an indication of the changing dynamics or composition of police departments) over our sample period. We plot the point estimates from our event-study approach in Appendix Figure B31. The top figure, B31a, plots the dynamic effects of the threat of litigation on the number of total sworn officers (per 1,000 residents), and panel B31b plots the estimates for the share of female officers. The figures show that pre-treatment effects are insignificant between treated and control counties. In addition, the findings indicate an immediate significant increase in the number of total sworn officers in treated counties relative to control counties following the threat of litigation. While the number of sworn officers seems to be increasing, in general, over the post-treatment period, the increase is not significant until event-year 24. Panel B31b also shows that the share of female officers increases immediately after litigation and continues to do so for all post-treatment years. Together, these

results indicate that litigated departments are becoming more diverse over our sample period.

Finally, we consider a method similar to Miller and Segal (2014) in which we regress the county share of Black and female sworn officers on the years since a county experiences litigation. Our results are presented in Table C7 and are similar to those of Miller and Segal (2014), where the share of Black officers increases as the number of years since litigation increases. We find similar results for female officers. We then split up the LEMAS data and run this model for each year of the LEMAS that is available. Appendix Figure B34 plots the point estimates and 95-percent confidence intervals on the variable, capturing years since litigation for each cross-section of LEMAS. Years since litigation is significantly and positively associated with a larger share of Black and female officers for almost every year of the LEMAS data.

To see how important diversity is in explaining our results, we use LEMAS data to account for the racial representation gap in police departments. We compare the racial composition of police departments to the racial demographics of treated cities twenty-five years after treatment since we do not have data on the initial racial representation gap. We use the most recent census year, closest to the twenty-five year mark, to obtain the racial composition of treated cities. For example, for a city treated in 1972, we use LEMAS data from 1997 and city demographic characteristics from 2000. Although not ideal, this provides a close approximation of representativeness in police departments at the end of our event-year window.[42] We then rank treated locations by their representation gap and compare locations above and below the median level of representativeness to non-treated locations.[43] Figure 8 plots event-study estimates for non-White civilian deaths. For the below median group, we see a large and immediate decrease in police killings of non-White civilians. We find little evidence of a long-run decrease in police killings for the above median group. The above median group consist heavily of counties from the south, suggesting that litigation may have been most effective at increasing diversity in locations with a large non-White applicant pool. However, it is likely that regionality influences police culture, resulting in heterogeneity in the treatment effect as previously noted.[44]

[42] We define representitiveness by the Black share of officers in the police department divided by the Black population share.

[43] Both the above and below-median groups will have the same control group.

[44] We do not have data on the initial racial representation gap, we do have data on Female representation from the early 1970s. Although a poor proxy in our context, this period also saw litigation against police departments that lacked female representation (Miller and Segal, 2014). Figure B37 shows the correlation between the share of Black and female officers in the LEMAS data. There is a positive correlation between the two variables. Using the fact

Panel (a) and (b) of Figure 9, compares pre-period (2-6 years before treatment) average police killings to long-run average police killings (20-25 years after treatment) against the share of police officers Black and the representation gap 25 years after treatment. There is weak evidence that the treatment had the greatest impact in locations with a lower Black share of officers and suggestive evidence of a tipping point. Initially, as the share of Black officers increase, the long-run effects grow more negative. Once the share of Black police officers grows beyond 30 percent, the decrease becomes less negative and eventually turns positive. Relatedly, the decrease is the largest in locations with the largest representation gap. A possible takeaway is that litigation matters the most when the representation gap is largest, and the share of Black officers is relatively low. However, diversity may not be enough to overcome institutional norms with long histories of state-sanctioned violence (Bjuggren et al., 2023). Although police departments continued to diversify post-litigation, the steady leveled long-run decline in non-white police killings that we find may also be attributed to diminishing returns in diversifying police departments. It is important to note that the level of diversity in a police department is endogenous to the treatment. However, this exercise suggests that federal intervention to promote diversity is effective even when closing the representation gap may be challenging.

5.2 Arrests

Our findings suggest that the threat of affirmative action leads to long-run declines in the police killing of non-White civilians as well as to changes in the demographic composition of police agencies. Another possible channel for a decrease in police violence is a change in policing due to federal intervention. Changes in policing could be reflected by a decrease in police contact for less serious crimes. Weisburst (2019) shows that racial disparities in police contact drive racial disparities in use of force incidents. Furthermore, McCrary (2007) finds evidence of a negative relationship between the share of Black arrestees and the threat of litigation. This could occur due to changes in police departments' racial and gender composition over time or as a reaction to federal intervention (not necessarily changing racial composition).

Using our main specification, equation (1), we plot event-study estimates for arrests by crime

that the two measures are highly correlated, we now run an analysis that stratifies our primary model by above and below the median in the female share of officers using the LEOKA data in Figure B38. We use the first year for which a department has the information available, which is 1971 for most departments. The point estimates are similar for both but more consistently statistically significant for female representation below the median.

type in Figure 10.[45] Panel (a) of Figure 10 provide evidence of a statistically significant decrease in non-White arrest rates in treated counties threatened with litigation beginning in event-year 11, the same year that police killings of non-Whites begin to significantly decrease. [46] When we disaggregate arrests rates by type of crime, a different picture emerges. Panels (b) and (c) of Figure 10 shows that the reduction in non-White arrests is driven by a reduction in arrests for property crimes. According to the point-estimate for event-year 11, property crime arrests decreased by 16 percent. This suggests that the decrease in arrests and police contact may have led a reduction in police-related fatalities. If the decrease in police killings is entirely driven by the decrease in property crime arrests, then it takes 394 fewer non-White arrests for property crime per 100,000 non-White residents to avert one non-White death. Or a one percent decrease in non-White property crime arrests decreases non-White police killings by 2.75 percent.

In panel (d) we check to see if the changes in property arrests are offset by changes in quality of life offenses. Quality of life offenses includes public intoxication, liquor law violations, disorderly conduct, gambling, suspicious behavior, vandalism, and vagrancy. We find no evidence that the decrease in property crime arrests are offset by changes in arrests for low-level offenses. The point estimates are negative and gradually decreasing but never statistically significant.[47]

Lastly, we explore how representation influence arrests analogous to Figure 8. We compare treated locations above and below the median level of representativeness 25 years after treatment to non-treated locations (irrespective of racial representation gap). In Appendix Figure B33, we see both violent and property crime arrests decrease in locations that were not as successful in closing their representation gap. The decrease in violent crime arrests is statistically significant for only a few event-years. Property crime arrests decreases for both above and below the above median treated counties, but we only see consistent statistically significant decreases in event-years for locations with above the median level of representativeness. This provide suggestive evidence that changes in police behavior reflected in decreased police contact may explain our

[45] Because the sample is limited to counties identified by McCrary (2007), which is comprised of agencies that report at least 52 years to the UCR program between 1960 and 2016 (UCR sample), we re-estimate our primary specification of police killings of non-White civilians using this subsample of counties in Appendix Figure B11 and find similar results.

[46] In Figure B26 we present the event-study estimates from our main model of police killings, after controlling for county-year arrests, and find similar results to our baseline estimates.

[47] In Appendix Table C8, we find suggestive evidence that the treatment is not associated with differences in the demand for police services. We also find no evidence that treatment is associated with differences in access to emergency systems.

results. However, it is important to note that we do not have information on all police contacts, and there may be other plausible mechanisms driving our findings.

The suggestive evidence in Figure 8 and Figure B33 indicates a decreases in arrests and police killings of non-Whites in areas with lower representation gaps, meaning that affirmative action litigation itself may have affected our findings independent of diversity. Moreover, we do not find a statistically significant decrease in quality of life arrests in panel (d) of Figure 10, implying that we cannot use this result as supporting the notion of a decrease in racially discriminatory policing as has been found in Chalfin and McCrary (2018). Instead, not finding a definitive effect of litigation on quality of life arrests may be due to less effort exerted by police post-treatment, which would coincide with previous studies such as Fryer (2018). They find that investigations of police departments that are preceded by highly publicized incidents of deadly force lead to lower quality policing, as evidenced by fewer police-civilian interactions and an increase in homicides. However, we do not find an effect on homicides (see Figure B28), but we do find a decrease in non-White arrests for property crimes. Also, recall that in Figure 8 we find suggestive evidence that the treatment had the greatest impact on police killing of non-Whites in locations with a lower Black share of officers. Although we find suggestive evidence that affirmative action litigation itself may have led to the decrease in non-White deaths due to legal intervention, we are unable to precisely tease out whether officers are putting forth less effort or whether there is a cultural shift leading to fewer discriminatory interactions as a result of diversity or otherwise.

6 Conclusion

Our results indicate a long-run decrease in police killings, particularly of non-White civilians, as a result of threatened affirmative action litigation. There is also suggestive evidence that police in those departments killed fewer White individuals as well. We find that a reduction in police contact with non-White civilians, measured by low-level arrests, may drive our findings. As in other studies, we confirm that a potential mechanism contributing to our main findings may be increased minority representation in police departments (Harvey and Mattia, 2022; Miller and Segal, 2014; McCrary, 2007). We also find suggestive evidence that litigation in and of itself may contribute to the reduction of police killings of civilians.

Given the recent call to diversify and restructure police departments, our results highlight

the vital role that federal interventions have in addressing the excessive use of police force in marginalized communities. Overall, our results show that diversifying the racial composition of police departments, or at the very least affirmative action litigation intended to do so, decreases non-White deaths at the hands of police. However, even with greater diversity within police departments today, Black citizens are still 2.5 times more likely to be killed by the police than White men (Buehler, 2017; Edwards et al., 2018). During encounters with Black citizens, police are also more likely to draw a gun and employ aggressive non-lethal force tactics (Fryer, 2019). If racial disparities stem from structural factors, such as a deep-rooted organizational culture that is in opposition to non-White residents (e.g., racial threat theory (Jacobs and O'brien, 1998)), political power (Gray and Parker, 2020), or other federal programs that encourage greater use of force (e.g., the 1033 program), then simply diversifying police personnel may be insufficient to address apparent racial disparities in police killing. Evidence of a more systemic issue may be present in our findings, given that our results suggest that diversification decreases non-White deaths in the long-run. Moreover, recent work suggests little impact of external factors (e.g., residential segregation) having spillover effects on deadly use of force relative to internal systemic changes to police departments.[48] Thus, it may be necessary to take other steps or policies to reform police culture for more immediate changes in police behavior.

[48]For instance, police union collective bargaining (**?**) have been shown to affect police killing of civilians. Conversely, factors such as extreme temperatures, segregation, and access to behavioral health have much weaker effects (Deza et al., 2023; Cox et al., 2022; **?**).

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1):1–19.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12):4088–4118.

Ba, B. A. and Rivera, R. (2019). The effect of police oversight on crime and allegations of misconduct: Evidence from chicago. U of Penn, Inst for Law & Econ Research Paper, (19-42).

Bailey, M. J. and Goodman-Bacon, A. (2015). The war on poverty's experiment in public medicine: Community health centers and the mortality of older americans. American Economic Review, 105(3):1067–1104.

Barber, C., Azrael, D., Cohen, A., Miller, M., Thymes, D., Wang, D. E., and Hemenway, D. (2016). Homicides by police: comparing counts from the national violent death reporting system, vital statistics, and supplementary homicide reports. American journal of public health, 106(5):922–927.

Bell, M. C. (2016). Police reform and the dismantling of legal estrangement. YAle lJ, 126:2054.

Bjuggren, C., Cox, R., Logan, T., and Williams, J. (2023). The legacy of lynchings and state-sanctioned violence against black americans. Working Paper.

Boggess, S. and Bound, J. (1997). Did criminal activity increase during the 1980s? comparisons across data sources. Social Science Quarterly, pages 725–739.

Buehler, J. W. (2017). Racial/ethnic disparities in the use of lethal force by us police, 2010–2014. American journal of public health, 107(2):295–297.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. Journal of Econometrics.

Campbell, T. (2021). Black lives matter's effect on police lethal use-of-force. Available at SSRN.

Carrell, S. E., Hoekstra, M., and West, J. E. (2019). The impact of college diversity on behavior toward minorities. American Economic Journal: Economic Policy, 11(4):159–82.

Carter, G. L. (1986). The 1960s black riots revisited: city level explanations of their severity. Sociological Inquiry, 56(2):210–228.

Chalfin, A. and McCrary, J. (2018). Are us cities underpoliced? theory and evidence. Review of Economics and Statistics, 100(1):167–186.

Close, B. R. and Mason, P. L. (2007). Searching for efficient enforcement: Officer characteristics and racially biased policing. Review of Law & Economics, 3(2):263–321.

Collaborators, G. . P. V. U. S. et al. (2021). Fatal police violence by race and state in the usa, 1980–2019: a network meta-regression. The Lancet, 398(10307):1239–1255.

Collins, W. J. and Margo, R. A. (2007). The economic aftermath of the 1960s riots in american cities: Evidence from property values. The Journal of Economic History, 67(4):849–883.

Commission, U. S. K. and on Civil Disorders, U. S. N. A. C. (1968). Report of the national advisory commission on civil disorders. US Government Printing Office.

Correll, J., Park, B., Judd, C. M., and Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. Journal of personality and social psychology, 83(6):1314.

Correll, J., Wittenbrink, B., Park, B., M., J. C., and Goyle, A. (2011). Dangerous enough: Moderating racial bias with contextual threat cues. Journal of experimental social psychology, 47(1):184–189.

Cox, R. and Cunningham, J. P. (2021). Financing the war on drugs: The impact of law enforcement grants on racial disparities in drug arrests. Journal of Policy Analysis and Management, 40(1):191–224.

Cox, R., Cunningham, J. P., Ortega, A., and Whaley, K. (2022). Black lives: The high cost of segregation. University of Southern California Working Paper.

Cunningham, J. P., Feir, D., and Gillezeau, R. (2020). The impact of access to collective bargaining rights on policing and civilian deaths. Working Paper.

Cunningham, J. P. and Gillezeau, R. (2019). Don't shoot! the impact of historical african american protest on police killings of civilians. Journal of Quantitative Criminology, pages 1–34.

Devi, T. and Fryer Jr, R. G. (2020). Policing the police: The impact of" pattern-or-practice" investigations on crime. Technical report, National Bureau of Economic Research.

Deza, M., Lu, T., Maclean, J. C., and Ortega, A. (2023). Behavioral health treatment and police officer safety. Technical report, National Bureau of Economic Research.

DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. Econometrica, 64(5):1001–1044.

Donohue III, J. J. and Levitt, S. D. (2001). The impact of race on policing and arrests. The Journal of Law and Economics, 44(2):367–394.

Edwards, F., Esposito, M. H., and Lee, H. (2018). Risk of police-involved death by race/ethnicity and place, united states, 2012–2018. American journal of public health, 108(9):1241–1248.

Edwards, F., Lee, H., and Esposito, M. (2019). Risk of being killed by police use of force in the united states by age, race–ethnicity, and sex. Proceedings of the National Academy of Sciences, 116(34):16793–16798.

Farhang, S. (2010). The litigation state: public regulation and private lawsuits in the US, volume 113. Princeton University Press.

Feigenberg, B. and Miller, C. (2018). Racial divisions and criminal justice: Evidence from southern state courts. Technical report, National Bureau of Economic Research.

Feldman, J. M., Gruskin, S., Coull, B. A., and 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. (2018). Reconciling results on racial differences in police shootings. AEA Papers and Proceedings, 108:228–33.

Fryer, R. G. (2019). An empirical analysis of racial differences in police use of force. Journal of Political Economy, 127(3):1210–1261.

Fyfe, J. J. (2002). Too many missing cases: Holes in our knowledge about police use of force. Justice Research and Policy, 4(1-2):87–102.

Garner, M., Harvey, A., and Johnson, H. (2019). Estimating effects of police force diversity: A replication and extension of previous research. Working Paper.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2):254–277.

Gray, A. C. and Parker, K. F. (2020). Race and police killings: examining the links between racial threat and police shootings of black americans. Journal of Ethnicity in Criminal Justice, pages 1–26.

Haines, M. R. et al. (2010). Historical, demographic, economic, and social data: the united states, 1790–2002. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Harvey, A. and Mattia, T. (2022). Reducing racial disparities in crime victimization: Evidence from employment discrimination litigation. Journal of Urban Economics, page 103459.

Hidalgo, C. (2019). Police Officer Data.

Hoekstra, M. and Sloan, C. W. (2020). Does race matter for police use of force? evidence from 911 calls. Working Paper.

Jacobs, D. and O'brien, R. M. (1998). The determinants of deadly force: A structural analysis of police violence. American journal of sociology, 103(4):837–862.

Loftin, C., Wiersema, B., McDowall, D., and Dobrin, A. (2003). Underreporting of justifiable homicides committed by police officers in the united states, 1976–1998. American Journal of Public Health, 93(7):1117–1121.

Masera, F. (2021). Police safety, killings by the police, and the militarization of us law enforcement. Journal of Urban Economics, 124:103365.

McCrary, J. (2007). The effect of court-ordered hiring quotas on the composition and quality of police. American Economic Review, 97(1):318–353.

McElvain, J. P. and Kposowa, A. J. (2008). Police officer characteristics and the likelihood of using deadly force. Criminal justice and behavior, 35(4):505–521.

Miller, A. R. and Segal, C. (2012). Does temporary affirmative action produce persistent effects? a study of black and female employment in law enforcement. Review of Economics and Statistics, 94(4):1107–1125.

Miller, A. R. and Segal, C. (2014). Do female officers improve law enforcement quality? effects on crime reporting and domestic violence escalation. University of Zurich, UBS International Center of Economics in Society, Working Paper, (9).

Miller, A. R. and Segal, C. (2019). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence. Review of Economic Studies, 86(5):2220–2247.

Myers, S. L. (1980). Why are crimes underreported? what is the crime rate? does it" really" matter? Social Science Quarterly, 61(1):23–43.

on 21st Century Policing, P. T. F. (2015). Final report of the president's task force on 21st century policing.

on Criminal Justice Standards, N. A. C. and Goals (1973). A national strategy to reduce crime. National Advisory Commission on Criminal Justice Standards and Goals.

Rabe-Hemp, C. E. (2008). Female officers and the ethic of care: Does officer gender impact police behaviors? Journal of criminal justice, 36(5):426–434.

Shannon, J. (2020). Usa today poll: Americans want major police reform, more focus on serious crime. USA Today.

Sherman, L. W. and Langworthy, R. H. (1979). Measuring homicide by police officers. J. Crim. L. & Criminology, 70:546.

Smith, B. W. (2003). The impact of police officer diversity on police-caused homicides. Policy Studies Journal, 31(2):147–162.

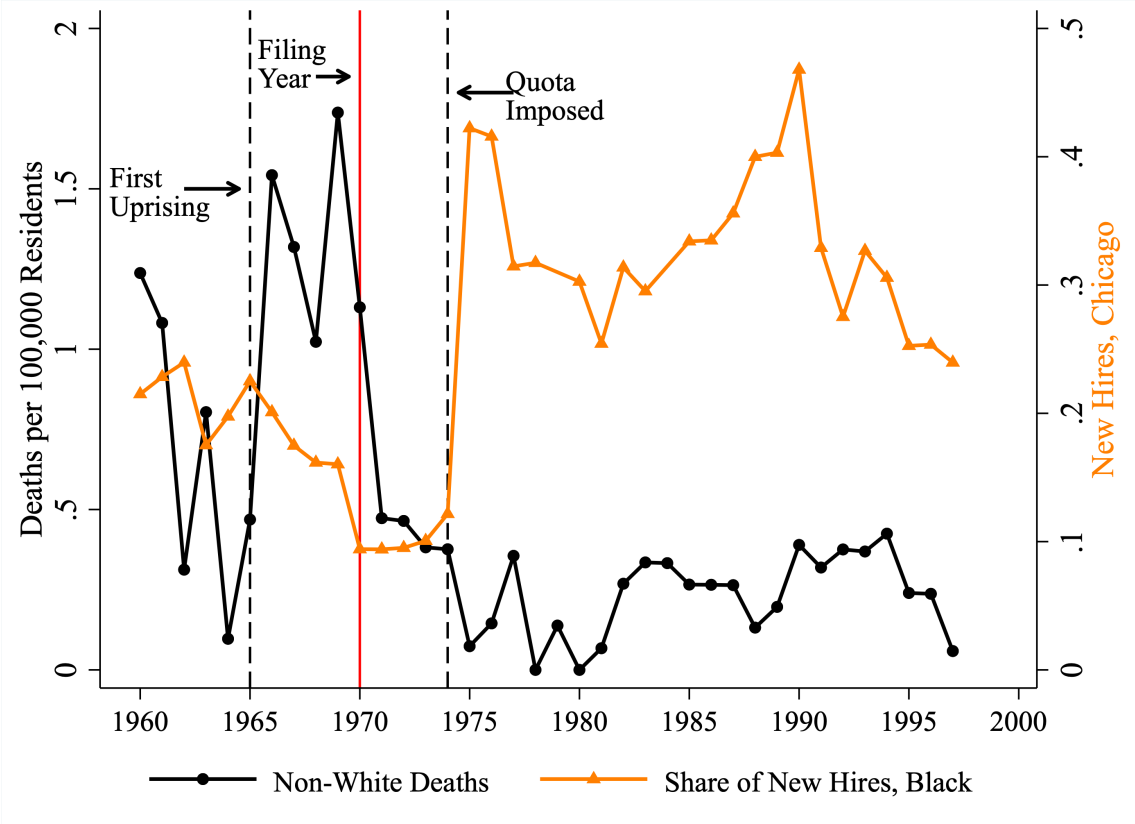
Sun, L. and Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics.

US Department of Health and Human Services (2007). Multiple cause of death public use files, 2000-2002. https://www.icpsr.umich.edu/web/NACDA/studies/4640.

Weisburst, E. K. (2019). Police use of force as an extension of arrests: Examining disparities across civilian and officer race. In AEA Papers and Proceedings, volume 109, pages 152–56.

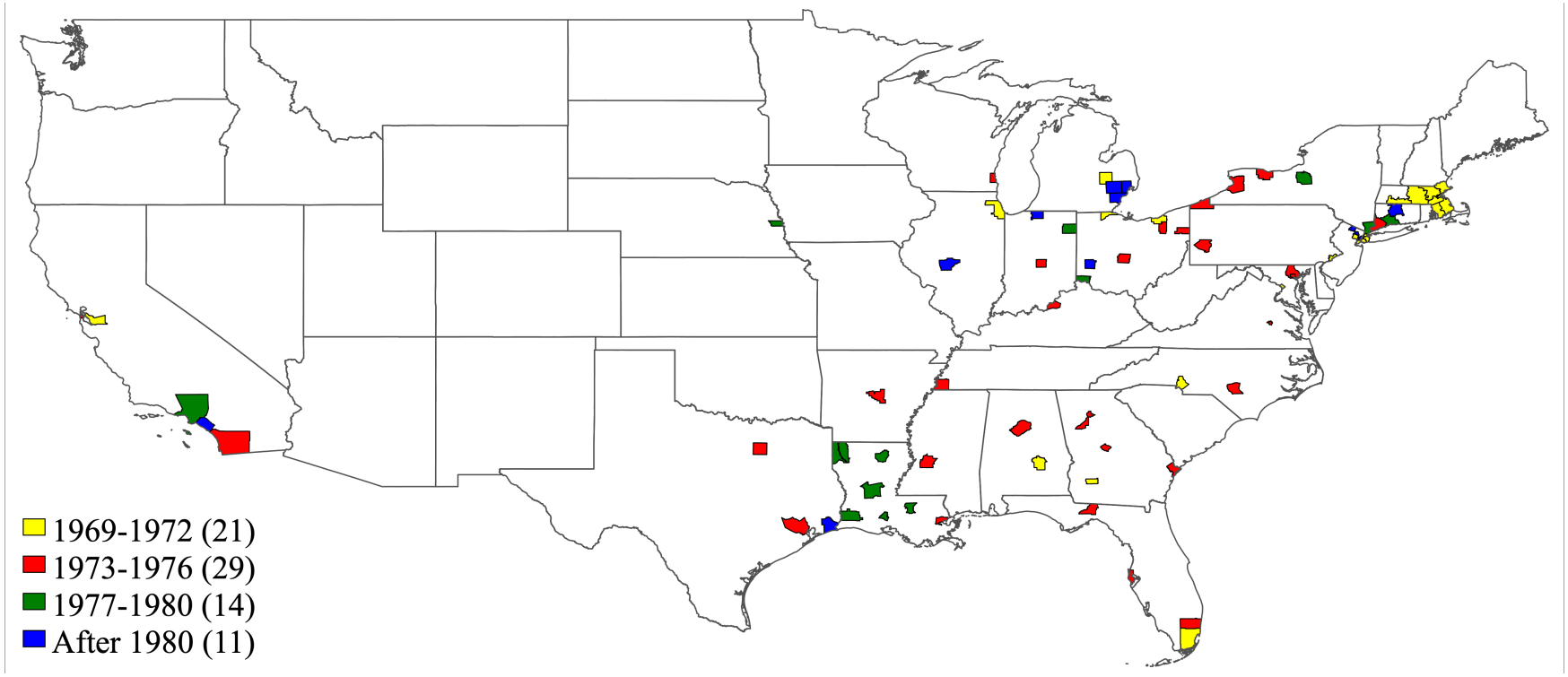
Figures

Figure 1: Cook County Police Killings Over Time



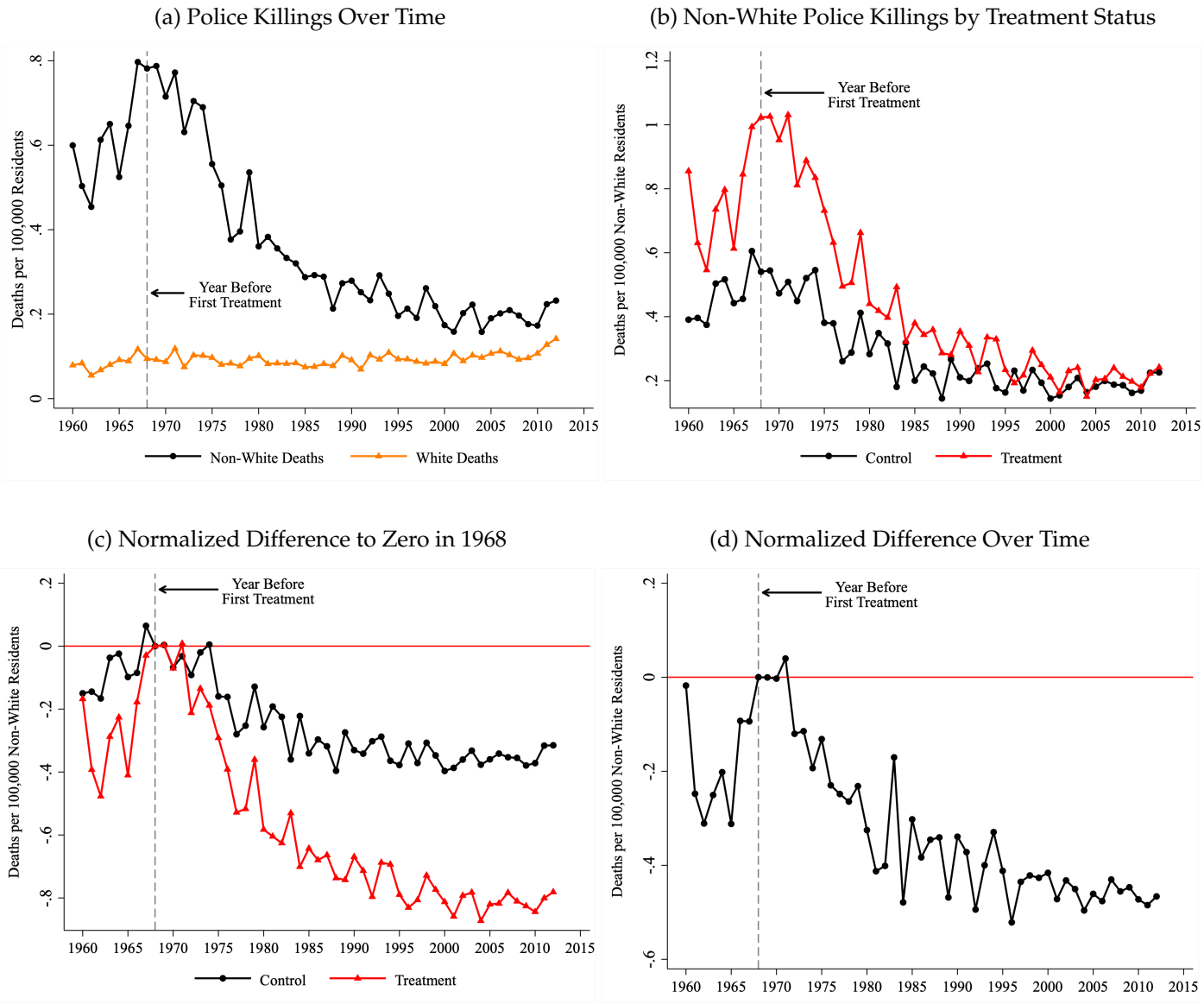
Notes: This Figure plots the data of Chicago Police Department on new hires come from McCrary (2007) and Cook county's police killings of non-White civilians over time. First uprising refers to year of first protests in response to police killings of African American citizens. Year filing is the initial year of Affirmative Action litigation and quota imposed refers to the court-ordered quota the Chicago police department was subject to after the lawsuit.

Figure 2: Location of Legal Action, 1969-2000



Notes: Legal action dates and locations come from McCrary (2007).

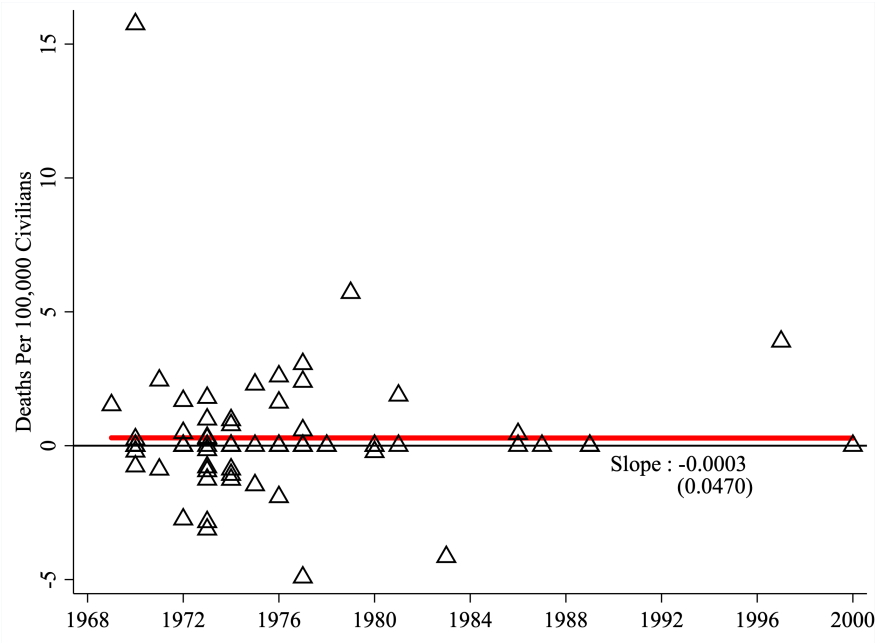
Figure 3: Police Killings Over Time – by Race & Treatment Status



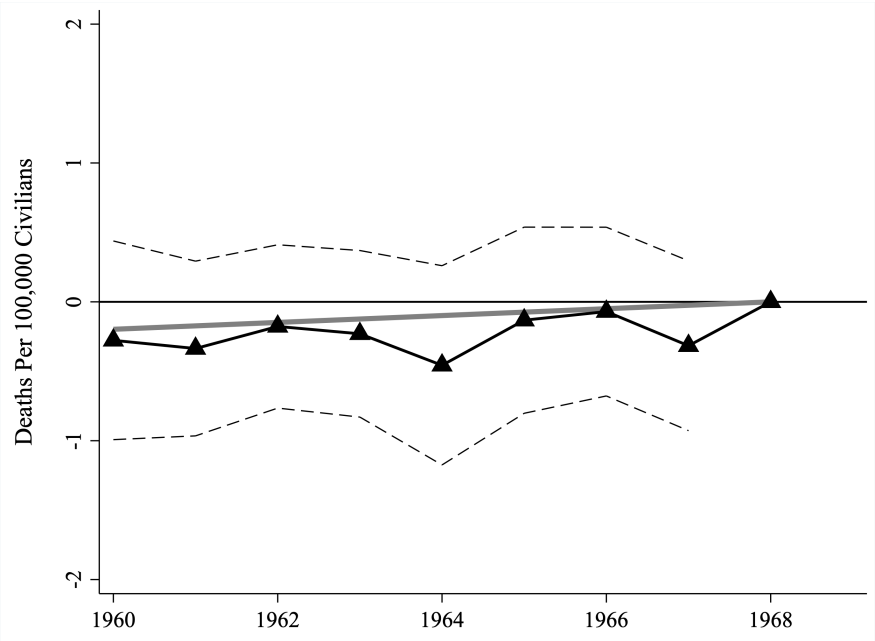
Notes: Panel (a) plots policing-related civilian fatalities per capita over time by race using Vital Statistics Data. Panel (b) plots policing-related fatalities of Non-White civilians over time by treatment status. Panel (c) normalized the difference in policing-related fatalities between control and treatment groups to zero in 1968. Panel (d) plots the difference in policing-related fatalities between treated and control groups. All panels use the full sample of counties and mortality rates are unweighted and not adjusted by age.

Figure 4: Test for Changes in Pre-Period Growth Rates by Treatment Timing and Status

(a) Pre-Trend Growth Rates in Non-White Deaths (1960-1968)

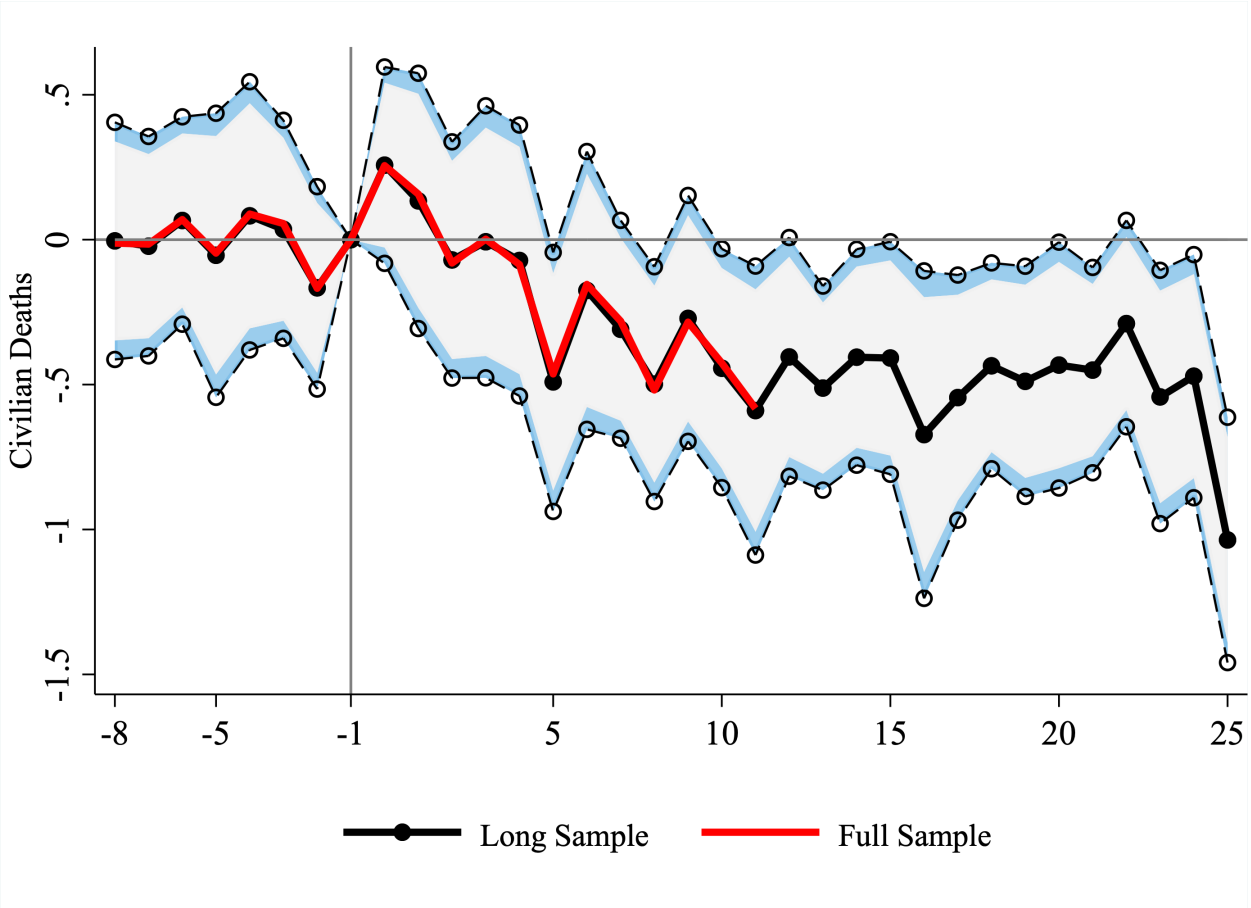


(b) Pre-Period Differences in Non-White Deaths



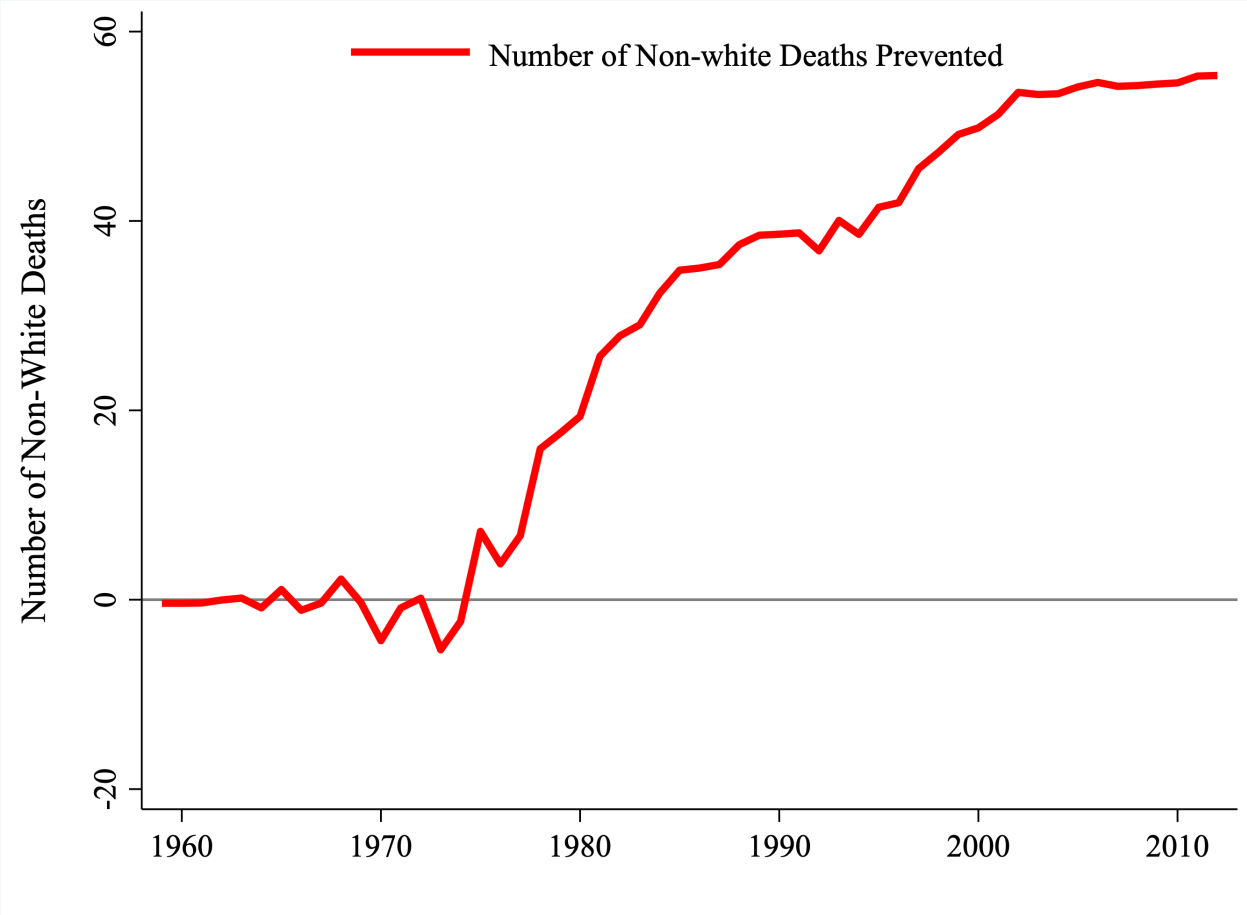
Notes: Panel (a) Regression coefficients and predicted values are from a univariate regression of the dependent variable changes in the non-White death rate on the year a county is threatened with litigation. Panel (b) The dependent variable is non-White deaths per capita. The independent variables are year fixed effects (1960-1968) – Y, treatment indicator (0/1 if ever litigated) –T, and year by treatment effects $T \times Y$. The coefficients plotted are the coefficients on the interaction terms.

Figure 5: Event Study – Non-White Deaths Due to Legal Intervention



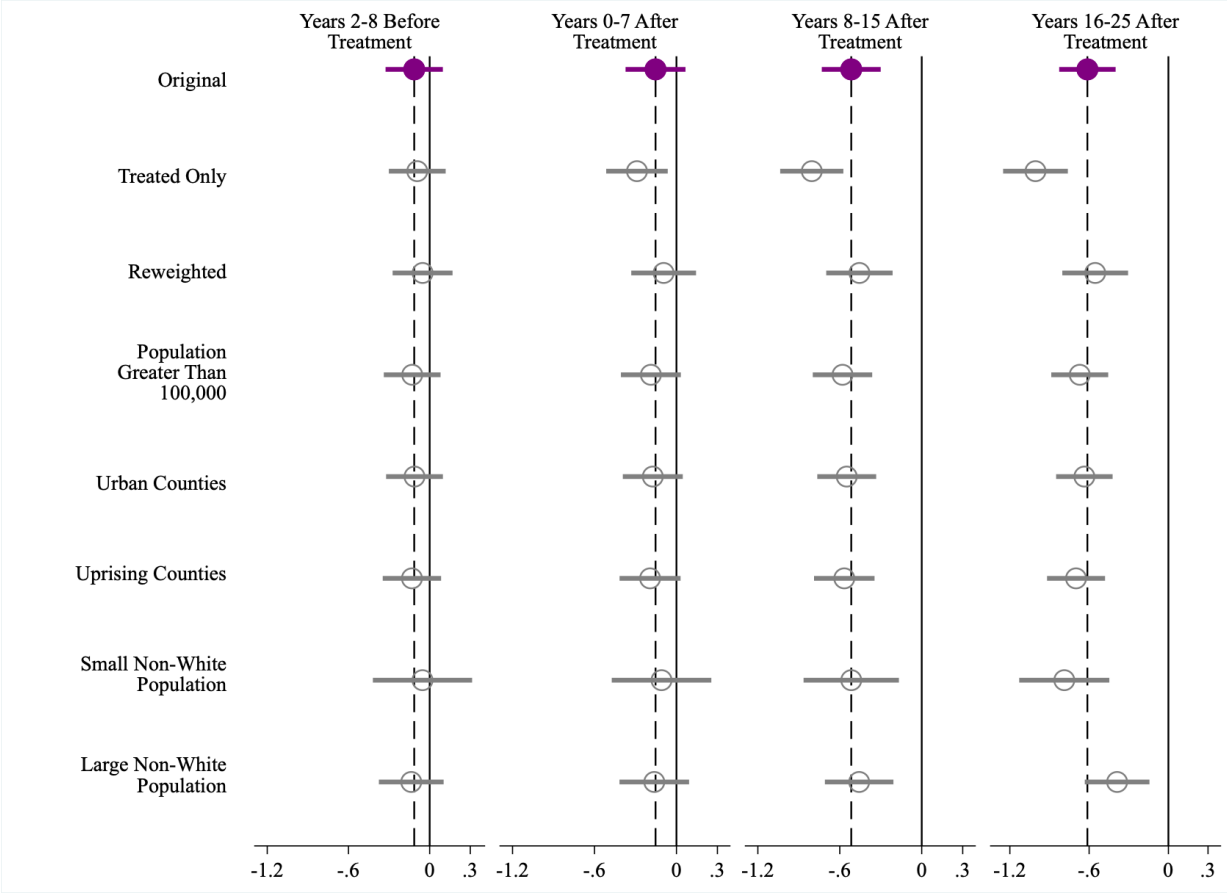
Notes: This figure plots Poisson regression estimates of the effect of the threat of court-ordered affirmative action on non-White policing-related civilian fatalities. This specification includes county and region-by-year fixed effects and accounts for exposure with non-White population. The red line corresponds to the full sample, counties treated between 1969 and 2000. The Black line with circle markers correspond to the long-sample, counties treated between 1969 and 1987. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the long-sample only. The horizontal axis represents event-years (years before and after litigation).

Figure 6: Number of Deaths Prevented



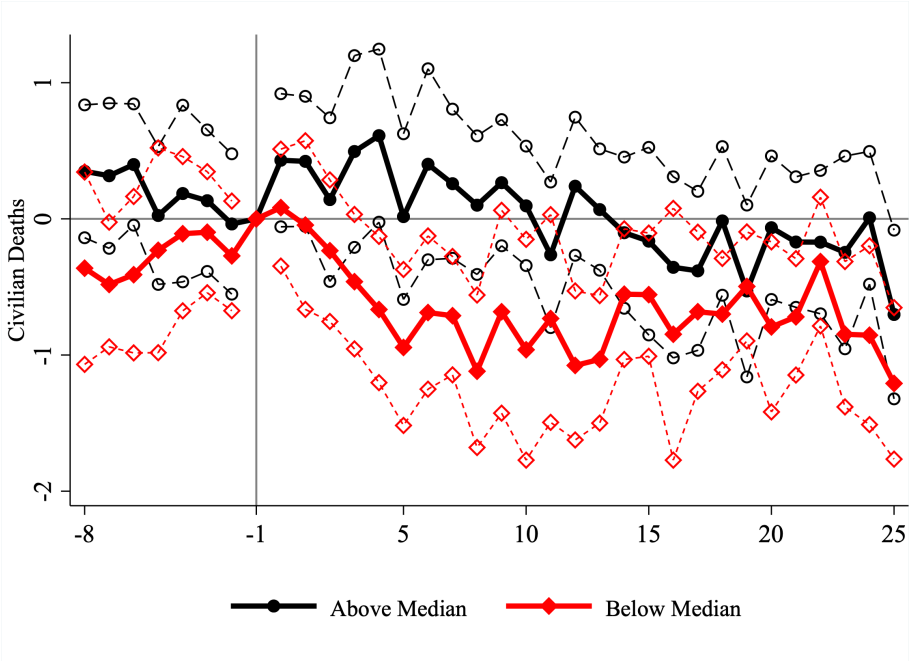
Notes: This figure plots the number of non-White policing-related civilian deaths averted by using coefficients from 5. The number of deaths averted is calculated by first transforming the pre-and-post-treatment effects from the Poisson specification, and then maps event years to actual calendar years.

Figure 7: Robustness Checks - Non-White Deaths



Notes: This figure displays Poisson estimates obtained from estimating Equation (1) after grouping event-years. The dependent variable is non-White policing-related civilian fatalities. All rows include county and region-by-year fixed effects and accounts for exposure with non-White population. Heteroskedasticity-robust standard errors clustered by county are presented by the bold line. Joint lest-square coefficients are presented by circle markers. Row 1 presents the baseline joint treatment effects; Row 2 restricts the sample solely to treated counties; Row 3 employs a semi-parametric reweighing design; Row 4 restricts the sample to counties with a population greater than 100,000; Row 5 limit the sample to counties where the proportion of residents residing in urban areas was higher than the median share of urbanization in 1960; Row 6 restricts the sample to counties that experienced at least one racial uprising in the 1960s; Rows 7 restricts the treatment group into counties where the share of the non-White population in 1960 is below the median share for the treatment group; Row 8 restricts the treatment group into counties where the share of the non-White population in 1960 is above the median share for the treatment group.

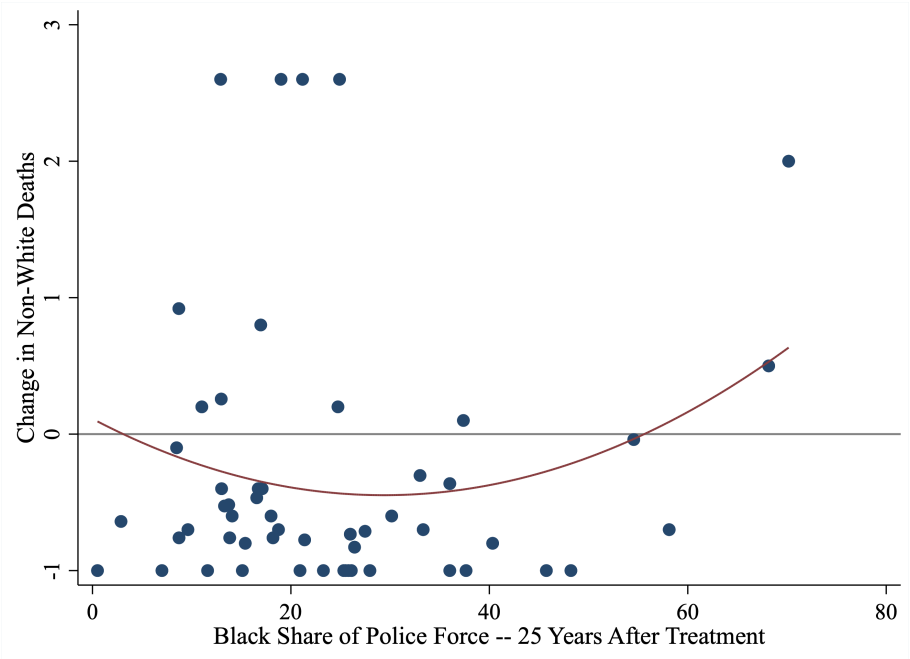
Figure 8: Event-Study Results - Police Representation



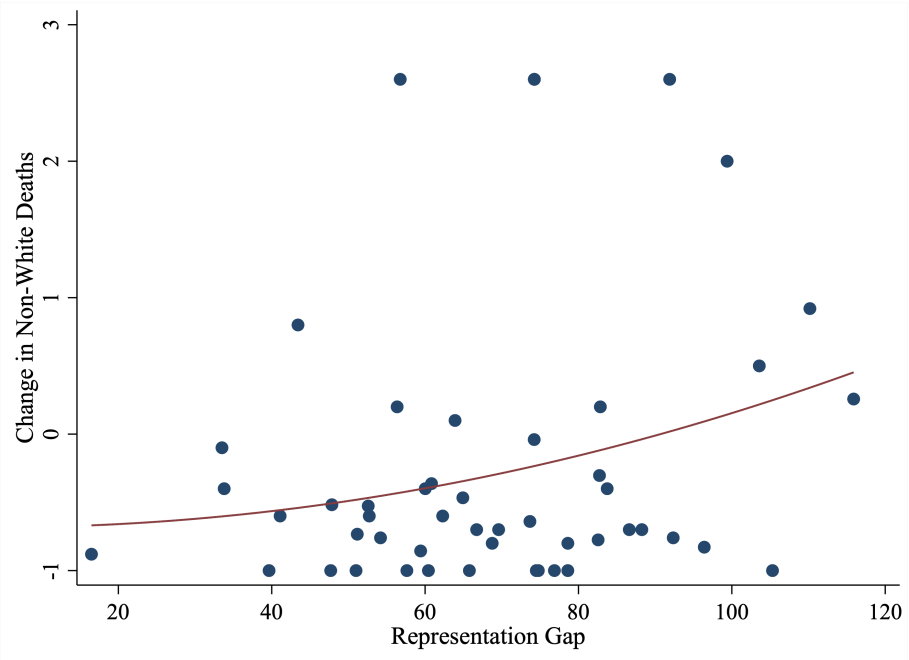
Notes: This figure stratifies the treated groups to those with litigated police departments with above (Black line with circle markers) and below (red line with square markers) median level of representation, 25 years after treatment. Representation is measured as the share of the police of officers that are Black divided by the share of the city population that is Black. The dependent variable is non-White policing-related civilian fatalities. This specification includes county and region-by-year fixed effects. The horizontal axis represents event-years (years before and after litigation).

Figure 9: Change in Actual Police Killings by Representation Gap

(a) Black Share

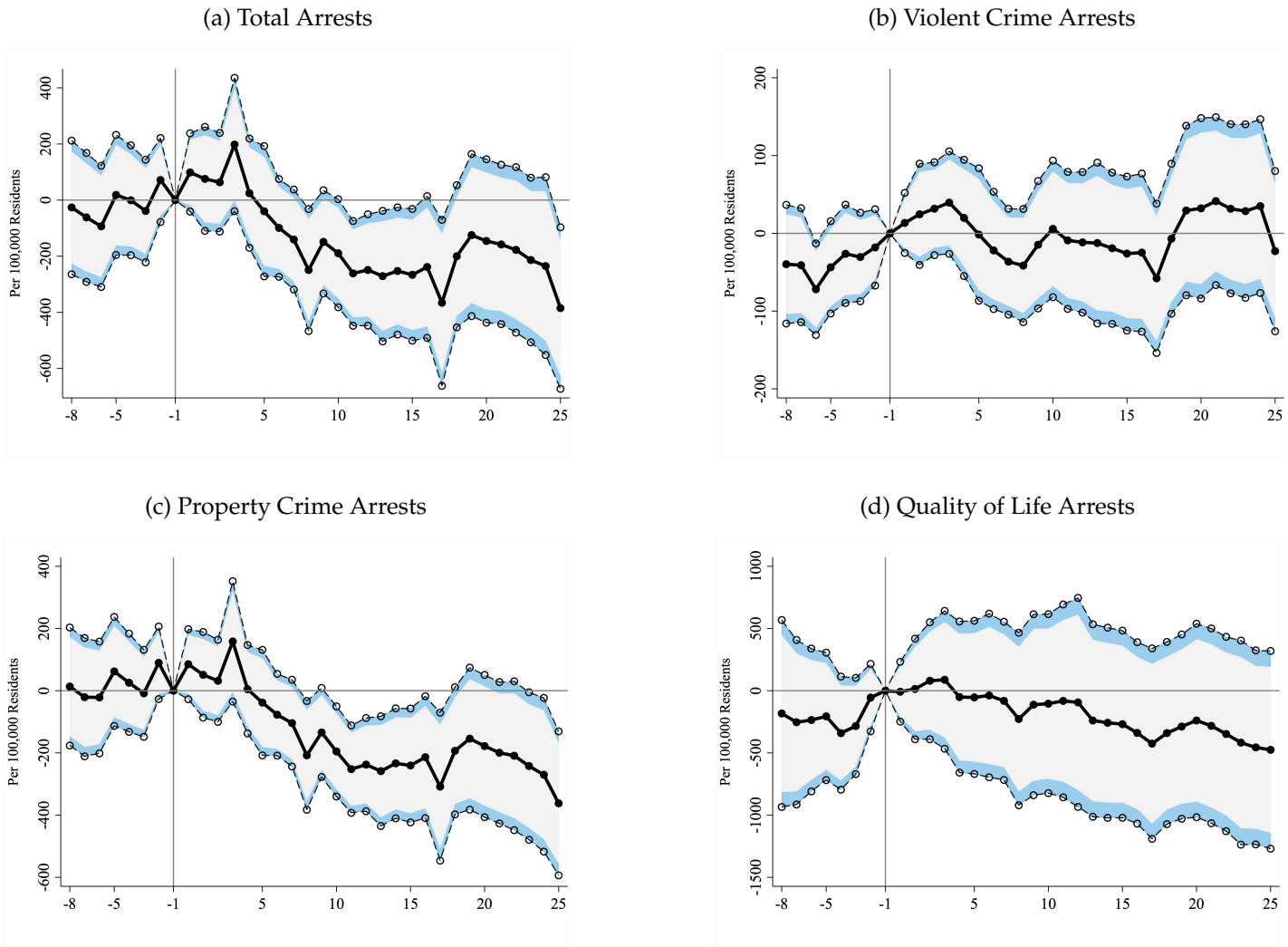


(b) Representation Gap



Notes: The figure shows the relationship between the change in non-White policing-related civilian fatalities relative to the Black share of officers (panel (a)) and the representation gap (panel (b)), 25 years after treatment. The change in police killings is measured by $\widehat{Killings}_{Long-Run} - \widehat{Killings}_{Pre-Period}$ where the long-run refers to average number of deaths in event years 20 to 25 and the pre-period refers to the average number of deaths in event years -8 to -2. Estimates of the share of Black officers and the representation gap, 25 years after treatment, are calculated using data from data from LEMAS. Each specification includes county and region-by-year fixed effects.

Figure 10: Event Study Estimates – UCR Non-White Arrests



Notes: This figure plots the regression estimates of the effect of the threat of court-ordered affirmative action on non-White arrests. Each specification includes county and region-by-year fixed effects. Panel (a) report total non-White arrests for index 1 crimes (violent and property crimes) while panels (b) and (c) report non-White arrests by crime type (violent and property). Panel (d) report quality of life arrest (public intoxication, liquor violations, disorderly conduct, gambling, suspicious behavior, vandalism, and vagrancy). The sample is limited to counties identified by McCrary (2007) and report at least 52 years between 1960 and 2016. Treatment group corresponds to counties with cities treated prior to 1987. Robust standard errors by county. The horizontal axis represents event-years (years before and after litigation).

Tables

Table 1: **Variation in Litigation Over Time**

Treatment Status	Number of Counties	Percent of Counties	Percent of 1960 Population	Percent of 1960 Non-White Population
Treated	75	2.51	32.27	46.13
Year Treated:				
1969	1	0.03	0.53	1.05
1970	11	0.40	7.95	9.01
1971	2	0.47	8.66	10.39
1972	7	0.70	11.79	19.68
1973	11	1.07	16.17	26.56
1974	9	1.37	18.58	30.22
1975	3	1.47	20.17	31.35
1976	6	1.68	21.93	34.43
1977	8	1.94	26.03	40.07
1978	2	2.01	26.66	40.43
1979	1	2.04	26.86	40.57
1980	3	2.14	27.98	41.52
1981	2	2.21	28.52	41.85
1983	1	2.24	28.93	42.14
1986	2	2.31	30.73	44.36
1987	1	2.35	30.87	44.56
1989	3	2.45	31.78	45.34
1997	1	2.48	32.19	46.09
2000	1	2.51	32.27	46.13
Untreated	**2910**	**97.49**	**67.73**	**53.87**

Note: This table report the roll out of counties that experience litigation and the changes in population over this time. Data on threats of litigation comes from McCrary (2007).

Table 2: Summary Statistics

1960 Characteristics	(1) Overall	(2) Treatment	(3) Control Group	(4) T-Test of Difference
Population	59,431	815,406	39,947	<0.01
Population per square mile	165.74	3,214.55	87.16	<0.01
% of counties that experienced uprisings	0.09	0.84	0.07	<0.01
Percentage of the Population				
residing in urban areas	32.50	87.21	31.09	<0.01
w/ 12 or more years of education	36.45	43.41	36.27	<0.01
w/ income greater than 10K	7.92	16.95	7.69	<0.01
w/ income less than 3K	35.62	17.95	36.07	<0.01
non-White	10.94	16.61	10.80	<0.01
Deaths Due to Legal Intervention				
White	0.04	0.61	0.03	<0.01
non-White	0.04	1.03	0.01	<0.01
Number of Counties	2,985	75	2,910	
joint F-test				3.58
p-value				<0.01

Note: This table reports summary statistics. Treatment refers to counties with a police department that experiences Affirmative Action litigation and control refers to all other counties.

Table 3: Event Study - Joint Effects

	(1) OLS	(2) Poisson	(3) WLS	(4) CS Estimator
Pre-Period Effect (Event Years -8 to -2)	-0.0691	-0.00501	-0.0262	-0.068
	[0.227]	[0.149]	[0.0880]	[0.253]
Shorter-Run Effect (Event Years 0 to 7)	-0.166	-0.0502	-0.0775	-0.166
	[0.262]	[0.149]	[0.108]	[0.290]
Medium-Run Effect (Event Years 8 to 15)	-0.596*	-0.413***	-0.249**	-0.595*
	[0.321]	[0.149]	[0.110]	[0.340]
Longer-Run Effect (Event Years 16 to 25)	-0.675**	-0.508***	-0.273**	-0.672**
	[0.319]	[0.148]	[0.113]	[0.333]
Mean DV	1.457	1.457	0.738	1.457
Number of Counties	2,980	2,980	2,707	2,980

Note: All regressions include county and region-by-year fixed effects. Columns (1) and (2) use then number of deaths as the dependent variable while column (3) uses the age-adjusted mortality rate. Column (1) includes the log of non-White population as an independent variable. Column (2) accounts for exposure with non-White population. Column (3) uses 1960 non-White population as weights. Column (4) reports estimates using Callaway and Sant'Anna (2020) estimator – analogous to column (1) OLS two-way fixed effects model. Robust standard errors are clustered by county. *** p<.01, ** p<.05, * p<0.1

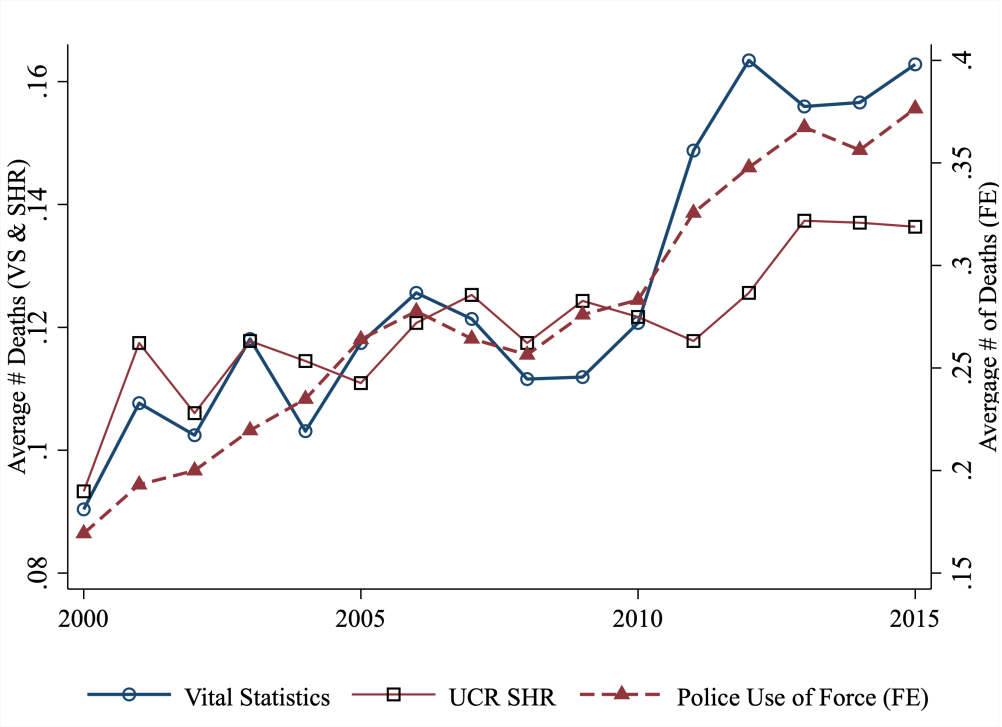
A Data Appendix

We use several data sources in our analysis. Below we describe the data we use and how we obtain the final data set.

A.1 Vital Statistics Data

Our measurement of police killings of civilians comes from the 1959 to 2016 Vital Statistics Multiple Cause of Death Files. The 1958-1988 files are publicly available at the ICPSR at the University of Michigan. For 1989-2016, we use restricted files obtained directly from the CDC. A major shortcoming of using Vital Statistics files to measure police-related fatalities is the severe under-reporting of police-caused deaths. Ideally, we would use newer, crowd-sourced data such as Fatal Encounters to capture police-related fatalities, but the series begins well after locations are treated. Despite the under-count, Vital Statistics does reasonably well with capturing national trends. Below, Appendix Figure A1 shows that the Vital Statistics and Fatal Encounters series on police-related fatalities mirror each other over time. Also, Appendix Figure A2 plot the time-series by treatment status.

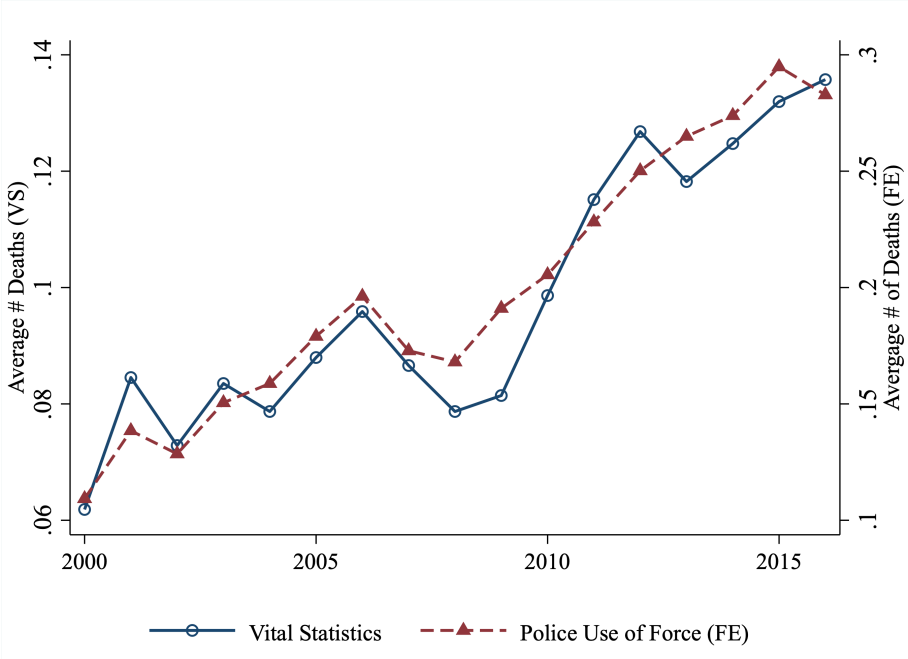
Figure A1: Police Killings, 2000-2016



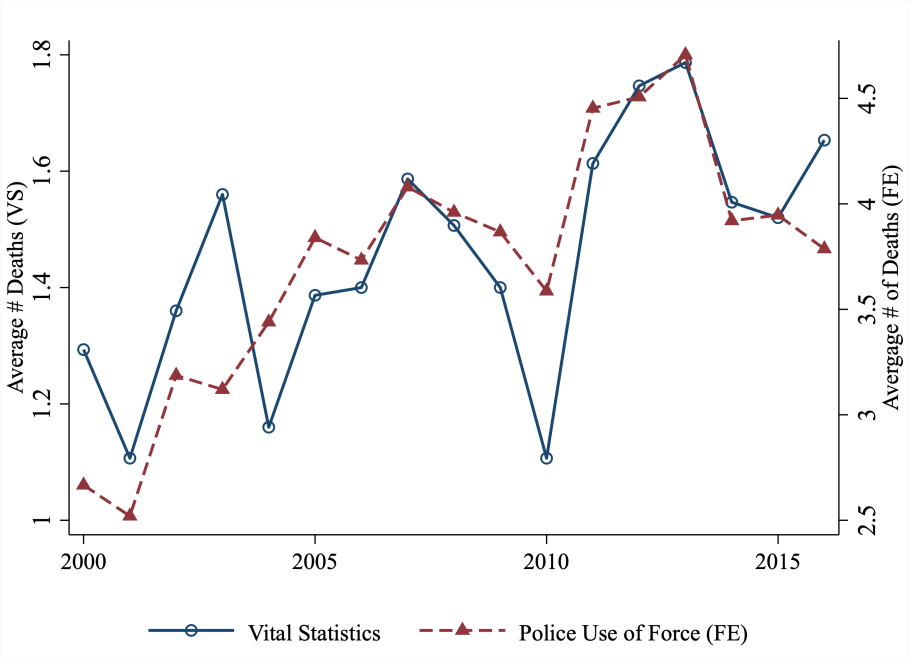
Notes: This figure shows the trend in non-White policing-related civilian fatalities from various data sources. Data are from the Vital Statistics Multiple Cause of Death Files, the UCR: Supplemental Homicide Report, and the Fatal Encounters Data (see Fatal Encounters Website).

Figure A2: Police Killings by Treatment Status, 2000-2016

(a) Control Group



(b) Treatment Group



Notes: This figure shows the trend in non-White policing-related civilian fatalities by treatment status. Treatment group refers counties that experience Affirmative Action litigation and the control group refers to all other counties. Data are from the 1998 to 2016 Vital Statistics Multiple Cause of Death Files and the Fatal Encounters Data (see Fatal Encounters Website).

The Vital Statistics files cover several decades with many changes to coding police-related fatalities. For years 1999 to 2016 we use the following ICD-10 codes to classify police-related deaths: Y35.0-Y35.4, Y35.6, Y35.7, and Y89.0. For 1968 to 1998, we use the following ICD-9 and ICD-8 codes: E970-E977. Finally, for 1959 to 1967, we use the following ICD-7 code: E984. For suicide deaths from 1999 and 2016, we use the following ICD-10 codes: X60-X84, Y870, U030. For suicide deaths from 1968 to 1998 we use the following ICD-9 and ICD-8 code: E950-E959. For 1959-1967, we use the following ICD-7 codes: E963, E970-E979.

A.2 UCR Data

Information for arrests and police employment comes from the Uniform Crime Report Data Series, which is publicly available at the ICPSR at the University of Michigan.

A.2.1 Police Employment

Uniform Crime Reporting (UCR): Law Enforcement Officers Killed or Assaulted UCR Law Enforcement Officers Killed or Assaulted (LEOKA) contains monthly counts of law enforcement officers killed and assaulted as well as annual law enforcement employment. The UCR reports the number of civilian officers and sworn officers as of October 31st of the reporting year. The UCR LEOKA is available publicly at the ICPSR website for years after 1974. For years 1960 to 1974 data are manually entered from hard copies of UCR LEOKA reports.

A.2.2 Arrest Data

UCR Arrests by Age, Sex, and Race is publicly available at the Inter-University Consortium for Political and Social Research (ICPSR). It provides information on the number of arrests reported by local and state law enforcement agencies to the Federal Bureau of Investigation (FBI). The data compiled for the UCR is submitted voluntarily by city, county, and state enforcement agencies. The FBI or state law agencies directly provide survey forms to local law agencies or state collecting programs, which are collected on a monthly basis. The data on arrests includes yearly information on the number of arrests within 43 categories including property and violent crime as well as drug-related offenses; it also includes information on the age, sex, and race of arrestees.

A.3 Police Diversity

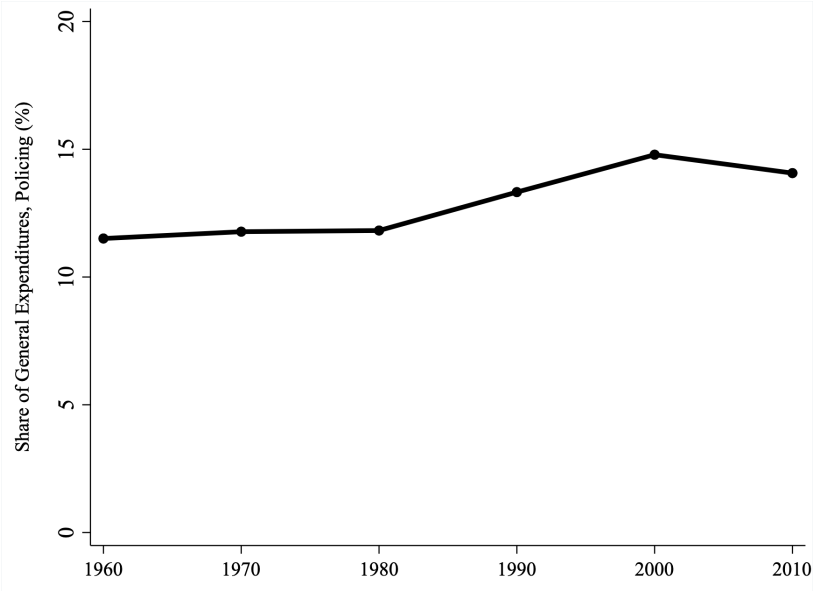
Law Enforcement Management and Administrative Statistics (LEMAS) is publicly available at the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan. LEMAS provides detailed survey data from police agencies on their equipment, personnel, and capabilities. The survey is conducted every three years and includes most large law enforcement agencies. The survey contains information on the demographic composition of law enforcement officers. Our analysis focuses on full-time sworn officers by race.

A.4 Demographic Information

County and City Data Books provide information on city demographics and local government expenditures for counties, cities, and incorporated areas of 25,000 inhabitants or more in the United States. Data on population counts as well as demographic information are from the decennial census conducted in the beginning of each decade. Data on local government operations, revenue, and expenditures are from the Annual Survey of Governments, which is conducted for a sample of governments by the Census Bureau. We use County and City Data Books for 1962 and 1967. Figure A3 plots police spending over time and shows a fairly consistent share of government spending allocated to policing. We obtain yearly population counts from the 1960 Census (Haines

et al., 2010) and the Surveillance, Epidemiology, and End Results (SEER) annual data, which begin in 1968 (interpolated between 1960 and 1968).

Figure A3: Police Spending Over Time, 1960-2010



Notes: This figure reports police spending over time. Data comes from City and County Data Books publicly available at the ICPSR and www.census.gov.

A.5 Uprising Data

Data for rioting was provided by Collins and Margo (2007) and was originally collected by Carter (1986). The data consist of detailed information on arrests, injuries, and deaths caused by riots in each city where rioting occurred. Using this information we create a dummy variable to identify the year and county in which rioting occurred. We also create an index of severity by summing the number of arrests, injuries, arsons, and deaths in the year and dividing them by the total sum of injuries, deaths, arsons, and arrests due to rioting. The severity index is similar to other rioting measures used in research that analyzes the impact of rioting.

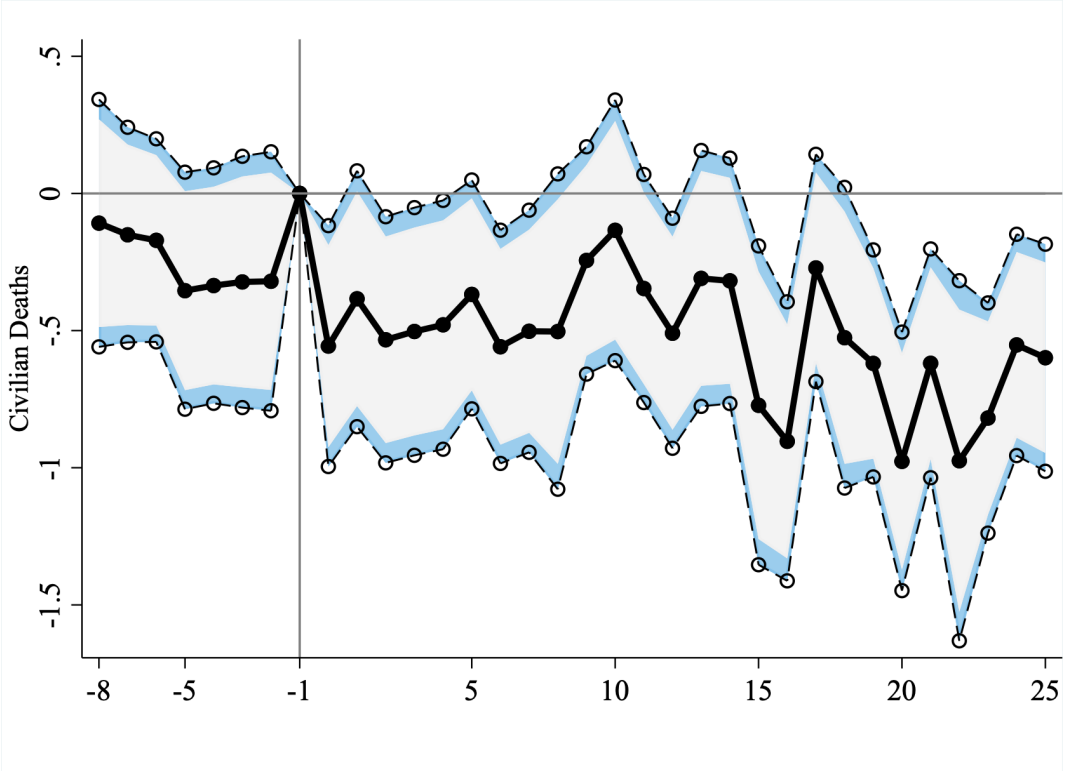
A.6 Final Sample

To compile our final sample we first collapse the litigation data, courtesy of McCrary (2007), and collapse the treated locations to their respective counties. We identify a treated county by the earliest year a city within a county is threatened with affirmative action litigation. We then merge the litigation data with Vital Statistics data using state and county FIPS identifiers. Note, all counties are identified by their 1960 definition. Although San Francisco and Washington D.C. are excluded from McCrary (2007), we include them because we know the litigation dates and outcome of interests. Lastly, we restrict our sample to counties that have a non-White population in every year of record (which excludes 30 counties).

B Appendix Figures

B.1 Alternative Specifications- Police Killings

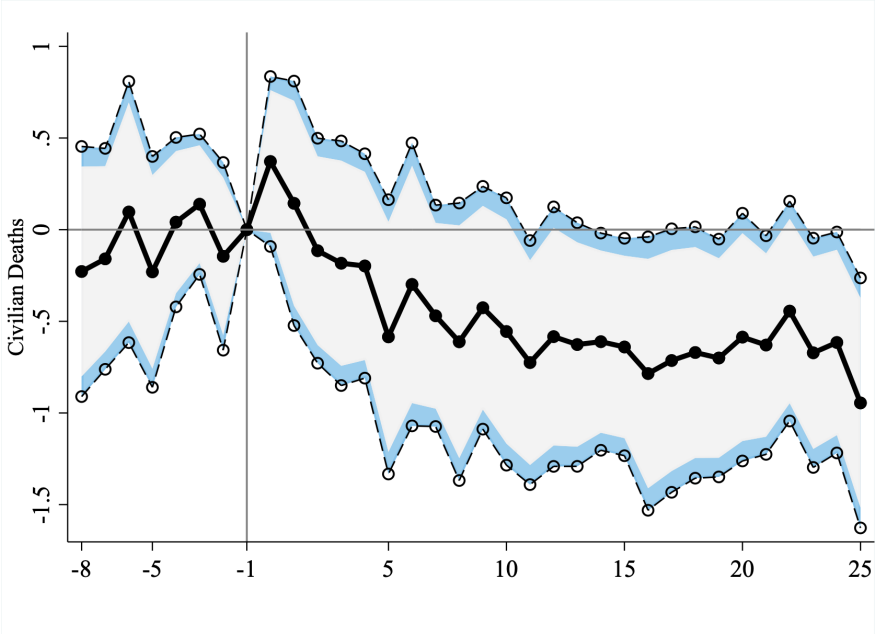
Figure B1: Event Study – White Deaths Due to Legal Intervention



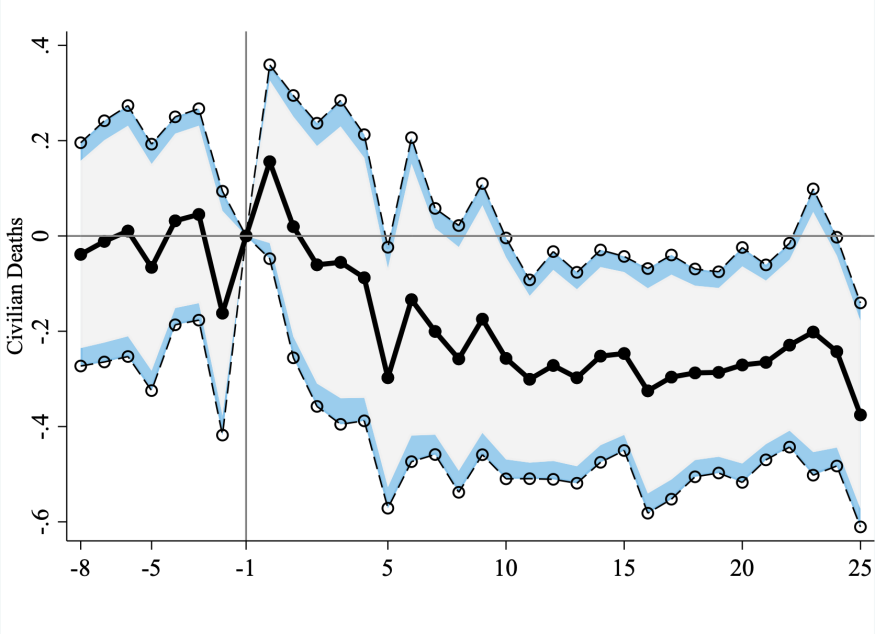
Notes: This figure plots Poisson regression estimates of the effect of the threat of court-ordered affirmative action on White policing-related civilian fatalities. This specification includes county and region-by-year fixed effects, and accounts for exposure with White population. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the long-sample only. The horizontal axis represents event-years (years before and after litigation).

Figure B2: Event Study – Least Squares Regressions – Non-White Deaths

(a) Non-White: OLS



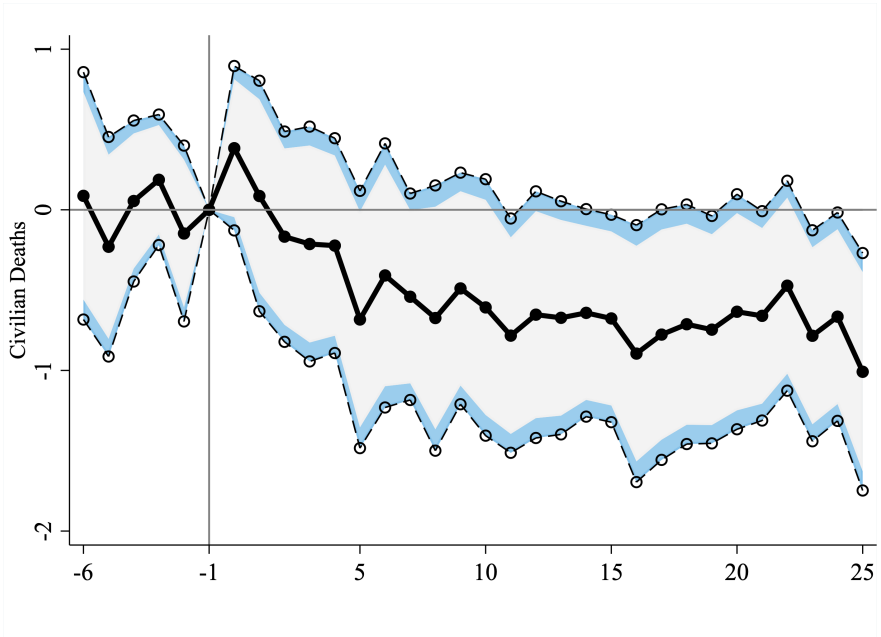
(b) Non-White: WLS



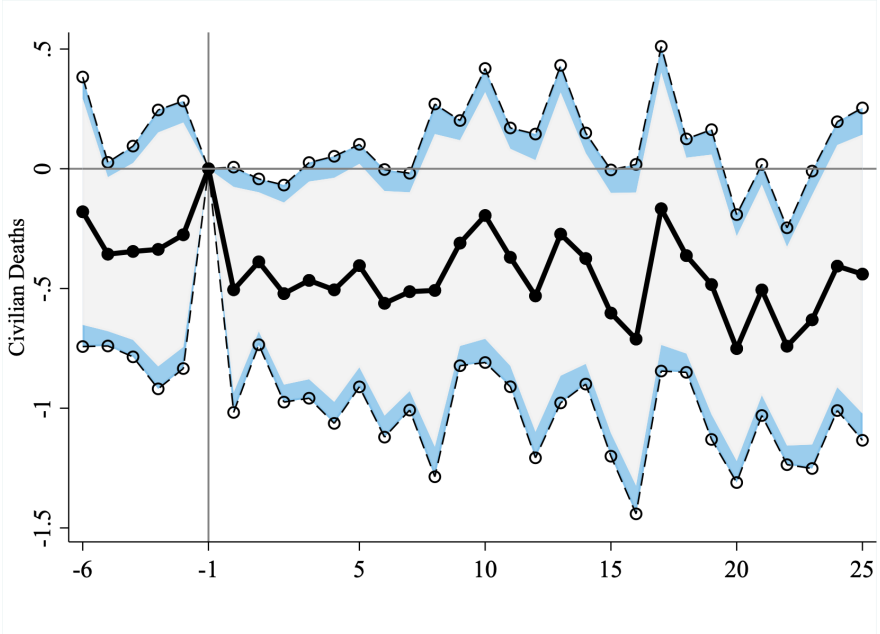
Notes: This figure plots OLS and WLS regression estimates of the effect of the threat of court-ordered affirmative action on non-White policing-related civilian fatalities. Each specification includes county and region-by-year fixed effects. The dependent variable is the number of non-White policing-related civilian fatalities in panel (a) and non-White deaths per 100,000 non-White residents in panel (b). Panel (a) also includes the log of non-White population as a regressor and panel (b) uses the non-White population in 1960 as weights. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the long sample only. The horizontal axis represents event-years (years before and after litigation).

Figure B3: Event Study – Non-White Deaths Due to Legal Intervention: State-by-Year Effects

(a) Non-White



(b) White



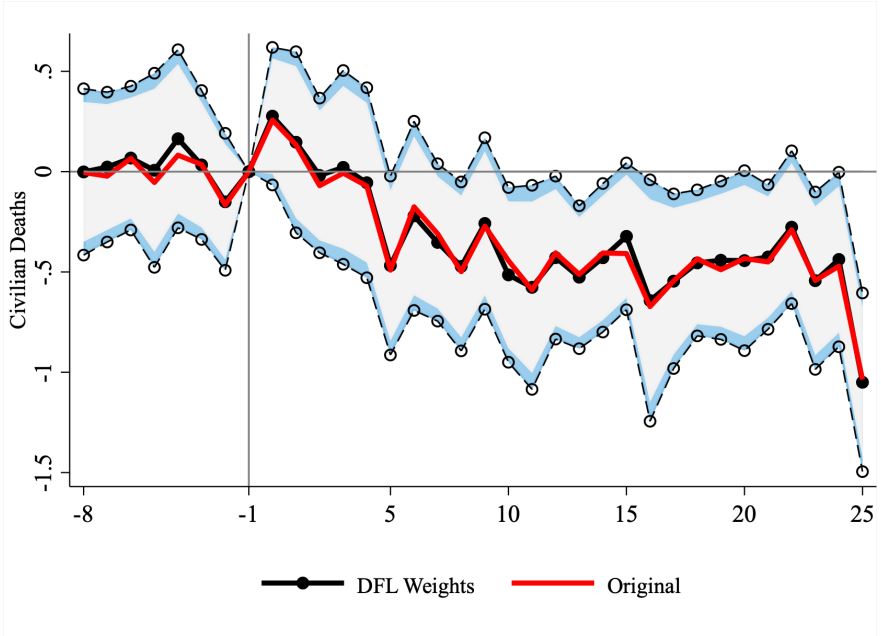
Notes: The ordinary least squares regression specification includes county and state-by-year fixed effects. The dependent variable is the number of non-White policing-related civilian fatalities in panel (a) and White policing-related civilian fatalities in panel (b). Each panel also includes the log of population for each respective racial group as a regressor. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

Figure B4: Event Study – Unadjusted vs. Age-Adjusted Mortality Rates

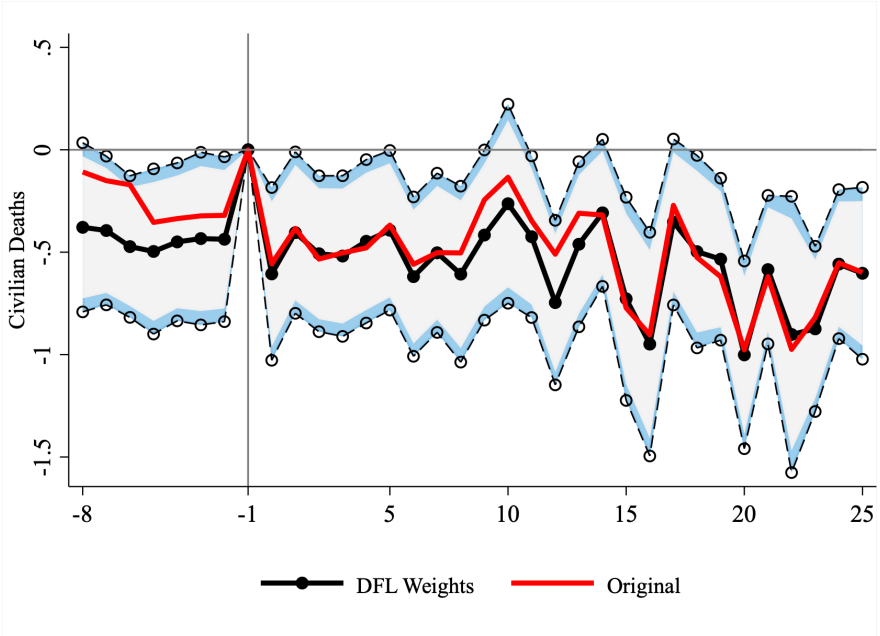
Notes: This figure presents age adjusted and unadjusted non-White policing-related civilian fatalities. Each specification includes county and region-by-year fixed effects. The Black line corresponds to changes in the age-adjusted mortality rate while he red line corresponds to the unadjusted mortality rate. Both regressions use the non-White population in 1960 as weights. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the age-adjusted regressions only. The horizontal axis represents event-years (years before and after litigation).

Figure B5: Event Study – Non-White Deaths Due to Legal Intervention: DFL Weights

(a) Non-White Deaths



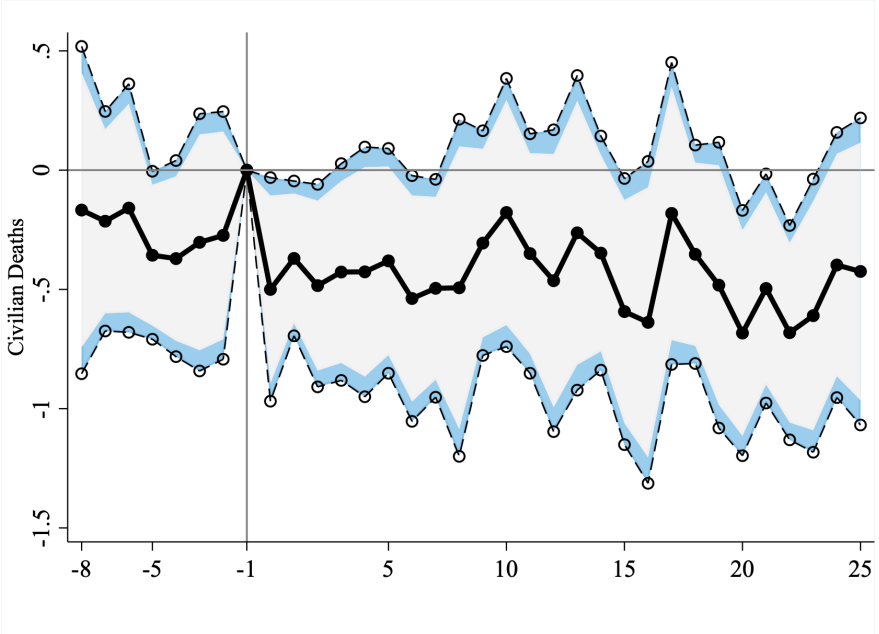
(b) White Deaths



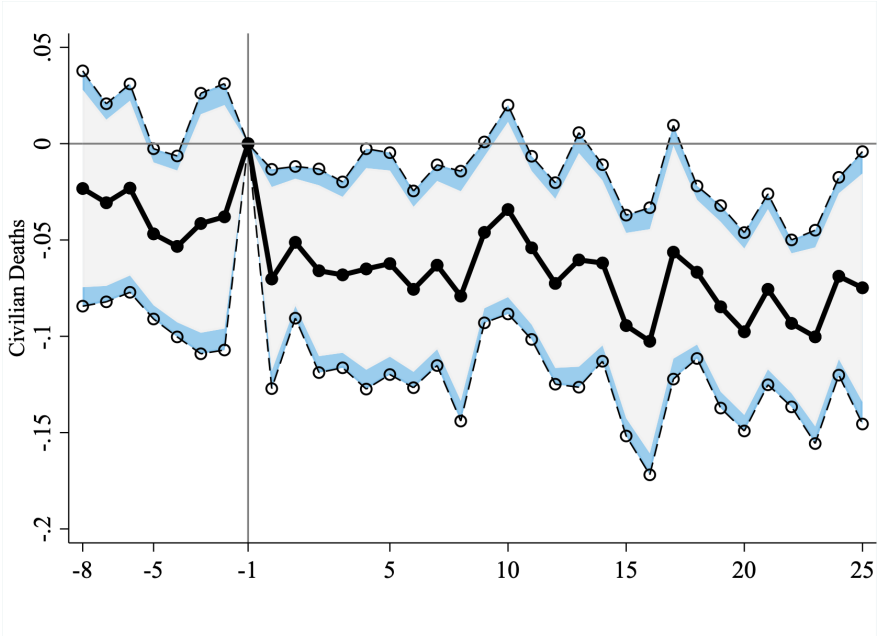
Notes: This figure plots regression estimates after employing propensity scores to reweigh counties in the control group by their inverse propensity scores to balance the distribution of covariates across groups (DiNardo et al., 1996). The dependant variable is the number of non-White policing-related civilian fatalities. This specification includes county and region-by-year fixed effects. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the long-sample only. The horizontal axis represents event-years (years before and after litigation).

Figure B6: Event Study – Least Squares Regressions – White Deaths

(a) White: OLS

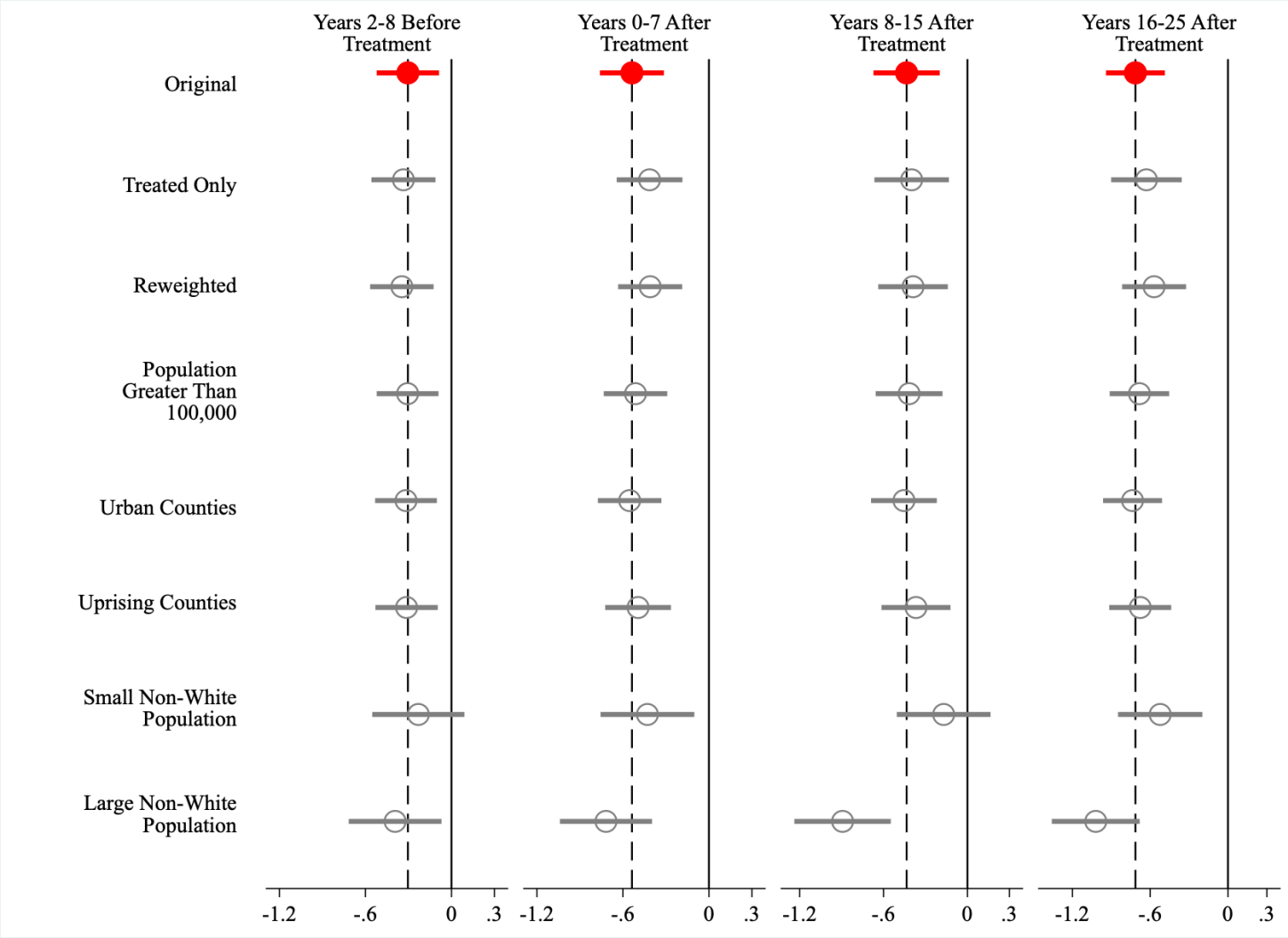


(b) White: WLS



Notes: This figure plots OLS and WLS regression estimates of the effect of the threat of court-ordered affirmative action on White policing-related civilian fatalities. Each specification includes county and region-by-year fixed effects. The dependent variable is the number of White policing-related civilian fatalities in panel (a) and White deaths per 100,000 White residents in panel (b). Panel (a) also includes the log of White population as a regressor and panel (b) uses the White population in 1960 as weights. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the long sample only. The horizontal axis represents event-years (years before and after litigation).

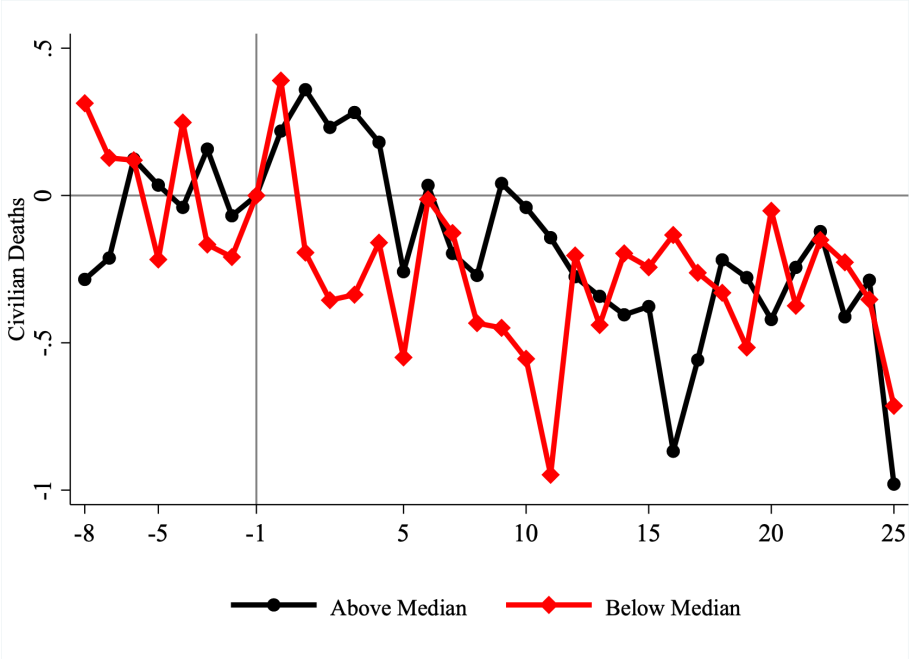
Figure B7: Robustness Checks - White Deaths



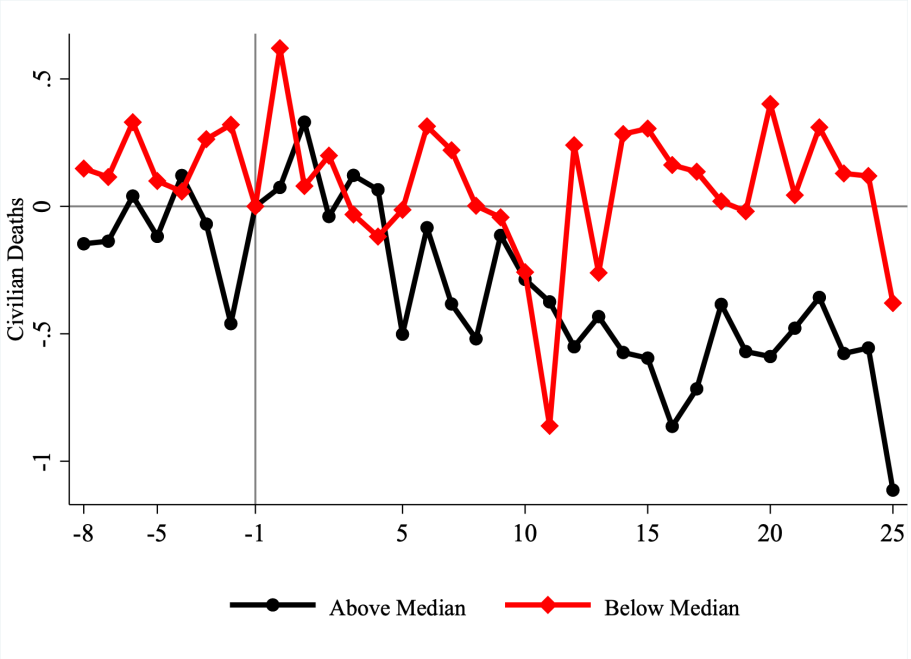
Notes: The figure displays Poisson estimates obtained from estimating Equation (1) by grouping event-years. All rows include county and region-by-year fixed effects. Heteroskedasticity-robust standard errors clustered by county are presented by the bold line. Joint effects are presented by circle markers. Row 1 presents the baseline joint treatment effects; Row 2 restricts the sample solely to treated counties; Row 3 employs a semi-parametric reweighing design; Row 4 restricts the sample to counties with a population greater than 100,000; Row 5 limit the sample to counties where the proportion of residents residing in urban areas was higher than the median share of urbanization in 1960; Row 6 restricts the sample to counties that experienced at least one racial uprising in the 1960s; Rows 7 restricts the treatment group into counties where the share of the non-White population in 1960 is below the median share for the treatment group; Row 8 restricts the treatment group into counties where the share of the Black population in 1960 is above the median share for the treatment group.

Figure B8: Event Study Results – Intensity of Treatment

(a) Total Population Covered

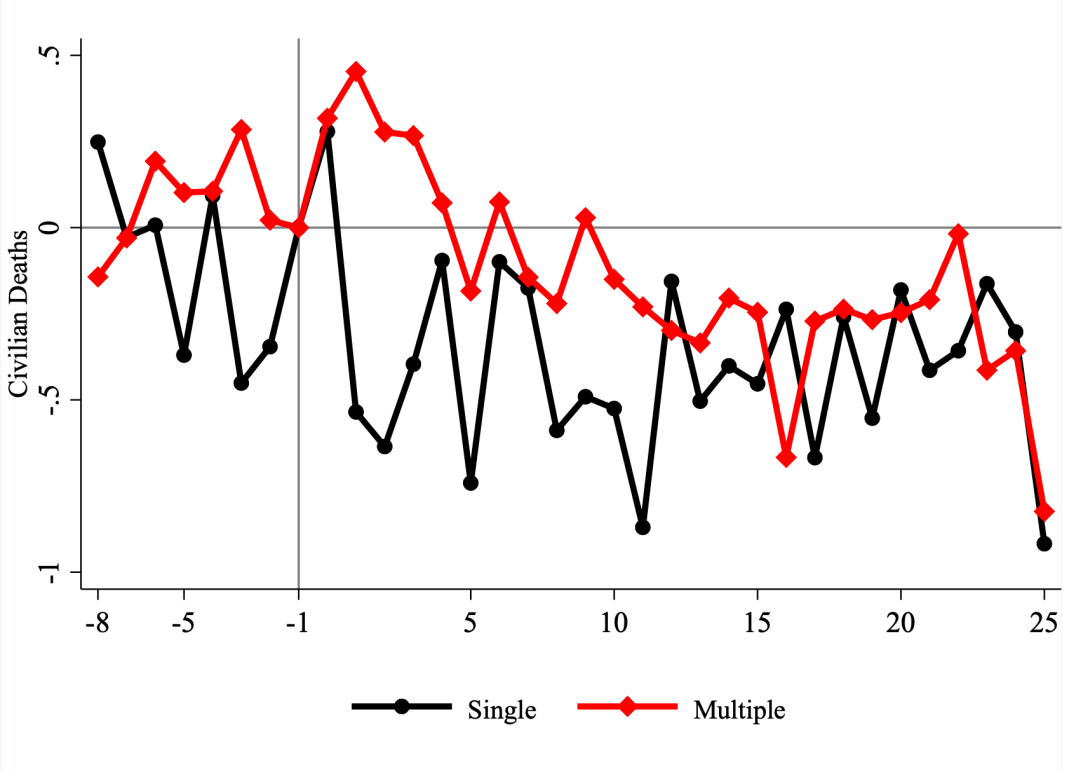


(b) Total Non-White Population Covered



Notes: This figure stratifies counties by treatment coverage, which is defined by the percentage of the county's population resides in treated cities within the county. The dependent variable is the number of non-White policing-related civilian fatalities. Panel (a) stratifies treated counties by locations with above-below median coverage. Panel (b) stratifies treated counties by locations with above-below median non-White population coverage. In both panels, the control group are non-treated locations, irrespective of county demographics. Each specification includes county and region-by-year fixed effects. The horizontal axis represents event-years (years before and after litigation).

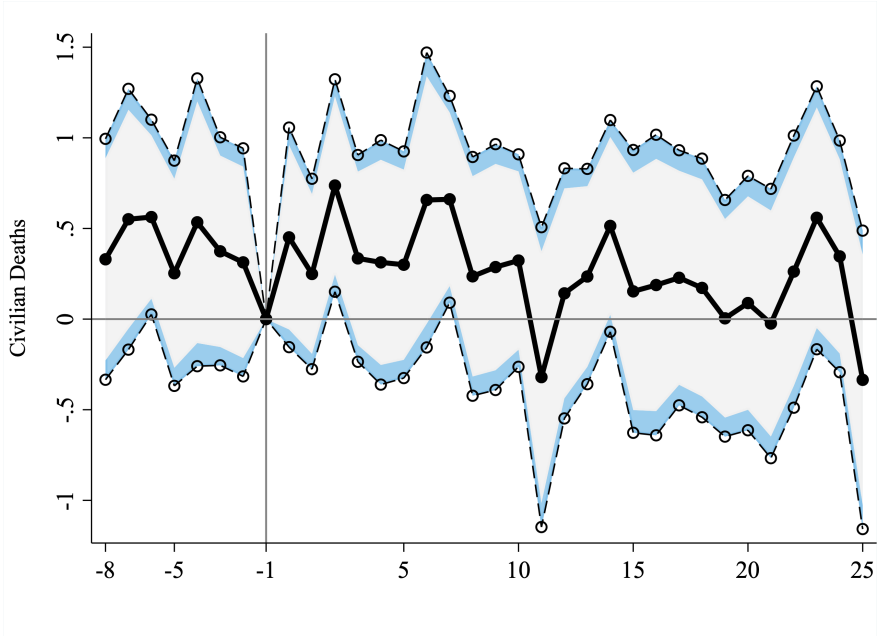
Figure B9: Event Study Results – Multiple Treatments



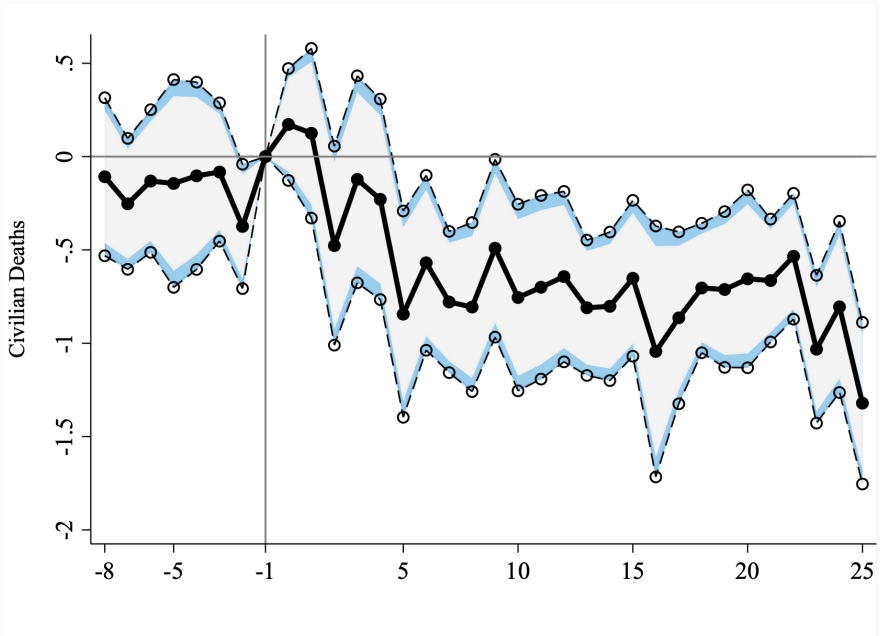
Notes: The figure shows pre- and post-treatment effects for counties with one treated city are compared to control group versus when counties with multiple treated cities are compared to the control group. The dependent variable is the number of non-White policing-related civilian fatalities. In each specification, the control group are non-treated locations, irrespective of county demographics. Each specification includes county and region-by-year fixed effects. The horizontal axis represents event-years (years before and after litigation).

Figure B10: Event Study Estimates – By Region

(a) Non-White - South

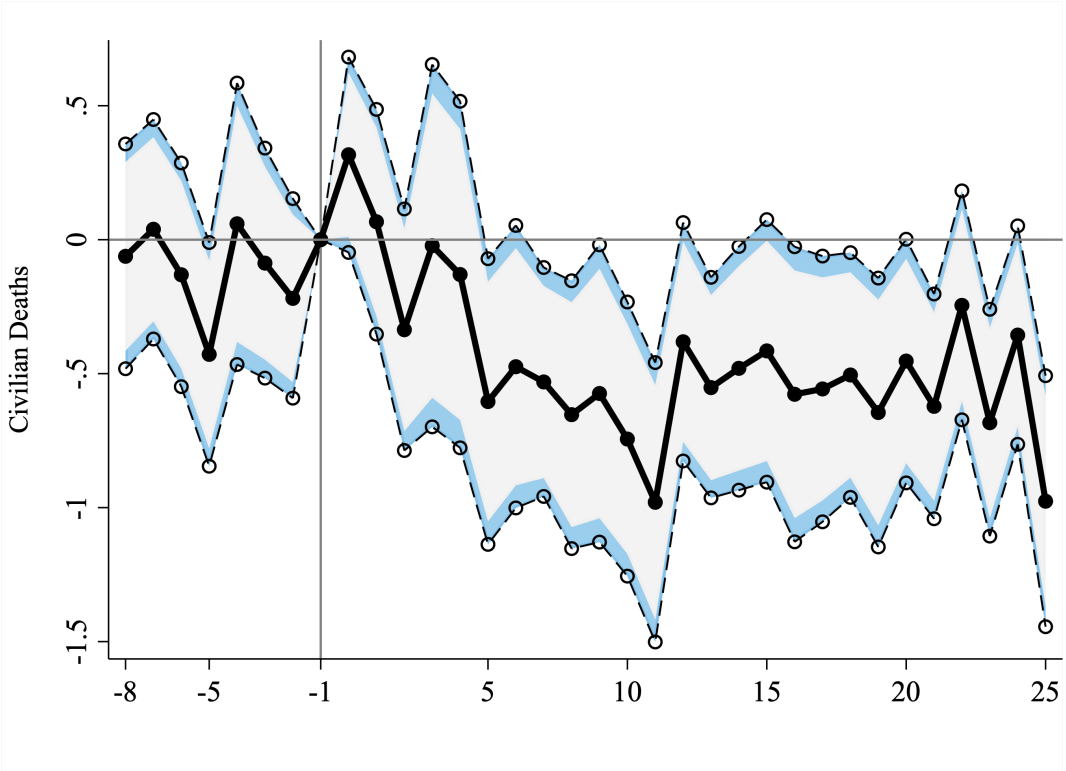


(b) Non-White - Rest of the Country



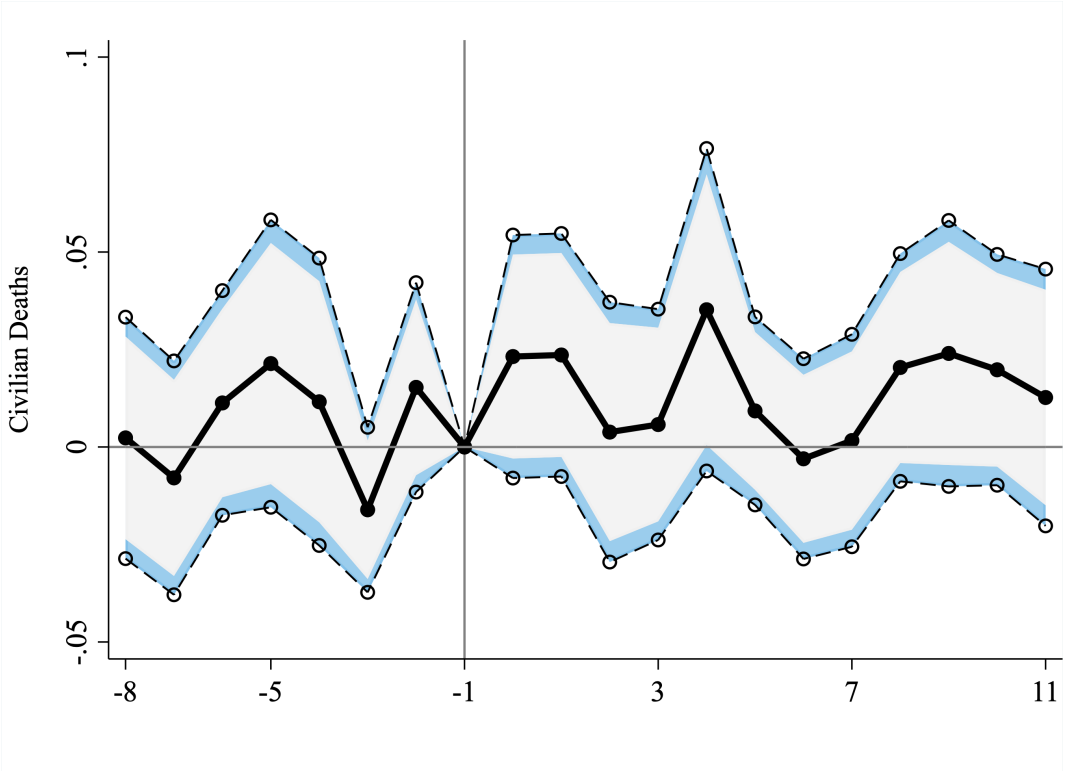
Notes: This figure plots estimates by geographic region. The dependent variable is the number of non-White policing-related civilian fatalities. Panel (a) reports estimates for the South and panel (b) reports estimates for the rest of the country and includes region-by-year fixed effects (Northeast, Midwest, and West). Robust standard errors by county. The horizontal axis represents event-years (years before and after litigation).

Figure B11: Event Study – Non-White Deaths – McCrary (2007) Sample



Notes: This figure plots estimates of the effect of Affirmative Action litigation on police killings of non-White civilians limited to counties identified by McCrary (2007) and that report at least 52 years between 1960 and 2016. This specification includes county and region-by-year fixed effects. Treatment group corresponds to counties with cities treated prior to 1987. Robust standard errors by county. The horizontal axis represents event-years (years before and after litigation).

Figure B12: Event Study Results – Contiguous Counties

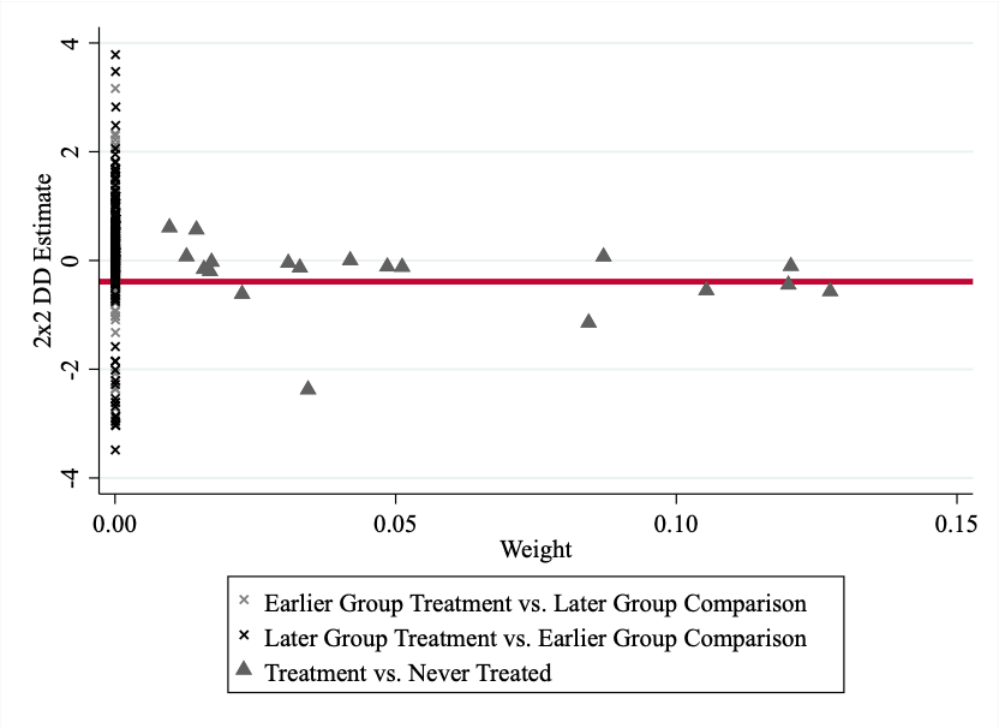


Notes: This figure plots pre-and-post-treatment effects when the treatment status and timing is applied to contiguous untreated counties. The comparison group consists of non-treated non-contiguous counties. Treated counties dropped from the sample. This specification includes county and region-by-year fixed effects and accounts for exposure with non-White population. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

B.2 Heterogeneity in the Treatment Effects Across Timing Groups

We explore heterogeneity in the treatment effects across timing groups for two reasons. One, treatment is a direct response to the 1960 uprisings; therefore, the initial uptick in police-related fatalities may be in response to uprisings and not police actions. Also, the number of newly treated locations drastically slows down during the Reagan Administration. This is especially important because successful litigation depends on involvement from the Equal Employment Opportunity Commission. Secondly, we explore heterogeneity in the treatment effects across timing groups due to potential bias from implementing the standard two-way fixed-effects difference-in-difference model (Goodman-Bacon, 2021). Figure B13 plots the decomposition proposed by Goodman-Bacon (2021).

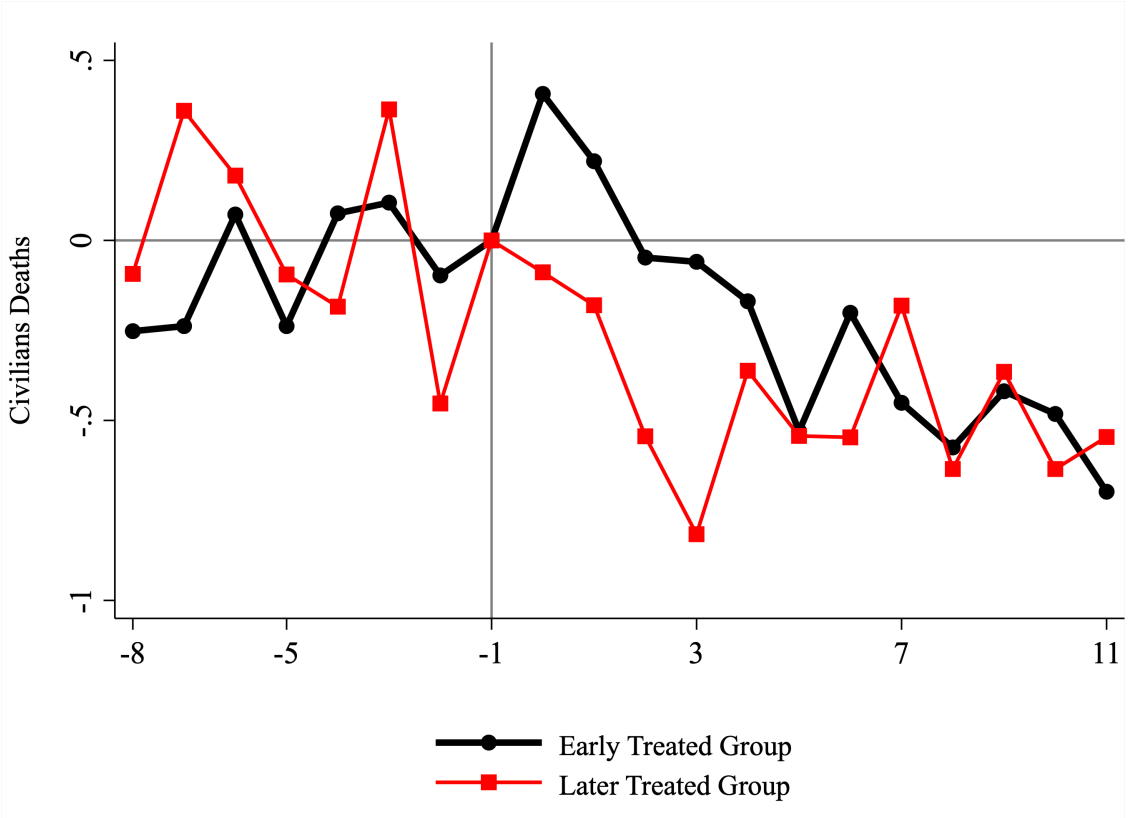
Figure B13: Difference-in-Difference Decomposition



Notes: This plots the decomposition proposed by Goodman-Bacon (2021) for the effect of Affirmative Action litigation on non-White policing-related civilian fatalities.

In Appendix Figure B14, we split treated locations into two groups: 1) those treated prior to 1981 and 2) those treated after 1980. Pre- and post-treatment effects are estimated using non-treated locations as the control group.

Figure B14: Event Study Heterogeneity in Timing – Non-White Deaths



Notes: This figure splits treated locations by those treated before 1981 (Black circle markers) and those treated after 1980 (red sqare markers. Pre- and post-treatment effects are estimated using non-treated locations as the control group. Each specification includes county and region-by-year fixed effects.

In Appendix Figure B15, we diverge from the traditional two-way-fixed effects model (TWFE) and use the estimator from Callaway and Sant'Anna (2020) for estimating unbiased average treatment effects on the treated (ATET) when there is variation in the timing of treatment.[49] The traditional TWFE model will produce biased ATET when there is heterogeneity in the treatment effect across treatment groups or heterogeneity in the treatment effect over time (equivalent to a change in the slope parameter). The bias in the TWFE is a result of how the ATET is calculated. In general, the DiD estimate is a combination of estimates comparing the treated group to the never-treated group, as well as a comparison between groups treated at different times. Appendix Figure B13 displays the weights applied to each group to produce the difference-in-difference estimate from a generic TWFE model with variation in the timing of treatment. Hence, treated groups serve as both treatments and controls when estimating the ATET. More specifically, the bias occurs when the later treated groups are compared to earlier treated groups. When making this comparison, the parallel trends assumption is likely violated (due to the treatment of earlier units), producing biased estimates (Goodman-Bacon, 2021). Relatedly, Sun and Abraham (2020) show that a similar bias exists in TWFE event-study analyses when there is heterogeneity in the dynamics of the ATET across timing groups. Therefore, TWFE will bias the ATET away from the true effect when there

[49]The traditional TWFE model includes both year and location fixed effects and a dummy variable to capture the treatment: $y_{it} = \alpha_i + \gamma_t + \beta^{DD} D_{it} + \varepsilon_{it}$.

is heterogeneity in the treatment effect over time or across groups.

To avoid the bias associated with TWFE models, we first employ the estimator from Callaway and Sant'Anna (2020) (CS) in an event-study framework. We stack our data so that stack k includes one timing group t_k^* and other counties that will be never be treated. The never-treated group serve as the control group for treatment group k. This method uses calendar time to compare and contrast treated locations with the control group. For example, for those counties treated in 1977, the comparison group will include the calendar years 1971 to 2002 for never-treated counties.

To estimate unbiased pre- and post-treatment effects, we create all possible stacks for timing groups as previously highlighted and calculate means by treatment group (g), time (t), and stack (k). Then we estimate the following equation:

$$Y_{kgt} = \alpha_{kg} + \gamma_{gt} + \sum_{y=-6}^{-2} \pi_y^{CS} D_g \mathbb{1}(t - T_g^* = y) + \sum_{y=0}^{25} \pi_y^{CS} D_g \mathbb{1}(t - T_g^* = y) + \epsilon_{kgt} \tag{3}$$

where Y is police-related fatalities by race. We include treatment group fixed effects for every stack α and time fixed effects for every stack γ. The Callaway and Sant'Anna estimator is captured by π^{CS}, which will provide pre- and post-treatment for group g. The estimator will calculate event-study estimates for each group. The reported pre- and post-treatment effects are a weighted average of each group's estimated effects using the number of treated counties in each timing group as weights. We construct 95 present confidence intervals from 250 draws of a block bootstrap that resamples the counties in our analysis. We aggregate the estimated average treatment effects by the number of resampled counties in each timing group and use this distribution for statistical inference. Due to the manner of computation, these results should be compared to the OLS results. If the TWFE model produced unbiased results, the CS estimator should produce similar estimates.

Figure B15: Event Study –Callaway and Sant'Anna Estimator

Notes: The figure plots regression coefficients estimated using Callaway and Sant'Anna (2020) estimators to avoid biases associated with two-way fixed effects models highlighted in Goodman-Bacon (2021). The dependent variable is the number of non-White deaths due to legal intervention. This specification includes county and region-by-year fixed effects. Both 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after threat of litigation).

B.3 Racial Disparities in Police Killings of Civilians

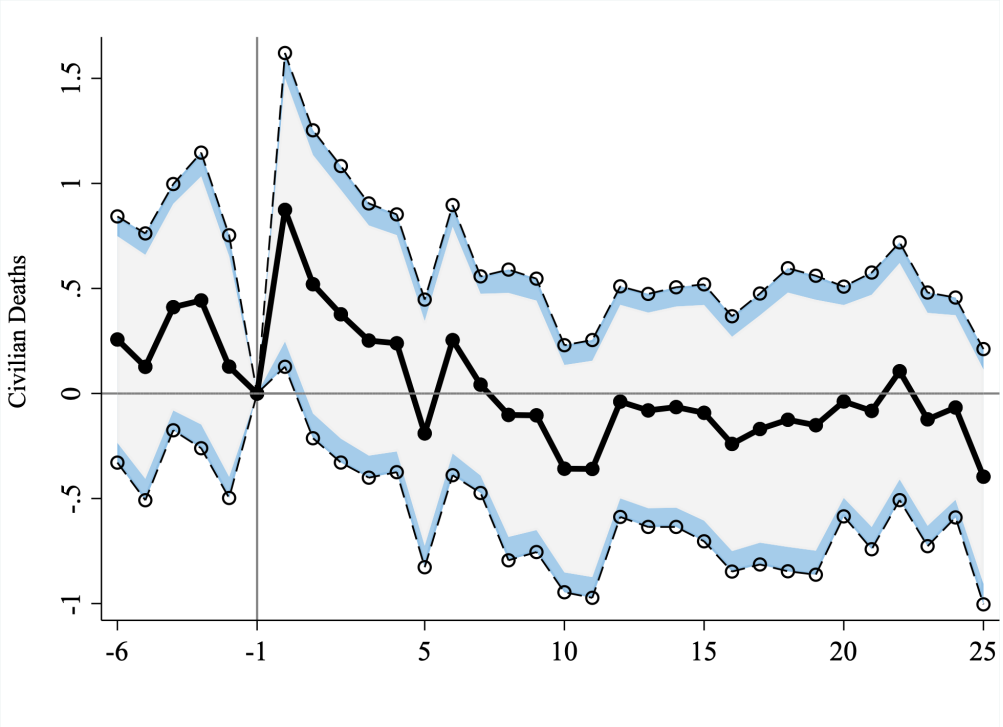
We test for racial disparities in police killings by converting equation (1) into a triple difference strategy (DDD) following Cox and Cunningham (2021). We estimate the following equation using OLS:

$$
\begin{aligned}
y_{ct} = {} & \alpha_c + \gamma_{r(c),t} + NW_k + NW_k\gamma_t + NW_k D_c \\
& + \sum_{j=2}^{7} \pi_j D_c \mathbb{1}\{t - t_c^* = -j\} + \sum_{j=0}^{26} \phi_j D_c \mathbb{1}\{t - t_c^* = j\} \\
& + \sum_{j=2}^{7} \lambda_j D_c \mathbb{1}\{t - t_c^* = -j\} NW_k + \sum_{j=0}^{26} \sigma_j D_c \mathbb{1}\{t - t_c^* = j\} NW_k + v_{ct}
\end{aligned}
$$

where the notation stays as earlier defined in equation 1, and NW_k is equal to one for non-White police killings and zero for White police killings. In this model, we can capture pre-existing trends in racial disparities in police killings of civilians. If the model is correctly specified, the pre-treatment effects will be indistinguishable from zero and any trend break in racial disparities in police killings will be attributed to the threat of litigation. Negative post-treatment effects would suggest that litigation reduces racial disparities in police killings, while positive post-treatment effects would imply that litigation contributes to racial disparities in police killings. If the post-treatment effects are zero, then litigation may not affect racial disparities in police killings.

Appendix Figure B16 plots pre- and post-treatment effects from the OLS model. Overall, pre-treatment effects are not statistically significant. Our model indicates an initial statistically significant increase in racial disparities in police killings at event-year zero. This is not surprising considering the initial rise in non-White deaths in Appendix Figure B2a and the initial decrease in White deaths in Appendix Figure B6a. After the initial increase in racial disparities of police killings, post-treatment effects decrease and eventually become negative, but are not statistically significant. Nonetheless, the initial increase in racial disparities could indicate a distinct change in police behavior in response to federal intervention.

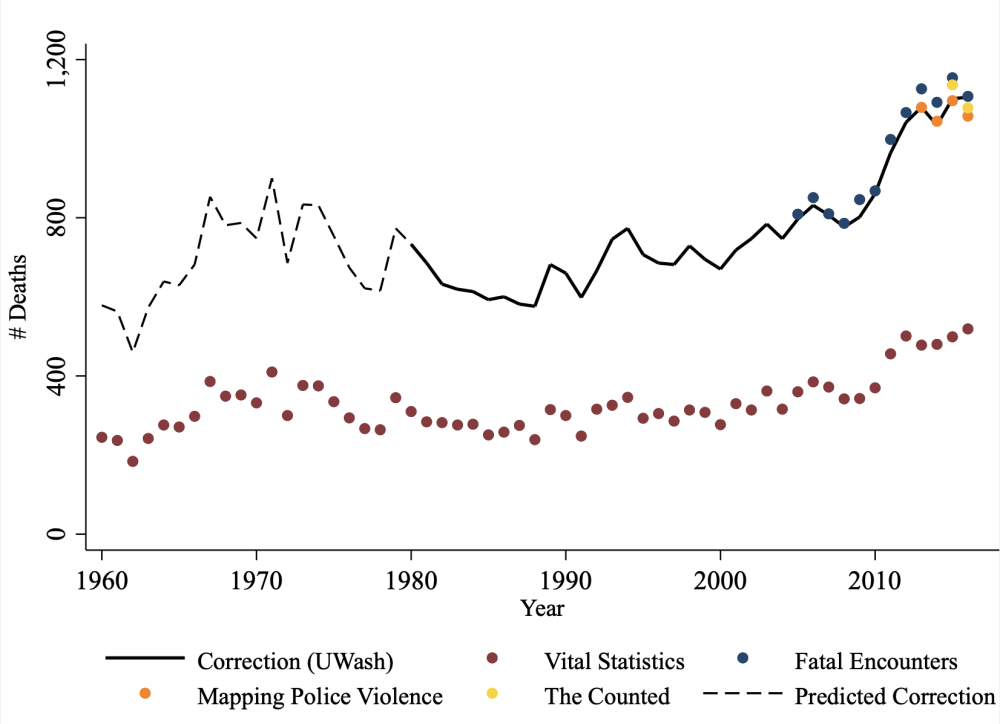
Figure B16: Triple Difference – Racial Disparities in Deaths Due to Legal Intervention



Notes: This figure plots estimates for the disparity in policing-related civilian fatalities. This ordinary least squares regression specification includes county and region-by-year fixed effects. The coefficients are estimated from a Difference-in-Difference-in-Difference model (see (Cox and Cunningham, 2021)). Marginal effects show the relative change in non-White deaths relative to White deaths in treated counties. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

B.4 Measurement Error and Internal Validity

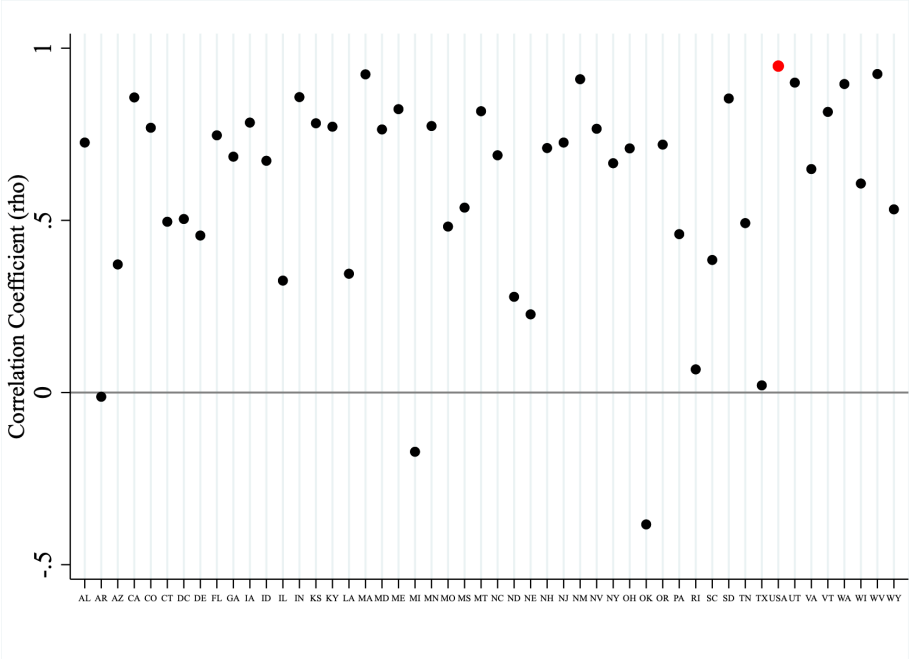
Figure B17: Police Killings of Civilians, 1960-2019



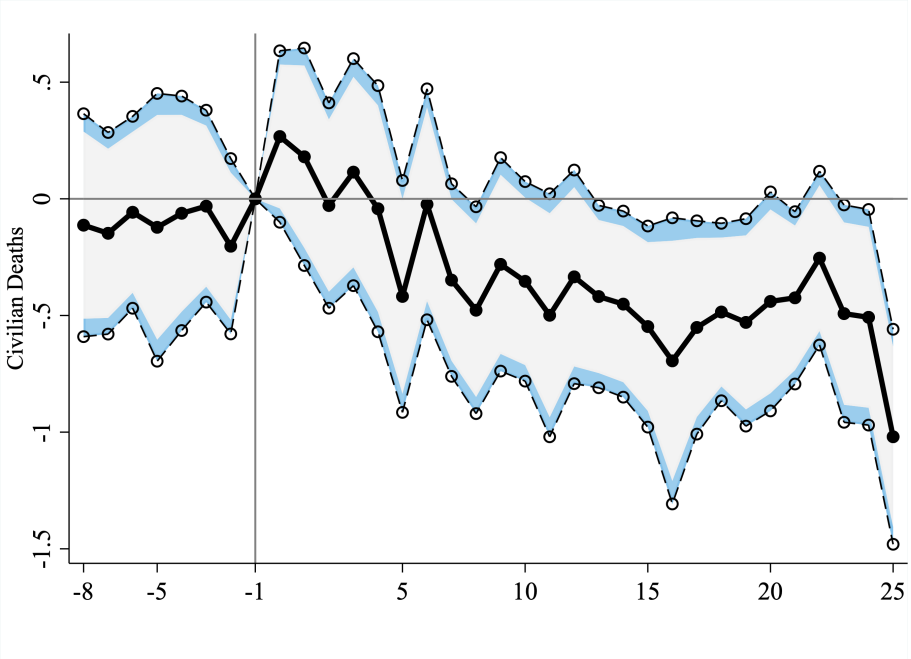
Notes: The figure replicates Collaborators et al. (2021), showing policing-related civilian fatalities from 1960 to 2019. We compute deaths from 1960 to 1979 as a function of reported deaths, bias in reporting, and state-and-year fixed effects.

Figure B18: Police Killings Correlation Ratios by States, 2000-2016

(a) Correlations by State



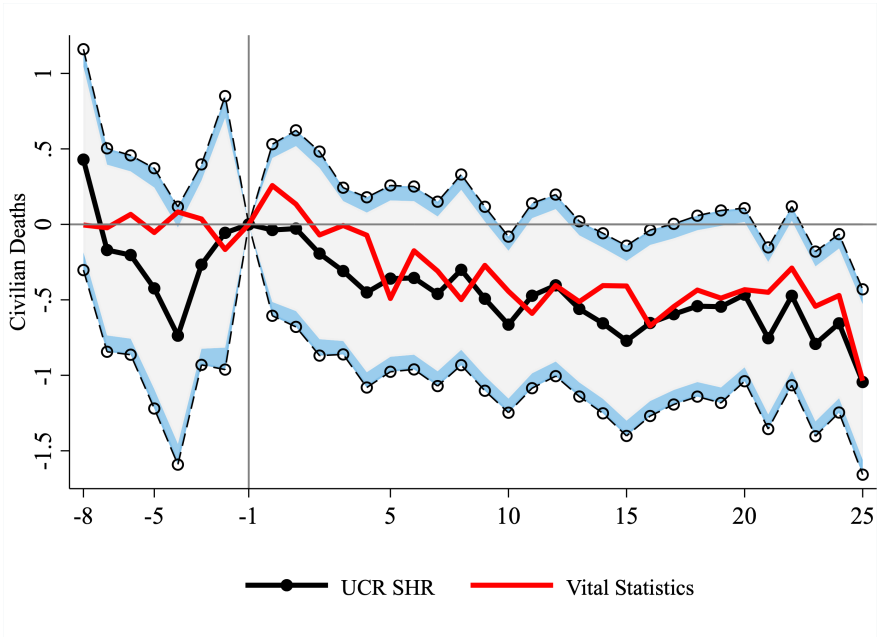
(b) Drop States with Low or Negative Correlation Ratios



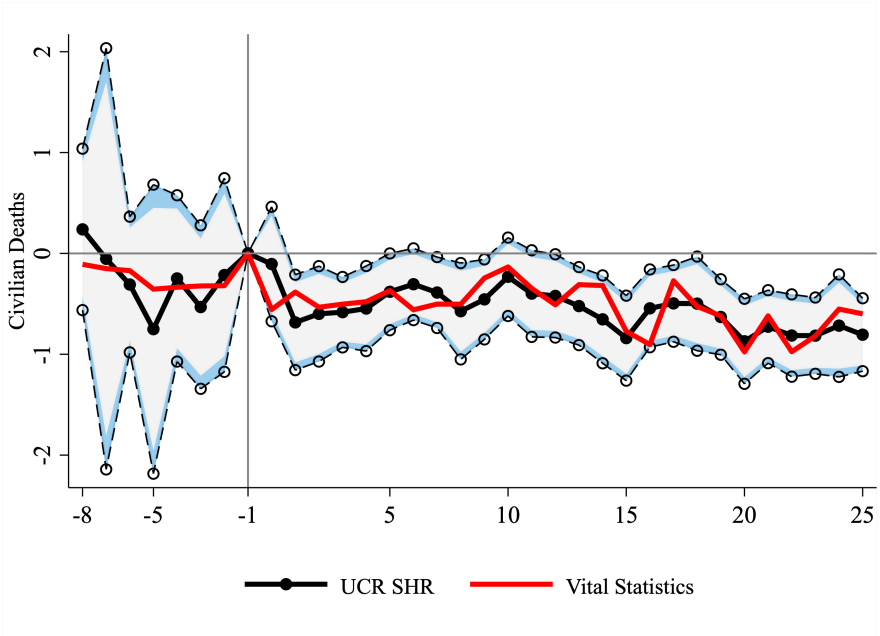
Notes: Panel (a) plots the correlation coefficient for non-White policing-related civilian fatalities for Fatal Encounters and Vital Statistics data by state. Panel (b) plots coefficients from a Poisson regression specification that includes county and region-by-year fixed effects and accounts for exposure with non-White population. Panel (b) drops states with low or negative correlation (below .3). The dependent variable is non-White policing-related civilian fatalities. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

Figure B19: Event Study – Deaths Due to Legal Intervention: UCR SHR

(a) Non-White Deaths



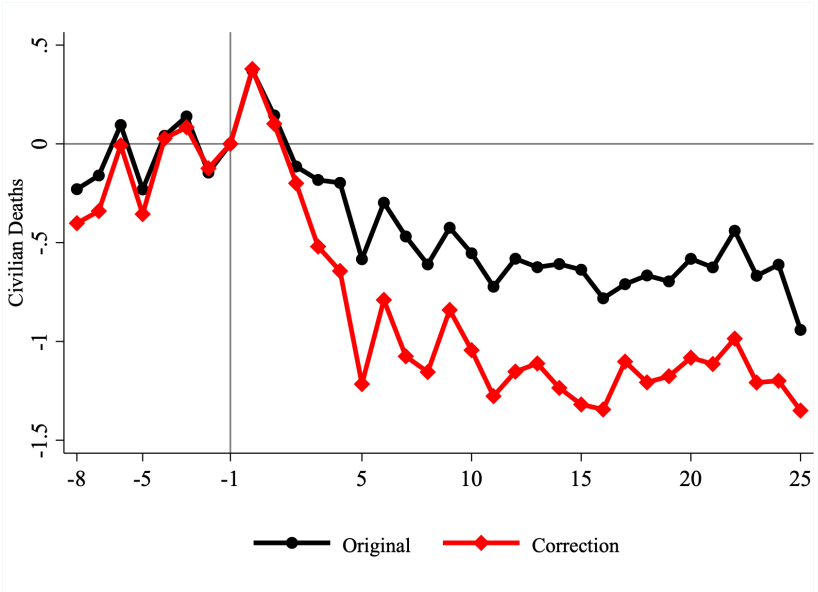
(b) White Deaths



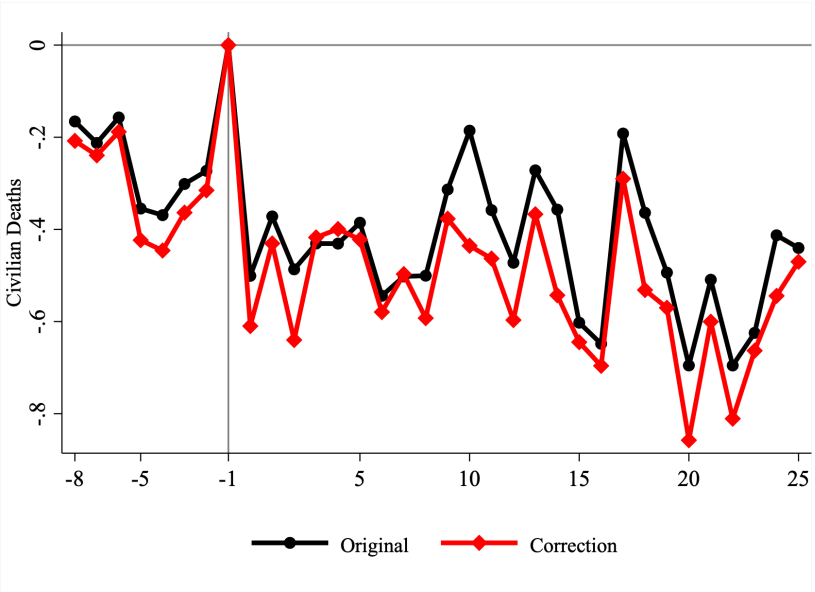
Notes: Data on deaths due to policing comes from the UCR: Supplemental Homicide Report which begins in 1976. Panels (a) and (b) plot pre- and post-treatment effects from an unbalanced panel for counties treated prior to 1987. All regressions include county and region-by-year fixed effects. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented for the vital statistics smaple only. The horizontal axis represents event-years (years before and after litigation).

Figure B20: Event Study – Estimated Deaths Due to Legal Intervention

(a) Non-White: OLS

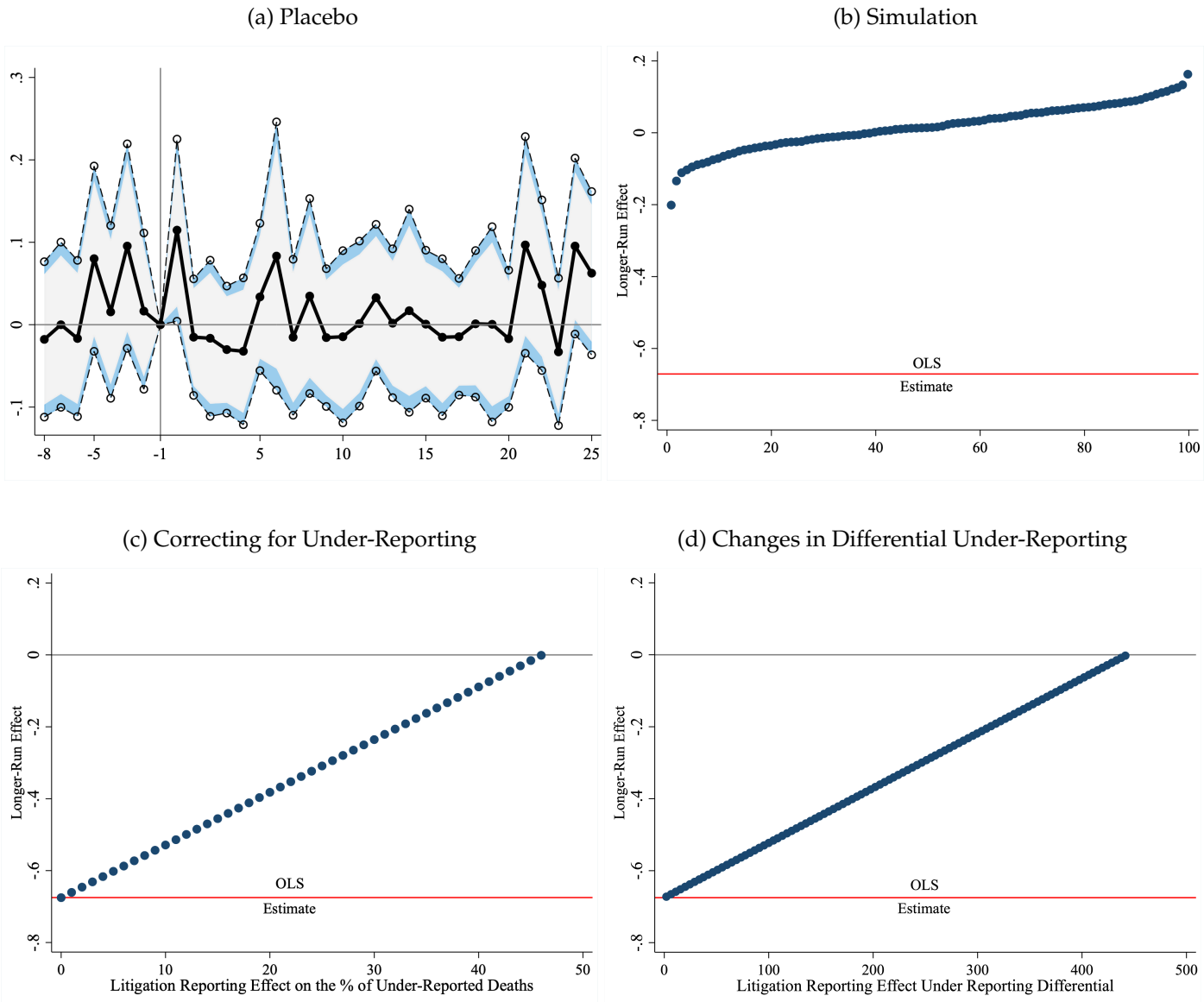


(b) White: OLS



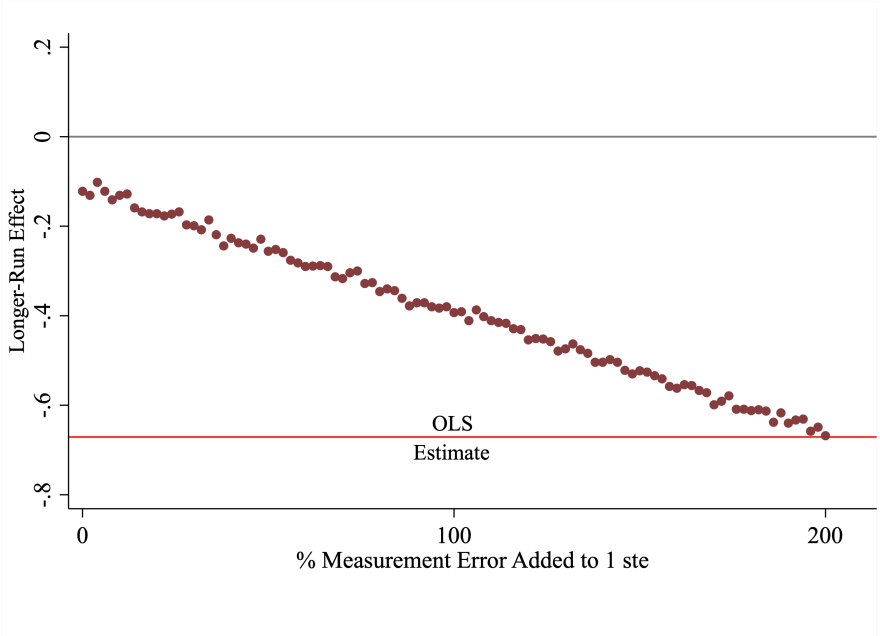
Notes: This figure plots the effect of Affirmative Action litigation on policing-related civilian fatalities after correcting for underreproting. The red line corresponds to the corrected estimated police killings using Collaborators et al. (2021), while the Black line refers to the original estimate. Each specification includes county and region-by-year fixed effects. The horizontal axis represents event-years (years before and after litigation).

Figure B21: Measurement Error & Changes in Under-reporting



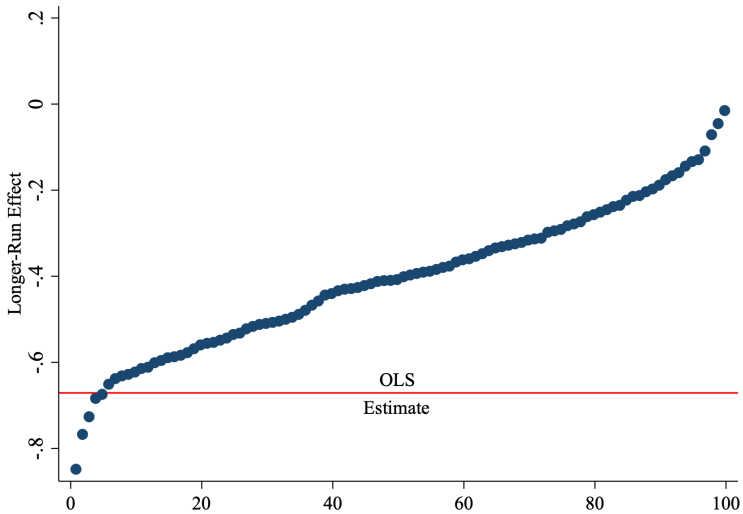
Notes: All Figures report estimates from an Ordinary Least Squares regression. Panel (a) plot pre- and post-treatment effects from a placebo that randomly assigns treatment status and treatment year to 60 counties that experienced uprisings in the 1960s. Panel (b) runs the placebo in panel (a) 500 times and plot long-run joint treatment effects (event-year 16 to 26). Both panel (c) and panel (d) plot long-run joint treatment effects. Panel (c) corrects for reporting to estimate how much bias would reduce the OLS result to zero. Panel (d) estimates how much differential under-reporting between the treatment and control group would litigation have to induce to reproduce our results.

Figure B22: Measurement Error in Police Killings of Non-White Civilians



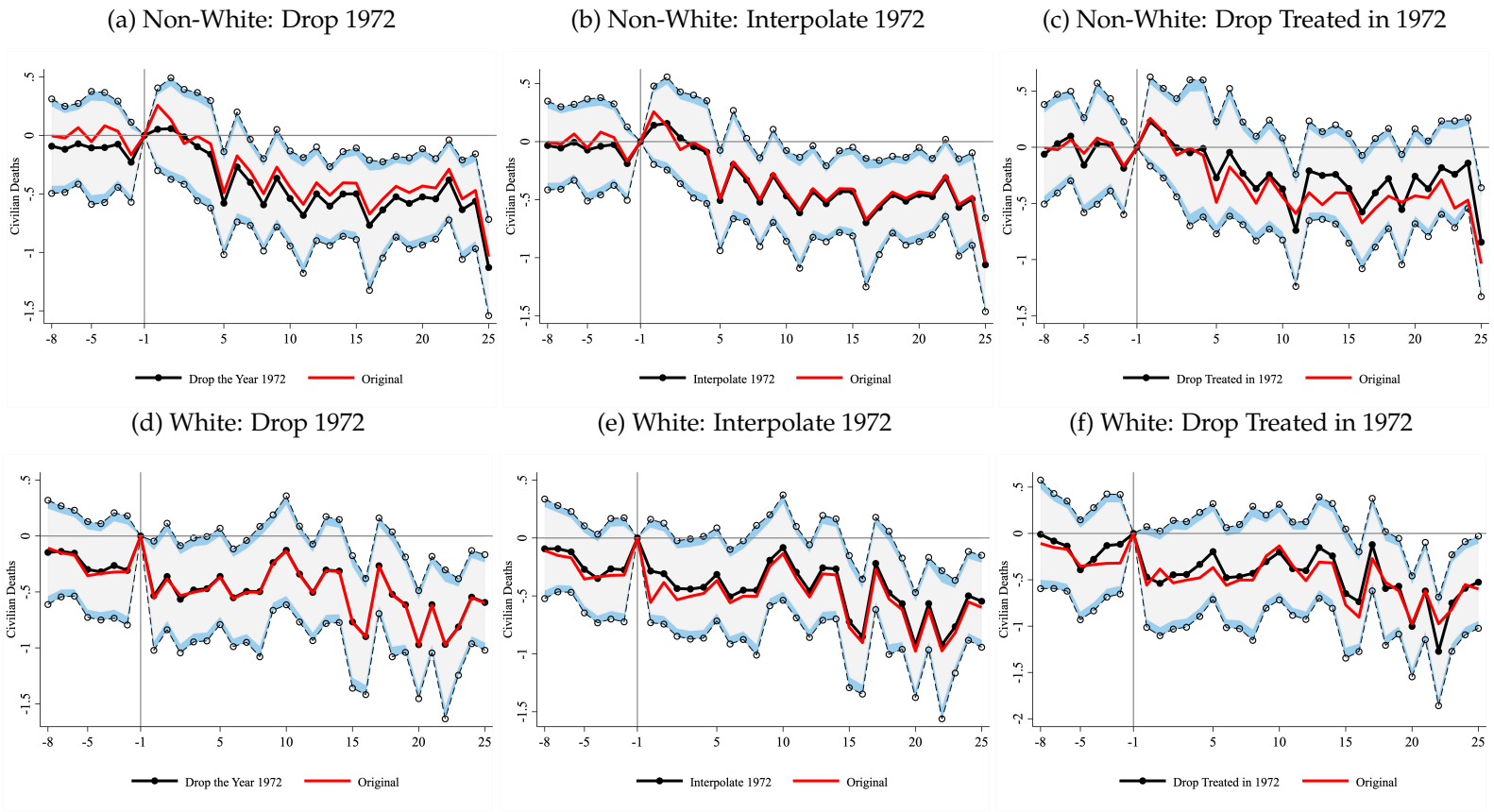
Notes: This figure plots the long-run effects from a placebo specification that randomly assigns treatment status and treatment year to 60 counties that experienced uprisings in the 1960s. The long-run effects are plotted after repeatedly subtracting measurement error, pulled from a uniform distribution, from the outcome of non-White policing-related civilian fatalities. Measurement error is reported as a percent of the standard deviation.

Figure B23: Measurement Error: Fatal Encounters vs. Vital Statistics



Notes: This figure plots the long-run effects from a placebo specification that randomly assigns treatment status and treatment year to 60 counties that experienced uprisings in the 1960s. The long-run effects are measured after repeatedly subtracting measurement error, pulled from a uniform distribution, from the difference in non-White policing-related civilian fatalities in NVSS and Fatal Encounters. XX

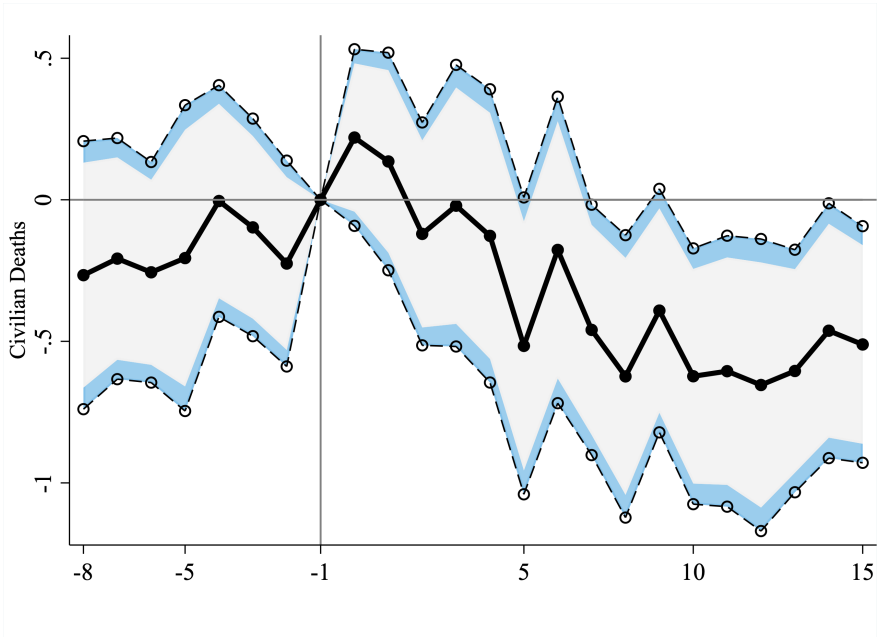
Figure B24: Different Imputation Methods for the Year 1972



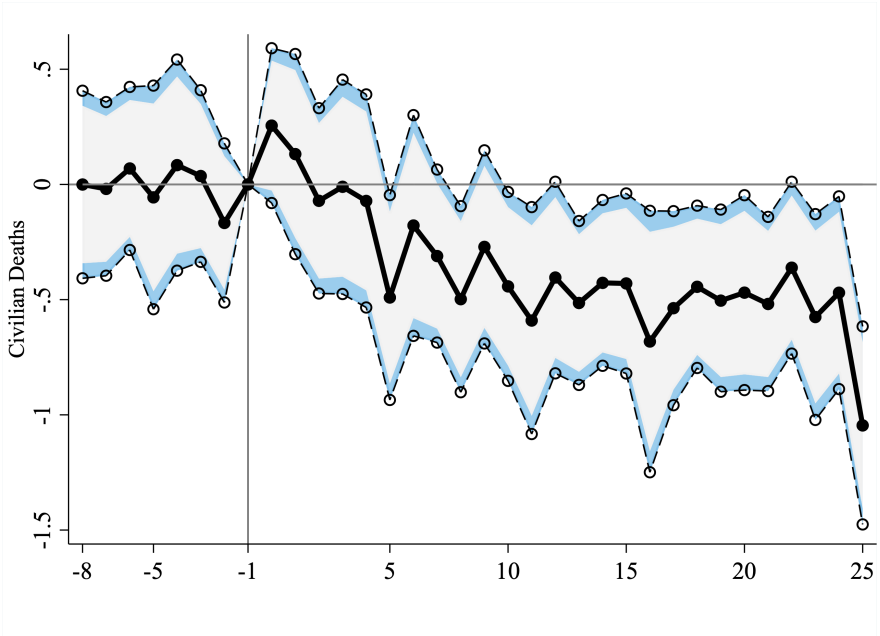
Notes: This figure plots the effect of Affirmative Action litigation on policing-related civilian fatalities after addressing issues with data for 1972. All panels estimates are from Poisson models that includes county and region-by-year fixed effects and accounts for exposure with population by race. Panels (a) and (d) drops the year 1972, panels (b) and (e) use the average number of deaths in 1971 and 1973 to estimate the number of deaths in 1972, and panel (c) and (f) drop locations treated in 1972. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

Figure B25: Event Study – Accounting for Pattern or Practice Investigations

(a) Pre-Pattern or Practice Investigations



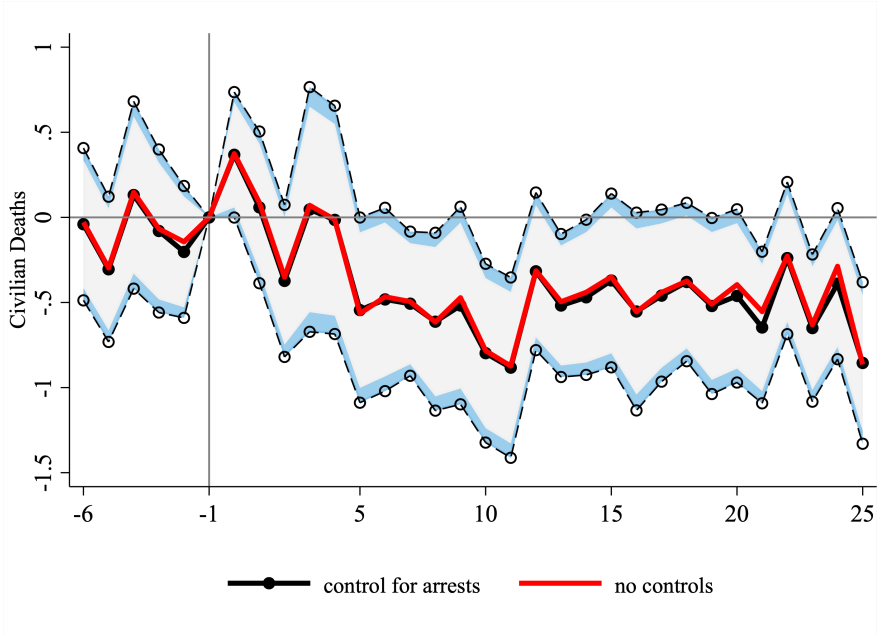
(b) Account for Federal Consent Decrees



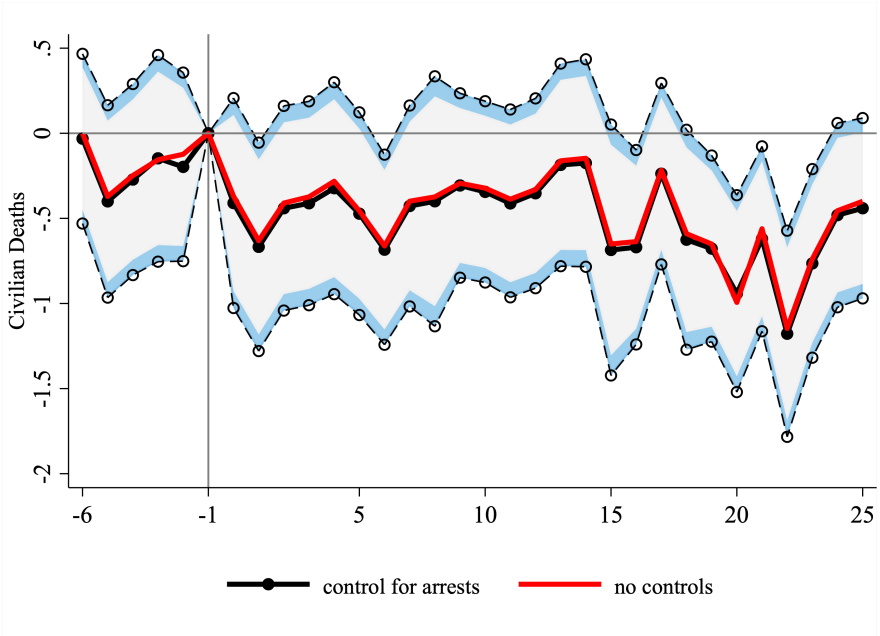
Notes: The figure shows pre- and post-treatment effects of the effect of affirmative action litigation on non-White policing-related civilian fatalities when accounting for practice and patterns investigation and federal consent decrees. Pattern or Practice Investigations begin in 1991. Panel (a) drops locations treated after 1975 and examines the impact of the threat of affirmative action pre-1991. In panel (b) we include indicators for the years before and the years after a consent decree. All regressions include county and region-by-year fixed effects. Robust standard errors are clustered at the county level, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

Figure B26: Control for Arrests

(a) Non-White Deaths

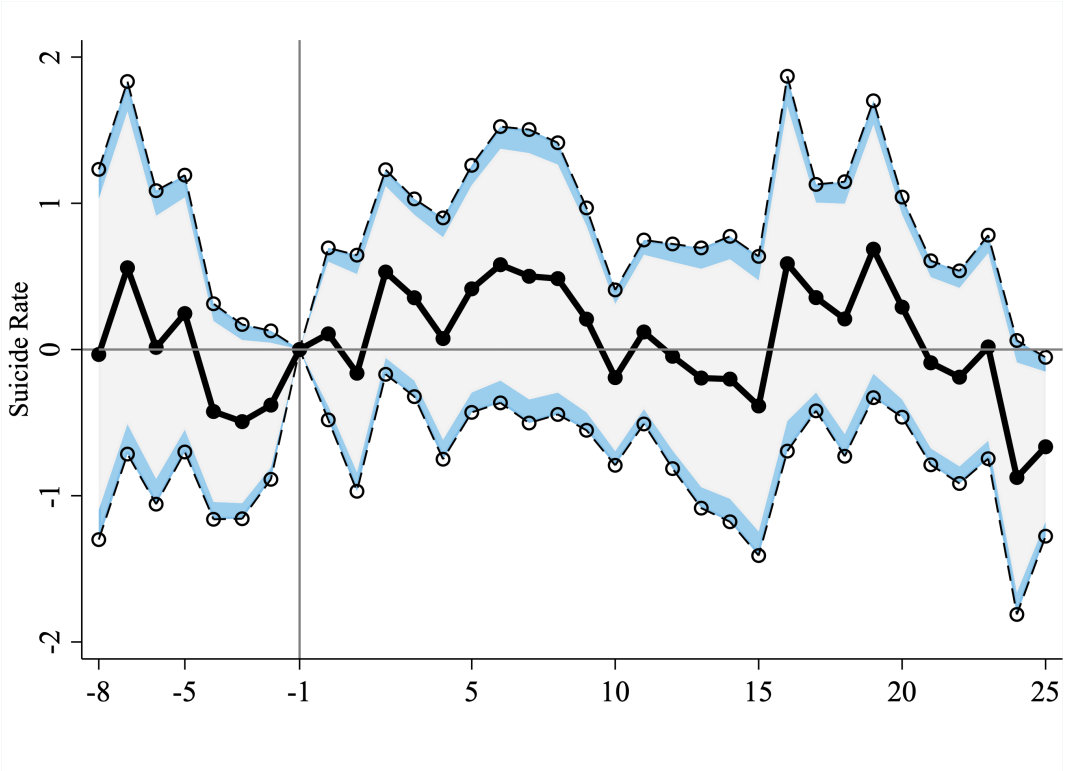


(b) White Deaths



Notes: This figure plots the effect of Affirmative Action litigation on policing-related civilian fatalities after controlling for arrests. The Poisson regression specification includes county and region-by-year fixed effects and accounts for exposure with non-White population. Results are from a sub sample of counties with arrests data for municipalities that report consistently. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

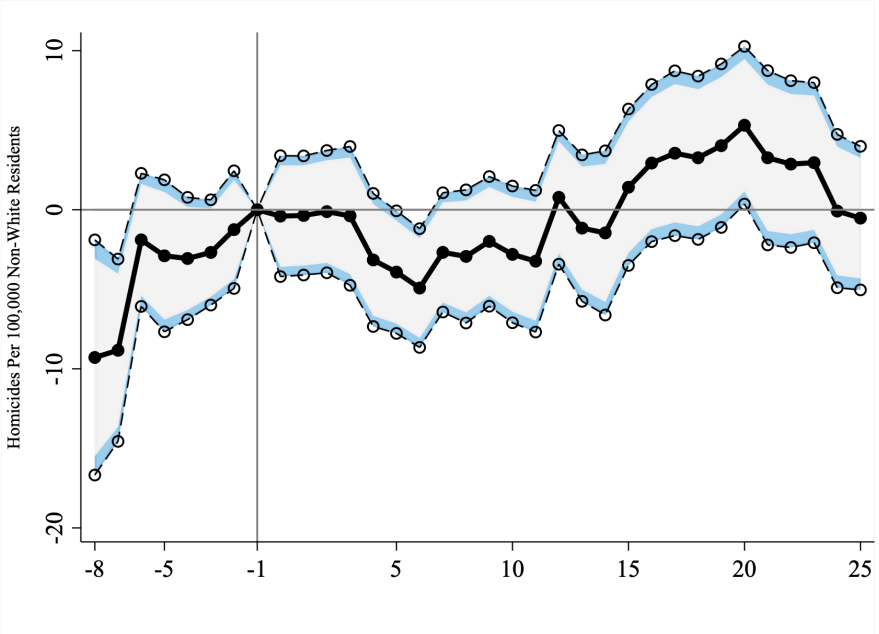
Figure B27: Event Study Results – Non-White Suicides



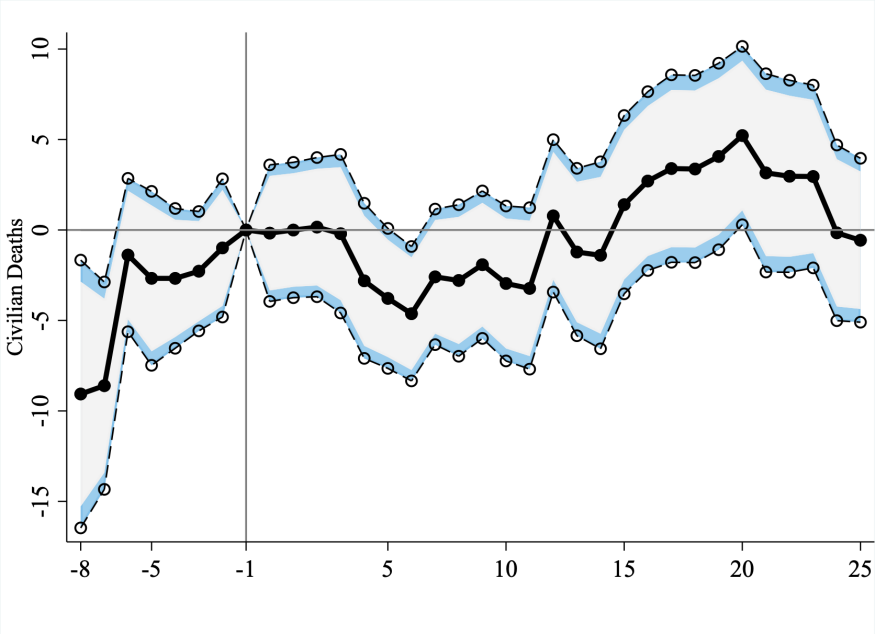
Notes: This figure plots the effect of Affirmative Action litigation on non-White suicides. This weighted least squares regression specification includes county and region-by-year fixed effects. Robust standard errors are clustered by county and 1960 non-White population is used as weights. The horizontal axis represents event-years (years before and after litigation).

Figure B28: Event Study – Non-White Homicide Victimization Rates

(a) Homicides



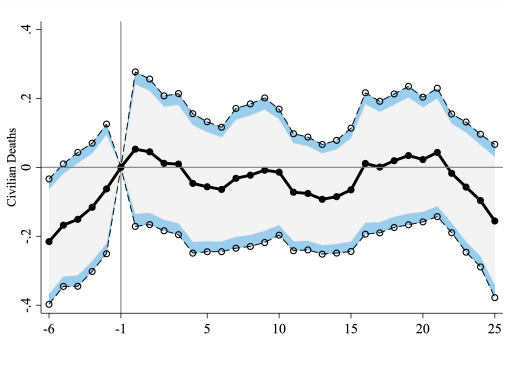
(b) Homicides & Police Killings



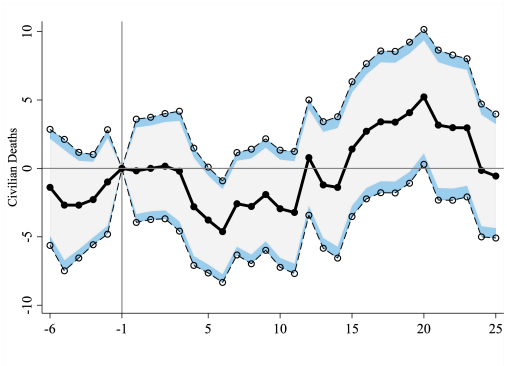
Notes: This figure plots the effect of Affirmative Action litigation on homicide victimization. In panel (a) the dependent variable is the non-White homicide victimization rates per 100,000 non-White residents. In panel (b) the dependent variable is the number of non-White homicides and policing-related civilian fatalities per 100,000 non-White residents. Each regression specification includes county, region-by-year, and urbanicity-by-year fixed effects. Robust standard errors are clustered at the county level, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

Figure B29: Event Study - Combine Homicides and Police Killings

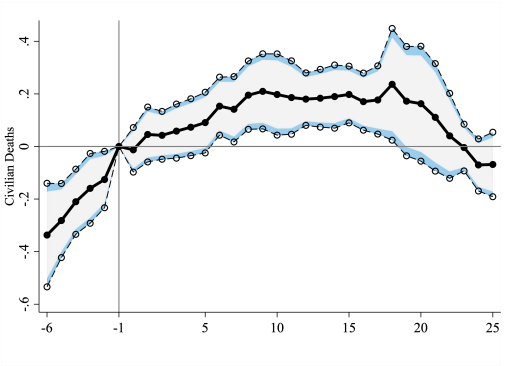
(a) Non-White Deaths



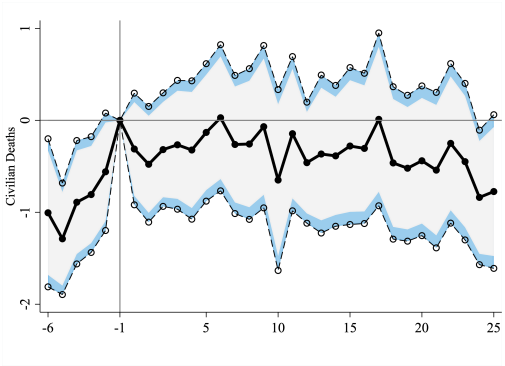
(b) Non-White Per Capita



(c) White Deaths



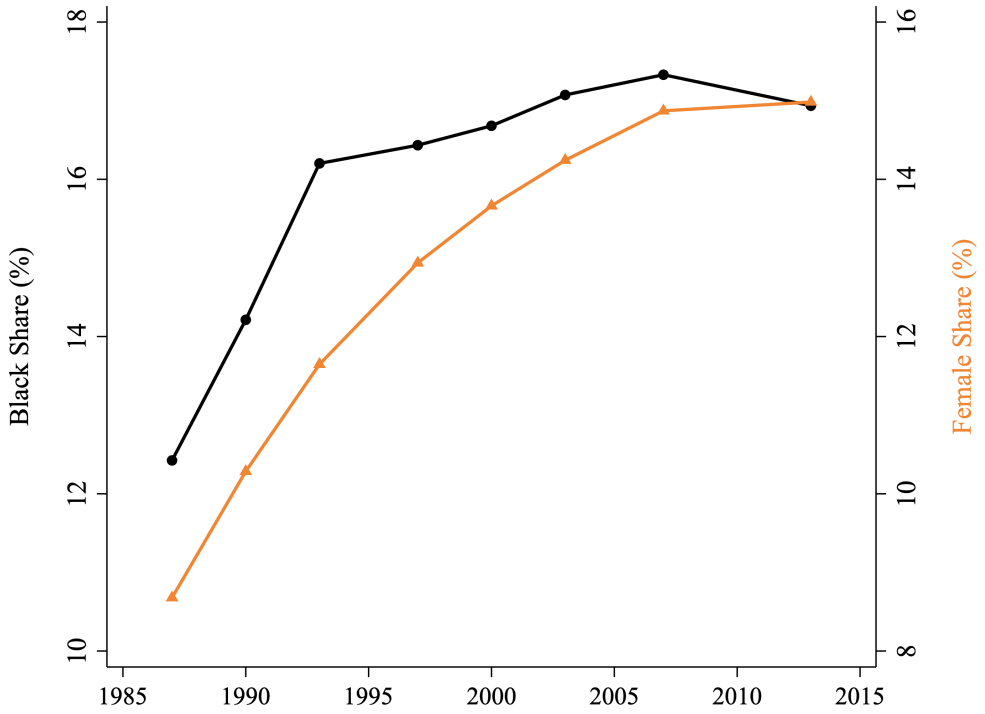
(d) White Deaths Per Capita



Notes: This figure plots the effect of Affirmative Action litigation on homicide victimization. The outcome in each panel contains all homicides for each respective racial group, combining homicides and policing-related civilian fatalities of civilians. In panel (a) the dependent variable is all non-White homicides and in panel (b) is the victimization rates per 100,000 non-White residents. Similarly, In panel (c) the dependent variable is the number of all White homicides and in panel (b) it is the policing-related civilian fatalities per 100,000 White residents. Each regression includes county, region-by-year, and urban-by-year fixed effects. Robust standard errors are clustered at the county level, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

B.5 Diversity vs. Intervention

Figure B30: LEMAS Police Composition Over Time

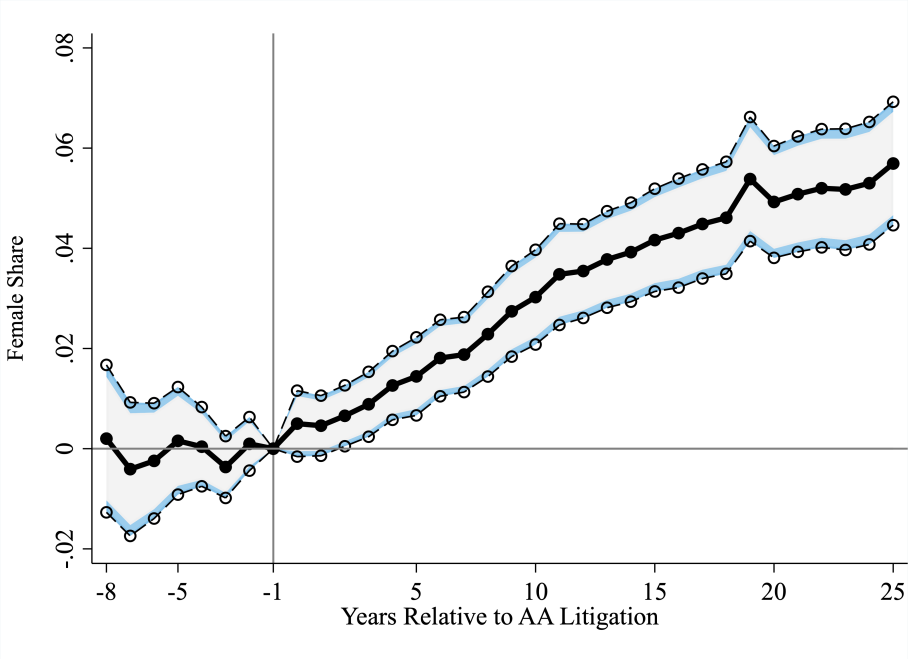


Notes: This figure plots the share of police departmetns that are Black and female, respectively, over time. Data comes from the Law Enforcement Management Statistics (LEMAS), publicly available at the ICPSR.

Figure B31: Event Study Results – Employment

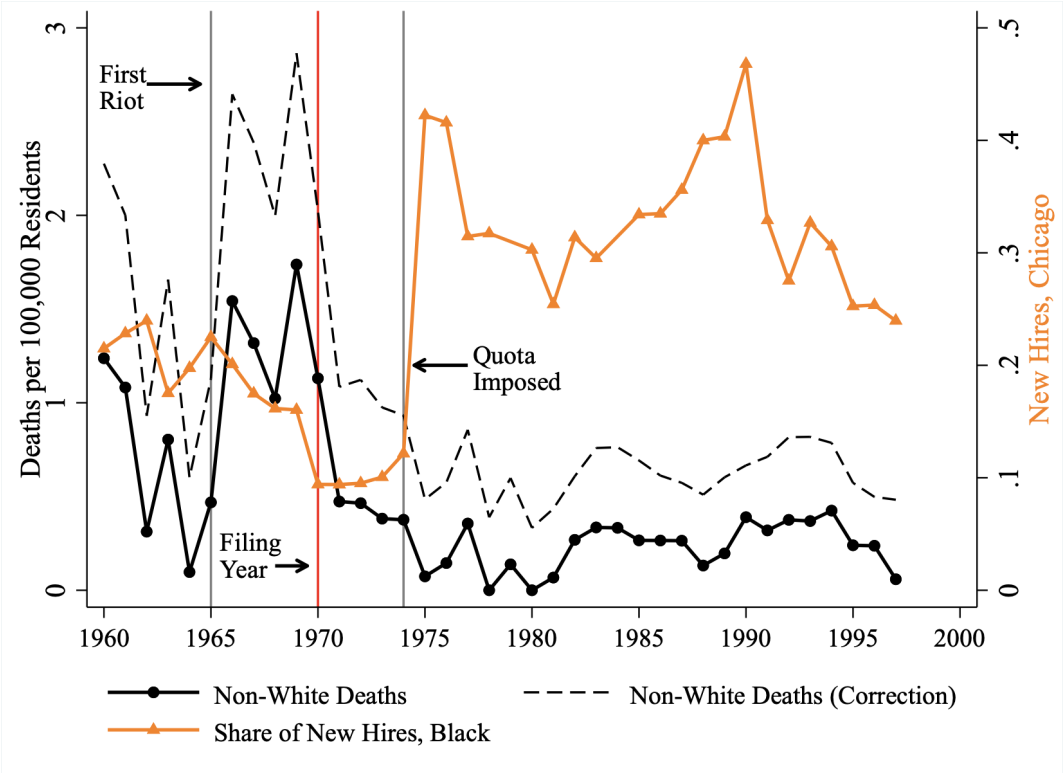
(a) Employment

(b) Female Share



Notes: This figure plots the effect of Affirmative Action litigation on police employment. The outcome in panel (a) is the number of police officers per 1,000 residents and in panel (b) it is the share of female officers. Each regression includes county and region-by-year fixed effects. Data comes from the UCR Law Enforcement Killed or Assaulted files. Employment by gender is available after 1971–therefore the results presented are from an unbalanced panel. Robust standard errors are clustered by county, and 95 and 90 percent confidence intervals are presented. The horizontal axis represents event-years (years before and after litigation).

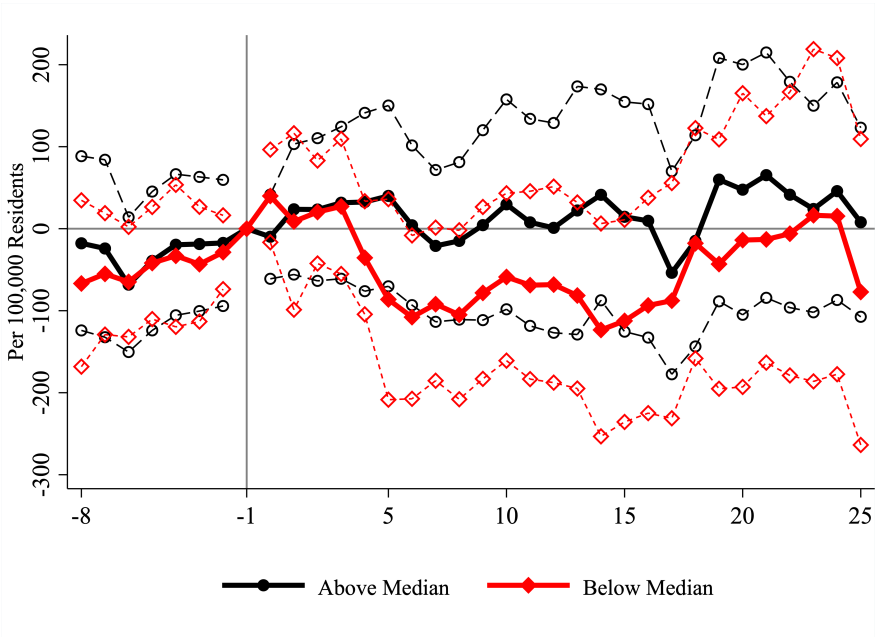
Figure B32: Cook County Police Killings Over Time, Estimated



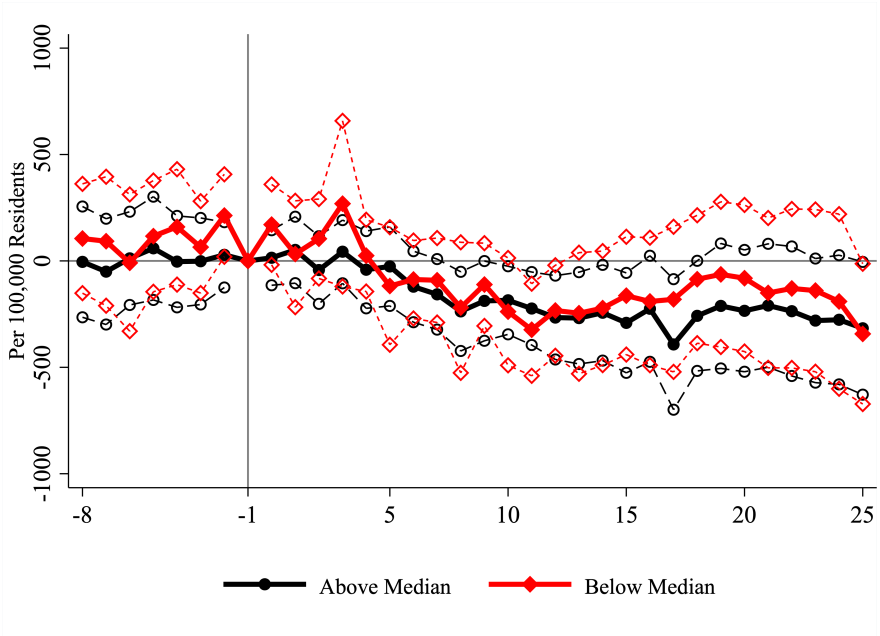
Notes: Estimated police killings displayed by dashed line. We assume that each county's contribution to the state level estimate of police killings is constant over time. The new county level measure of police killings reflects reported deaths and the bias attributed to each county contribution of the state's total.

Figure B33: Event Study Estimates – Non-White Arrests by Police Representation Gap

(a) Violent Crime Arrests



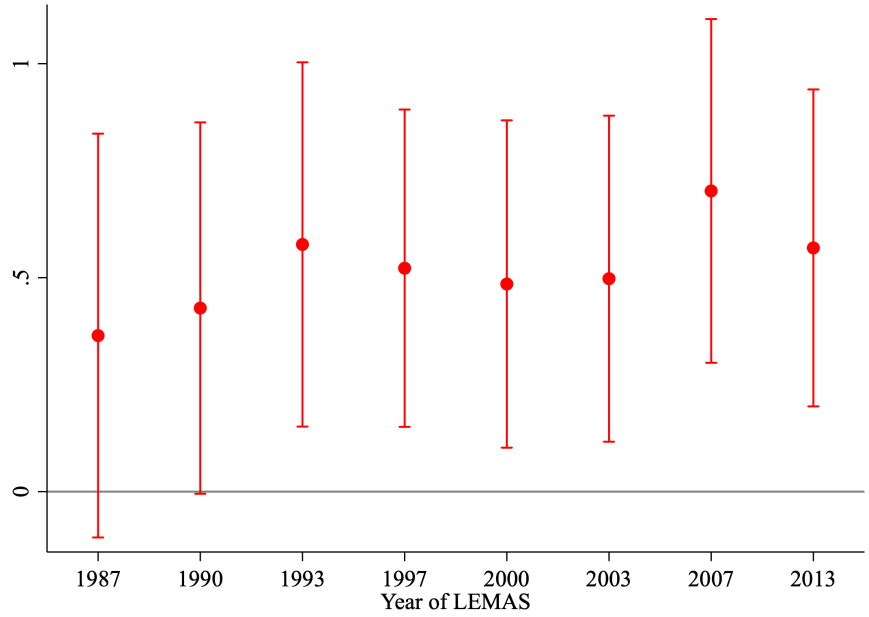
(b) Property Crime Arrests



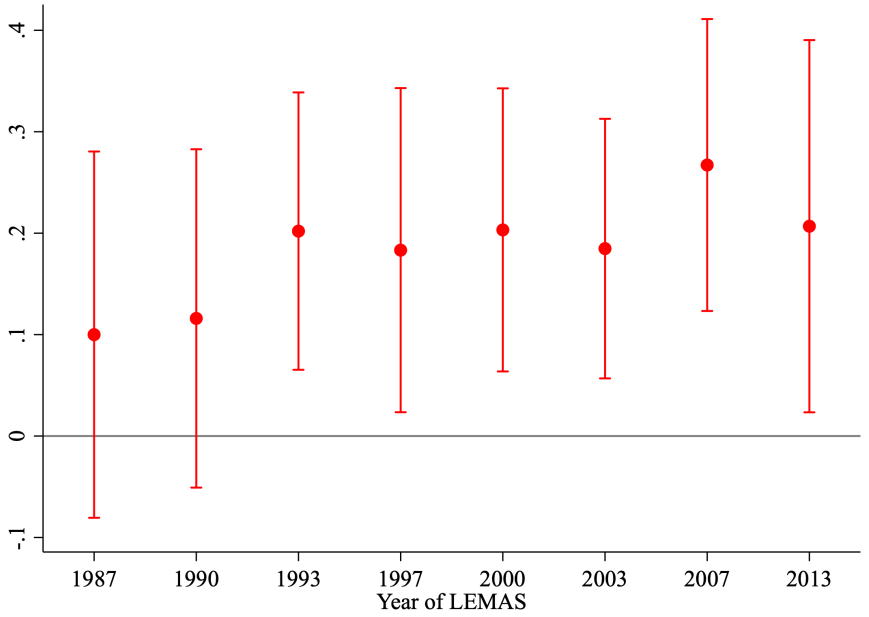
Notes: This figure plots the effect of Affirmative Action litigation on arrests by representation gap. Panel (a) report total non-White arrests for index 1 crimes (violent and property crimes) while panels (b) and (c) report non-White arrests by crime type (violent or property). Panel (b) report quality of life arrest (public intoxication, liquor violations, disorderly conduct, gambling, suspicious behavior, vandalism, and vagrancy). The sample is limited to counties identified by McCrary (2007) and report at least 52 years between 1960 and 2016. Treatment group corresponds to counties with cities treated prior to 1987. Each specification includes county and region-by-year fixed effects. Estimates of the share of Black officers and the representation gap, 25 years after treatment, are calculated using data from data from LEMAS. Robust standard errors are clustered by county. The horizontal axis represents event-years (years before and after litigation).

Figure B34: Years Since Litigation and Police Composition

(a) Black Share

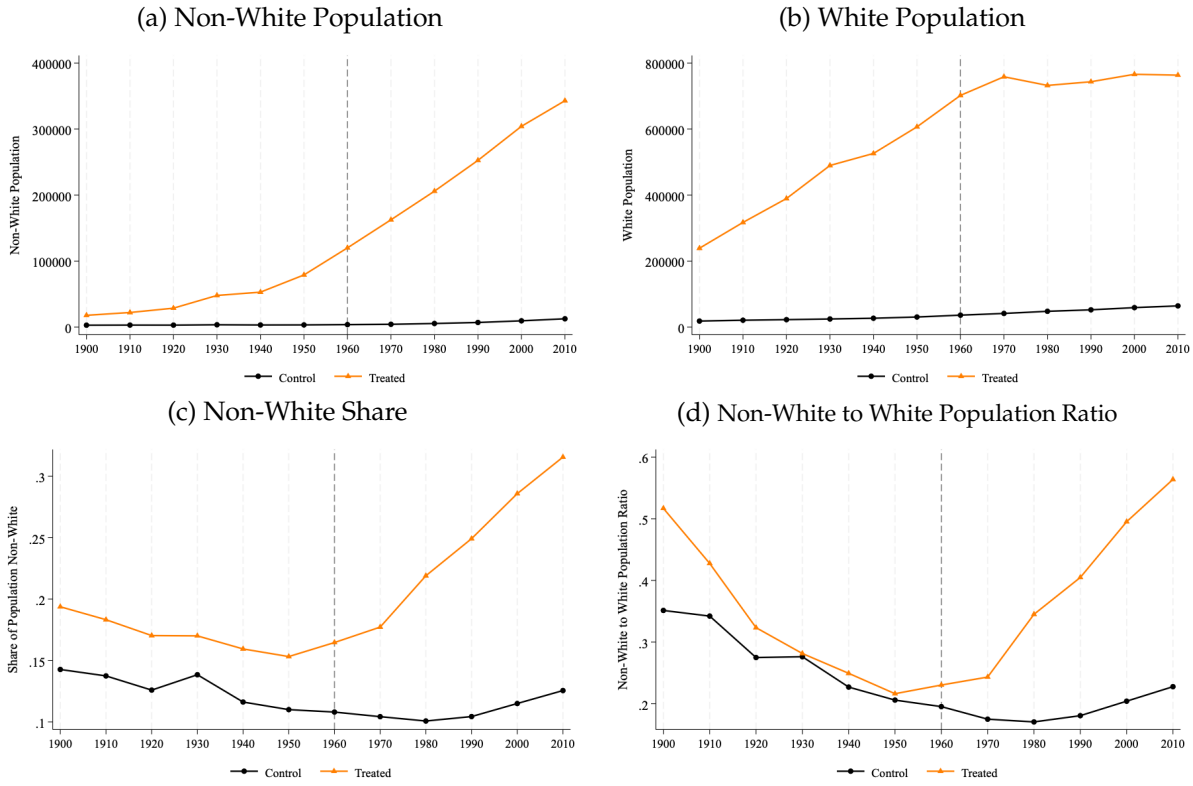


(b) Female Share



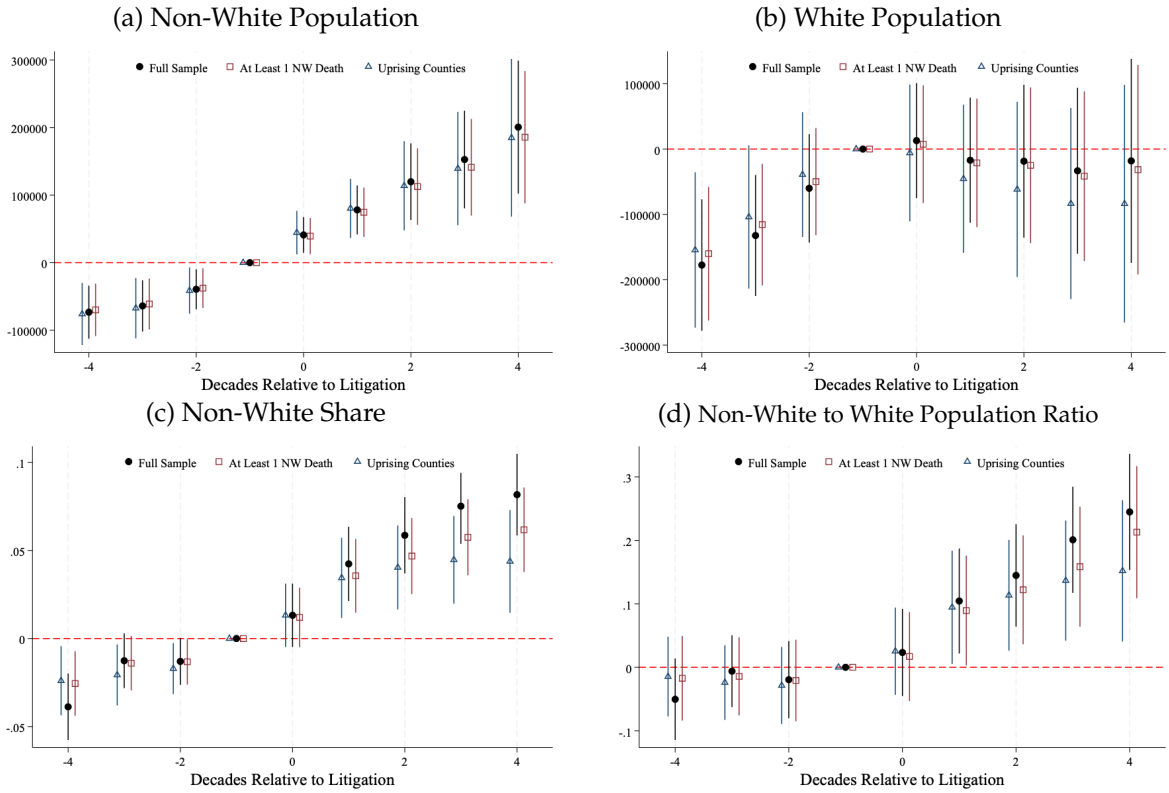
Notes: This figure plots coefficient estimate of the effect of years since a county experiences Affirmative Action litigation on the share of Black or female officers, respectively. Each coefficient is from a separate cross-sectional regression that includes region fixed effects. Data comes from the Law Enforcement Management Statistics (LEMAS).

Figure B35: Demographic Trends: 1900-2010



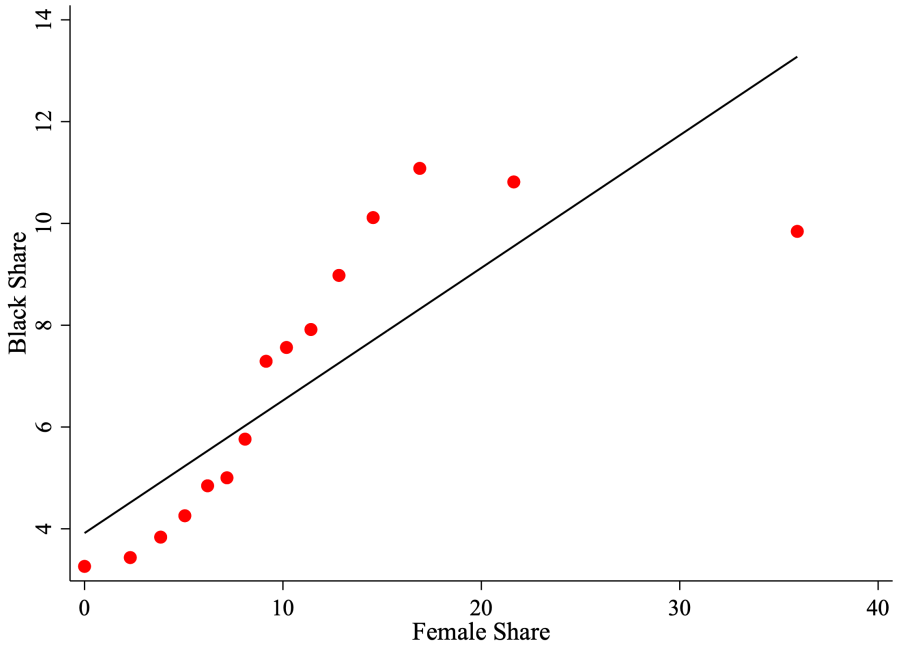
Notes: This figure plots population changes for each Census year from 1900-2010. Data for historical population counts by county comes from Haines et al. (2010). Population counts for 1970 to 2010 comes from Surveillance, Epidemiology, and End Results (SEER) annual data.

Figure B36: Demographic Trends: 1900-2010



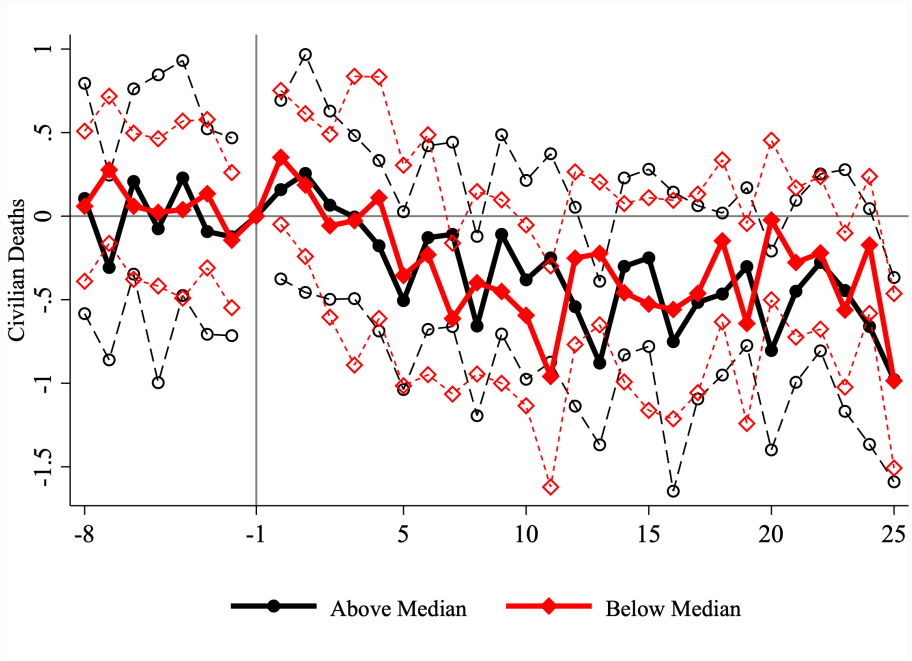
Notes: This figure plots the effect of Affirmative Action litigation on population changes. Each regression includes county, region-by-year, and, urban-by-year fixed effects. Data for historical population counts by county comes from Haines et al. (2010). Population counts for 1970 to 2010 comes from Surveillance, Epidemiology, and End Results (SEER) annual data.

Figure B37: Correlation Between the Share of Black and Female Officers



Notes: This figure is a binned scatter plot of the share of Black and female officers. Data comes from the Law Enforcement Management Statistics (LEMAS).

Figure B38: Event-Study Results - Female Police Representation



Notes: This figure stratifies the treated groups to those with litigated police departments with above (Black line with circle markers) and below (red line with square markers) median level of female representation in the first year that this data is available for a police department (starting in 1971). The dependent variable is non-White policing-related civilian fatalities. This specification includes county and region-by-year fixed effects. The horizontal axis represents event-years (years before and after litigation).

C Appendix Tables

Table C1: Event Study - Joint Effects – White Deaths

	(1) OLS	(2) Poisson	(3) WLS	(4) CS Estimator
Pre-Period Effect (Event Years -8 to -2)	-0.264	-0.252	-0.0369	-0.260
	[0.222]	[0.173]	[0.0253]	[0.232]
Shorter-Run Effect (Event Years 0 to 7)	-0.453**	-0.487***	-0.0655***	-0.458
	[0.215]	[0.174]	[0.0242]	[0.222]
Medium-Run Effect (Event Years 8 to 15)	-0.374	-0.384**	-0.0631***	-0.386*
	[0.263]	[0.179]	[0.0241]	[0.275]
Longer-Run Effect (Event Years 16 to 25)	-0.495*	-0.662***	-0.0822***	-0.515**
	[0.268]	[0.176]	[0.0251]	[0.286]
Mean DV	1.171	1.171	0.146	1.171
Number of Counties	2,980	2,980	2,707	2,980

Note: This sample removes counties treated after 1987 from the analysis. All regressions include county and region-by-year fixed effects. Columns (1) and (2) use then number of deaths as the dependent variable while column (3) uses the age-adjusted mortality rate. Column (1) includes the log of non-White population as an independent variable. Column (2) accounts for exposure with non-White population. Column (3) uses 1960 non-White population as weights. Column (4) reports estimates using Callaway and Sant'Anna (2020) estimator – analogous to column (1) OLS two-way fixed effects model. Robust standard errors are clustered by county. *** p<.01, ** p<.05, * p<0.1.

Table C2: Joint Effects - Counties With At Least One Non-White Death

	(1)	(2)	(3)
		Non-White	
	OLS	Poisson	WLS
Pre-Period Effect (Event Years -8 to -2)	-0.0715	-0.00501	0.0121
	[0.231]	[0.149]	[0.0948]
Shorter-Run Effect (Event Years 0 to 7)	-0.161	-0.0502	-0.0246
	[0.263]	[0.149]	[0.111]
Medium-Run Effect (Event Years 8 to 15)	-0.591*	-0.413***	-0.216*
	[0.322]	[0.149]	[0.116]
Longer-Run Effect (Event Years 16 to 25)	-0.683**	-0.508***	-0.238*
	[0.319]	[0.148]	[0.126]
Number of Counties	844	844	784
Mean DV	1.478	1.478	0.414

Note: This sample includes counties that report at least one non-White death over the sample period. All regressions include county and region-by-year fixed effects. Columns (1) and (2) use then number of deaths as the dependent variable while column (3) uses the age-adjusted mortality rate. Column (1) includes the log of non-White population as an independent variable. Column (2) accounts for exposure with non-White population. Column (3) uses 1960 non-White population as weights. Column (4) reports estimates using Callaway and Sant'Anna (2020) estimator – analogous to column (1) OLS two-way fixed effects model. Robust standard errors are clustered by county. *** p<.01, ** p<.05, * p<0.1

Table C3: Joint Effects - Counties With At Least One Non-White Death in the 1960s

	(1)	(2)	(3)
		Non-White	
	OLS	Poisson	WLS
Pre-Period Effect (Event Years -8 to -2)	-0.117	-0.127	-0.0693
	[0.293]	[0.147]	[0.101]
Shorter-Run Effect (Event Years 0 to 7)	-0.214	-0.0625	-0.0359
	[0.320]	[0.144]	[0.107]
Medium-Run Effect (Event Years 8 to 15)	-0.707*	-0.406***	-0.222*
	[0.391]	[0.144]	[0.116]
Longer-Run Effect (Event Years 16 to 25)	-0.798**	-0.470***	-0.237*
	[0.385]	[0.145]	[0.133]
Number of Counties	369	369	350
Mean DV	1.800	1.800	0.462

Note: This sample includes counties that report at least one non-White death prior to 1969. All regressions include county and region-by-year fixed effects. Columns (1) and (2) use then number of deaths as the dependent variable while column (3) uses the age-adjusted mortality rate. Column (1) includes the log of non-White population as an independent variable. Column (2) accounts for exposure with non-White population. Column (3) uses 1960 non-White population as weights. Column (4) reports estimates using Callaway and Sant'Anna (2020) estimator – analogous to column (1) OLS two-way fixed effects model. Robust standard errors are clustered by county. *** p<.01, ** p<.05, * p<0.1

Table C4: Data Problems: Year 1972

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Non-White				White		
	Original	Drop 1972	Interpolate 1972	Drop if Treated in 1972	Original	Drop 1972	Interpolate 1972	Drop if Treated in 1972
Pre-Period Effect (Event Years -8 to -2)	-0.00501	-0.108	-0.0568	-0.0292	-0.252	-0.231	-0.212	-0.159
	[0.149]	[0.142]	[0.130]	[0.167]	[0.173]	[0.181]	[0.161]	[0.225]
Shorter-Run Effect (Event Years 0 to 7)	-0.0502	-0.151	-0.0633	-0.00905	-0.487***	-0.474***	-0.394**	-0.419*
	[0.149]	[0.144]	[0.134]	[0.169]	[0.174]	[0.182]	[0.163]	[0.224]
Medium-Run Effect (Event Years 8 to 15)	-0.413***	-0.506***	-0.436***	-0.322*	-0.384**	-0.377**	-0.330**	-0.340
	[0.149]	[0.141]	[0.133]	[0.168]	[0.179]	[0.187]	[0.168]	[0.227]
Longer-Run Effect (Event Years 16 to 25)	-0.508***	-0.599***	-0.532***	-0.365**	-0.662***	-0.656***	-0.608***	-0.640***
	[0.148]	[0.140]	[0.132]	[0.165]	[0.176]	[0.184]	[0.165]	[0.224]
Observations	44,096	42,900	43,725	43,725	88,139	86,320	87,980	87,768
Mean DV	1.457	1.593	1.657	1.238	1.171	1.254	1.214	0.905

Note: All panels' estimates are from Poisson models that include county and region-by-year fixed effects and account for exposure with population by race. Columns (1) and (5) report the original estimates for non-White and White police killings of civilians, respectively. Columns (2) and (6) drop the year 1972, columns (3) and (7) use the average number of deaths in 1971 and 1973 to estimate the number of deaths in 1972, and columns (4) and (8) drop locations treated in 1972. Robust standard errors are clustered by county. *** p<.01, ** p<.05, * p<0.1.

Table C5: Summary Statistics Accounting for Key Cross-Sectional Differences

1960 Characteristics	(1) Overall	(2) Treatment	(3) Control	(4) T-Test of Difference	(5) Reweighted Group	(6) T-Test of Difference
Population	59,431	815,406	39,947	<0.01	462,850	<0.05
Population per square mile	165.74	3,214.55	87.16	<0.01	1,415.17	.1
% of counties that experienced uprisings	0.09	0.84	0.07	<0.01	0.84	.99
Percentage of the Population						
residing in urban areas	32.50	87.21	31.09	<0.01	86.62	.8
w/ 12 or more years of education	36.45	43.41	36.27	<0.01	38.94	.07
w/ income greater than 10K	7.92	16.95	7.69	<0.01	14.74	.19
w/ income less than 3K	35.62	17.95	36.07	<0.01	18.19	.86
non-White	10.94	16.61	10.80	<0.01	12.34	.25
Deaths Due to Legal Intervention						
White	0.04	0.61	0.03	<0.01	0.09	<0.05
non-White	0.04	1.03	0.01	<0.01	0.17	<0.01
Number of Counties	2,985	75	2,910		2,910	
joint F-test				3.58		1.84
p-value				<0.01		.06

Note: This table reports summary statistics. Treatment refers to counties with a police department that experiences Affirmative Action litigation, and control refers to all other counties. Column (5) reweigh counties in the control group by their inverse propensity scores and column (6) reports the p-value associated with the $t-$test difference in means between groups.

Table C6: Composition of Police Department, 1987-2013

	(1) Black Share	(2) Female Share
Treated in 1970s	0.131***	0.0511***
	[0.0141]	[0.00477]
Treated in 1980s	0.118***	0.0383**
	[0.0454]	[0.0149]
Treated in 1990s	0.0393	0.0112***
	[0.0245]	[0.00331]
Observations	9,267	9,267
R-squared	0.179	0.036
Mean DV	0.0550	0.0790

Note: This table reports the regression estimates of share of sworn police officers by demographic group relative to the timing of treatment. The reference group consists of counties that were never treated. All columns include region and year fixed effects. Mean dependent variables report the average share of officers for the control group. Robust standard errors are clustered at the county level and presented in brackets. Source: Law Enforcement and Administrative Statistics (LEMAS), 1987,1990, 1993, 1997, 2000, 2003, 2007, and 2013.

Table C7: Composition of Police Department, Years Since Litigation

	(1) Black Share	(2) Female Share
Years Since Litigation	0.00437**	0.00159**
	[0.00172]	[0.000675]
Observations	564	564
R-squared	0.276	0.194
Mean DV	0.0810	0.0820

Note: This table reports the regression estimates of share of sworn police officers by demographic group relative to timing of treatment. All columns include region and year fixed effects. Mean dependent variables report average share of officers for the initial year of litigation. Robust standard errors are clustered at the county level and presented in brackets. Source: Law Enforcement and Administrative Statistics (LEMAS), 1987,1990, 1993, 1997, 2000, 2003, 2007, and 2013.

Table C8: Service Calls by Treatment Status in 1987

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Service Calls Per Capita			Operational 911 System		
Treatment	0.0509	0.00866	0.128	-0.0615	-0.107	0.0627
	[0.105]	[0.109]	[0.192]	[0.0782]	[0.0732]	[0.196]
Add Covariates		X			X	
Treated Only			X			X
Observations	211	211	68	211	211	68
R-squared	0.147	0.242	0.233	0.113	0.189	0.256

Note: Regression of police services to 1987 treatment status. All columns include the percentage of the population Black (1980) and region fixed effects. The dependent variable in columns (1) - (3) is service calls per capita and (0/1) operational 911 system in columns (4) - (6). Columns (2) and (5) adds the following covariates from the 1980 census: population per square mile, median age, median income, percentage of the population age 5 and under, percentage of the population age 64 and older, percent of the population with 12 or more years of education, and the percent of households headed by a female. Columns (3) and (6) restrict the sample to cities litigated between 1969 and 2000. Robust standard errors presented in brackets. Source: Law Enforcement and Administrative Statistics (LEMAS), 1987.

Table C9: Synthetic Diff-in-Diff

	(1) Non-White	(2) White
Treatment	-0.258***	-0.019
	[0.087]	[0.015]
Observations	157940	157940
Mean DV	0.383	.143

Note: This table reports the estimates from a synthetic difference-in-differences design proposed by (Arkhangelsky et al., 2021). The outcome is the age-adjusted mortality rate of police killings for each racial group, respectively. Robust standard errors are clustered at the county level and presented in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$