Racial Profiling and Use of Force in Police Stops: How Local Events Trigger Periods of Increased Discrimination[1]

Joscha Legewie
Yale University

Racial profiling and the disproportionate use of police force are controversial political issues. I argue that racial bias in the use of force increases after relevant events such as the shooting of a police officer by a black suspect. To examine this argument, I design a quasi experiment using data from 3.9 million time and geocoded pedestrian stops in New York City. The findings show that two fatal shootings of police officers by black suspects increased the use of police force against blacks substantially in the days after the shootings. The use of force against whites and Hispanics, however, remained unchanged, and there is no evidence for an effect of two other police murders by a white and Hispanic suspect. Aside from the importance for the debate on racial profiling and police use of force, this research reveals a general set of processes where events create intergroup conflict, foreground stereotypes, and trigger discriminatory responses.

Urban ethnographies vividly describe how aggressive policing targets young black men in poor urban communities (Anderson 1990; Venkatesh 2002; Bourgois 2003; Rios 2011; Stuart 2011; Goffman 2014). Based on extensive fieldwork in the late 1970s and 1980s, Elijah Anderson's *Streetwise* chroni-

[1] This project was supported by the Russell Sage Foundation (grant no. 87-14-03). I thank Delia Baldassarri, Peter Bearman, David Brady, Tom DiPrete, Jeffrey Fagan, Anne Nassauer, Merlin Schaeffer, Patrick Sharkey, and Mathijs de Vaan for helpful comments and discussions. Direct correspondence to Joscha Legewie, Yale University, Department of Sociology, PO Box 208265, New Haven, Connecticut 06520. E-mail: joscha.legewie@yale.edu

0002-9602/2016/12202-0002$10.00

cles the perils faced by black men of constant police presence including stops, harassment, and even arrests whether or not they had committed a crime (Anderson 1990, chap. 7). Fast-forward to the early 21st century, and debates on racial bias in policing continue to reverberate across the country, making national headlines. In New York City, the controversial "stop, question, and frisk" (SQF) policy was endorsed by some as essential for reducing crime rates (MacDonald 2001) and challenged by others as racially biased with a heavy burden placed on affected individuals and communities (Fagan et al. 2010). Previous research has documented racial bias in various areas of policing including racial profiling in pedestrian and vehicle stops, the use of police force, and even an officer's decision to shoot black and white criminal suspects in computer simulations (Geller and Toch 1996*b*; Plant and Peruche 2005; Fagan, Conyers, and Ayres 2014). Yet research on the causal dynamics underlying racial profiling, the excessive use of police force, and other forms of discrimination has largely ignored an important aspect of the social environment in which discriminatory behavior occurs: the temporal embeddedness of social interactions and the potential importance of events. Whereas much research has examined the role of place for racial bias in policing and police behavior more broadly, a temporal perspective considers the role that events play in altering racial disparities. The focus on events helps us to extend previous work from *where* discrimination takes place to *when* it happens.

This study examines how acts of extreme violence towards law enforcement affect the subsequent police treatment of residents. Building on conflict and racial threat theories, I argue that racial bias in policing and discrimination more broadly is not static but fluctuates, partly driven by significant events that provoke intergroup conflict and foreground racial stereotypes. Events strengthen cohesion within the police department and invoke the notion of the police versus black youth. Police increase the use of force against minority groups to mitigate (perceived) threat, retaliate against the offending group, and preserve social order. But the increase might not necessarily be based on a conscious response to the event. Events might also foreground *implicit* racial stereotypes that portray blacks as violent and increase concerns about personal safety among officers. In contrast to previous work on minority threat, the focus on events elaborates a mechanism that might explain short-term (and potentially long-term) changes in the rates of discriminatory behavior.

To study the role of events for racial bias, I design a quasi experiment that examines the effect of events on the use of police force against racial minorities. The design is based on 3.9 million time and geocoded police stops of pedestrians in New York City. The focus is not on incidents of extreme police brutality. Instead, the data provide a unique lens on millions of everyday police-citizen interactions and the potentially disproportionate use of police

force. The design compares similar stops before and right after four acts of extreme violence against police officers. In particular, I match the stops in the two weeks after the events to similar stops before the events (i.e., same location and time of day) and use these matched stops to construct a counterfactual trend that captures what would have happened in the absence of the event. The findings show that, relative to similar stops before two shootings of New York City Police Department (NYPD) police officers by black suspects in 2007 and 2011, the use of physical force by police officers against blacks increased substantially by 16.0% and 13.3%, respectively, in the days after the shootings. The use of force against whites and Hispanics, however, remained unchanged, and I found no evidence for an effect of two police murders by a white and a Hispanic suspect. Complementary estimates based on an regression discontinuity (RD) design reaffirm the findings, and a simulation of placebo treatments provides further support. The race-specific nature of the response whether driven by perceptions of threat, concerns about officer safety, or implicit stereotypes indicates racial bias.

The findings provide quasi-experimental evidence showing that incidents of extreme violence against police officers can lead to periods of substantially increased use of force against African-Americans but not against other groups. The effect sizes are likely conservative considering that the analyses are based on police-reported data. The temporal duration of the effect is modest, but the relatively frequent nature of the events makes the consequences profound particularly at a time of intense tensions between the police and black communities. The interpretation of my findings extends beyond acts of extreme violence against police officers. The findings reveal a general set of processes where local events trigger discriminatory responses both with the police and with other actors who might engage in discriminatory behavior (employers, landlords, teachers, etc.). From this perspective, discriminatory behavior arises not only from static conditions but also from temporal sequences of events and responses. This argument contributes to a small but growing line of research in criminology that uses events to better understand the mechanisms that trigger the diffusion of intergroup violence. Aside from these broader theoretical contributions, the quasi-experimental design introduces a new method to measure bias based on sudden changes in the rates of certain behavior.

POLICING THE URBAN POOR

Minority Threat and the Use of Police Force

Previous research has documented racial and ethnic disparities in many areas of criminal justice including police-citizen encounters such as SQF operations, the use of police force, arrests, sentencing, and imprisonment

(Sampson and Lauritsen 1997; Cole 2000; Pettit and Western 2004). Many of these differences can be attributed to the higher involvement of African-Americans in criminal offending (Sampson and Lauritsen 1997). But research also points to racial profiling and bias as a possible explanation. Racial profiling is a form of discrimination by which law enforcement uses a person's race or ethnicity as a key reason to engage in various forms of enforcement. Profiling violates basic human rights, undermines trust in public institutions, and has severe consequences for the victims and for society at large (Weitzer and Tuch 2002; Lundman and Kaufman 2003; Thompson and Lee 2004; Mays, Cochran, and Barnes 2007; Gee and Ponce 2010; Geller et al. 2014; Tyler, Fagan, and Geller 2014). Since the late 1990s, these concerns have been popularized under the notion "driving while black" (Harris 1999; Weitzer 2000; Lundman and Kaufman 2003; Antonovics and Knight 2009). Along these lines, studies have documented racial bias in pedestrian and vehicle stops (Fagan, Zimring, and Kim 1997; Fagan and Davies 2000; Fagan 2002; Antonovics and Knight 2009), the use of police force (Skolnick and Fyfe 1993; Geller and Toch 1996*b*; Schuck 2004), and even an officer's decision to shoot black and white criminal suspects in computer simulations (Plant and Peruche 2005). Gelman, Fagan, and Kiss (2007), for example, compare racial disparities in stop-and-frisk operations across New York police precincts in 1998 and 1999. They find that African-Americans and Hispanics were stopped more often than whites conditional on population shares and race-specific arrest rates as an estimate of criminal offending (see also Fagan et al. 1997; Greene 1999; Fagan and Davies 2000; Fagan 2002; Coviello and Persico 2013). More recent work by Jeffrey Fagan and colleagues (Fagan et al. 2014) confirms this pattern for later periods and was a central part of the court case against the NYPD (*Floyd et al. v. City of New York et al.*).[2] The literature on the use of police force similarly finds pronounced differences in the use of force against ethnic minorities (Geller and Toch 1996*b*; Alpert and Dunham 1999; Terrill and Mastrofski 2002; Schuck 2004). The stop-and-frisk data used in this study, for example, reveal that 16.5% of police stops involve some use of force for whites compared with 22.2% for African-Americans (see table 1). Using data from 3,116 police-suspect encounters documented

[2] Some studies also challenge these findings and argue that the higher number of minority stops largely reflects higher crime rates among these groups (Grogger and Ridgeway 2006; Ridgeway 2007; Worden, McLean, and Wheeler 2012). Grogger and Ridgeway (2006), e.g., use an innovative strategy to test for racial profiling in traffic stops. They argue that darkness makes it more difficult to identify the race of a motorist before making a stop. Based on this assumption, the authors compare the race distribution of stops before and after sunset using data from Oakland, California. They find no evidence for racial profiling in vehicle stops.

TABLE 1
Stop-and-Frisk Operations in New York City, 2006–2012

	Whites	Blacks	Hispanics	Overall
Number of stops	384,346	2,046,241	1,215,788	3,782,662
% of stops	10.2	54.1	32.1	100
Population share in New York City	44.6	25.1	27.5	100
Characteristics of stops:				
Male	89.7	93.1	93.4	92.9
Average age (years)	29.3	28.1	27.6	28.1
Police force	16.4	22.2	23.9	22.0
Person frisked	41.4	55.3	56.3	53.8
Person searched	8.3	8.3	9.0	8.5
Person arrested	6.1	5.8	6.0	5.9
Weapon found	1.5	0.9	1.1	1.0
Contraband found	2.2	1.7	1.7	1.7

Note.—Overall includes stops of other racial groups. Data are percentages unless otherwise specified.

as part of an observational study in two urban settings, Terrill and Mastrofski (2002) find similar differences indicating that officers treated nonwhite, poor, and younger suspects more forcefully than others even conditional on their behavior. Other research, however, attributes racial disparities in the use of police force to suspects' behavior and argues that citizen behavior is the leading cause of police treatment (Friedrich 1980; Garner, Maxwell, and Heraux 2002; Durna 2011). While the majority of police-citizen interactions do not involve police force and some controversy remains about the correct explanations, this research indicates that racial bias in pedestrian and vehicle stops and the use of police force at least partly explains the stark racial disparities.

In the search for factors influencing police behavior, sociologists have primarily focused on situational and structural characteristics of police-citizen interactions including the social class, race, and gender of suspects, their demeanor, and the location of encounters (e.g., Smith 1986; Worden 1996; Holmes 2000). Based on broader sociological theories of race relations and group threat (Blumer 1958; Blalock 1967), a prominent account in research on variations in formal levels of social control such as aggressive policing has focused on (perceived) levels of threat posed by racial and ethnic minorities (Turk 1966; Liska 1992; Jacobs and O'Brien 1998; Baumer, Messner, and Rosenfeld 2003; Jacobs, Carmichael, and Kent 2005; Stults and Baumer 2007; Smith and Holmes 2014). This argument focuses on the ways in which dominant or privileged groups use the police, criminal law, and other state instruments to control subordinate groups who threaten their interests. From this perspective, police behavior including coercive crime control mechanisms such as the use of police force partly reflects deeply rooted social divisions that separate dominant and subordinate racial and ethnic groups (Liska

1992; Chambliss 2001; Smith and Holmes 2014). An important empirical prediction from this theoretical perspective is the *minority threat hypothesis*. It suggests that the relative size of minority groups in an area influences the perceived level of political and economic threat among majority members. This increase in perceived threat, in turn, raises support for crime control mechanisms or at least gives the police additional leeway. Increased support or lack of restraint translate to more expenditures for criminal justice, higher arrest and imprisonment rates, and an increase in coercive police behavior such as police use of force (Liska 1992). This argument does not necessarily imply that privileged groups directly demand the use of aggressive policing. The lack of restraint might result in increased use of police force as an efficient way to ensure social control (Jacobs and O'Brien 1998).

The racial threat thesis has generated a large number of empirical studies that examine whether the relative size of the black population is related to different aspects of social control such as police use of deadly force, the size of the police force, arrests, incarceration rates, and so on (Jackson and Carroll 1981; Liska 1992; Jacobs and O'Brien 1998; Eitle, D'Alessio, and Stolzenberg 2002; Jacobs and Carmichael 2002; Stults and Baumer 2007; Smith and Holmes 2014; Legewie and Fagan 2016). A number of studies also focus on police use of nonlethal force. Smith and Holmes (2014), for example, study sustained complaints about the use of excessive police force from multiple cities to show that the proportion of minority residents in a city increases the use of police force. Research looking at the perception of police-based racial discrimination among black youth mirrors these findings. It shows that perception of police misconduct is higher in neighborhoods that experienced a recent in-migration of African-Americans (Stewart et al. 2009).

Other research on police behavior similarly highlights the importance of place without relying on a minority threat argument. Werthman and Piliavin (1967), for example, maintain that neighborhoods influence expectations regarding appropriate behavior and function as an important indicator used by police to identify suspects, which in turn influences police conduct (see also Black's theory of law 1976). Terrill and Reisig (2003), for example, show that the use of police force is substantially higher in disadvantaged neighborhoods and in those with higher homicide rates conditional on suspect resistance and other situational factors (see also Smith 1986; Weitzer 1999; Kane 2002). The NYPD itself pioneered a form of aggressive street policing targeted at high crime areas in the 1990s (Zimring 2013). Reflecting the importance of place for police behavior, the strategy heavily relied on stop-and-frisk operations in hot-spot areas with the goal to prevent crimes by stopping suspicious pedestrians and arresting them for minor offenses (Zimring 2013). For some, the racial disparities in stop rates were simply a consequence of targeting high-crime areas (MacDonald 2001), whereas others (including the court ruling *Floyd et al. v. City of New York et al.*) denounced the strategy

as racially biased (Fagan et al. 2010). Fagan and colleagues (Fagan and Davies 2000; Fagan et al. 2010), for example, show that the racial composition, poverty rate, and other neighborhood characteristics are important predictors of race- and crime-specific stops even net of local crime rates and physical disorder.

Resonating with the vivid ethnographic descriptions of aggressive policing in poor black communities (Rios 2011; Stuart 2011; Goffman 2014), the racial threat thesis and other research on policing highlight the importance of place for police behavior. From this perspective, neighborhoods are a critical situational factor that influences all forms of police behavior including racial bias and misconduct. Yet research on the causal dynamics underlying racial profiling, the excessive use of police force, and other forms of discrimination has largely ignored an important aspect of the social environment in which discriminatory behavior occurs: the temporal embeddedness of social interactions and the potential importance of events. While prior work has examined *where* discrimination takes place, we know little about *when* it happens.

Events and the Use of Police Force against Minority Groups

Building on conflict and racial threat theories, I argue that relevant events increase the use of police force against minority groups through their effect on intergroup conflict, perceptions of minority threat, racial stereotypes, and concerns about personal safety among officers. Significant crime events such as the shooting of a police officer increase internal group cohesion within the police department. The events evoke a long history of tensions between the police and black communities including a time of high-profile murders of police officers around the country. They create and sustain the notion of the police versus black youth and increase the perceived level of threat both in the general public and among officers. Attacks against officers stir anger and emotions in the police community. Reports about the aftermath of officer killings in New York City and other places highlight the emotional nature of the event including gatherings around the home of killed officers, ceremonial funerals attracting thousands of officers, hospital visits, and public denunciations of the vicious nature of the crime.[3] Portes and Sensenbrenner's (1993) concept of "bounded solidarity" reinforces this argument. Bounded solidarity emerges "out of the situated reaction of a class of people faced with common adversities. . . . It is limited to members of a particular

[3] These ceremonies are documented in news coverage of police officer killings. For example, these *New York Times* articles cover the four events examined in this study: http://nyti.ms/1GIzJLB, http://nyti.ms/1HWaN3M, http://nyti.ms/1JVRokN, and http://nyti.ms/1S9y7D2 (all accessed on June 30, 2015).

group who find themselves affected by common events in a particular time and place" (Portes and Sensenbrenner 1993, p. 1325). Research in criminology similarly highlights the importance of group solidarity and links solidarity to intergroup violence. Donald Black's influential work on violence as social control (Black 1983) for example, describes interpersonal violence in terms of "self-help." Violent acts "express a grievance by one person or group against another" (Black 1983, p. 34) and are an effective way to assert social control. They are partly motivated by collective responsibility or liability whereby all group members are accountable for the conduct of their peers. Gang researchers similarly identify "generalized violence" as violent acts targeting group members that were not part of the initial dispute (Jacobs 2004; Papachristos 2009). Papachristos (2009) argues that gang murders are embedded in structured social relations and are part of a status struggle between groups independent of individual actors. Acts of violence against group members are perceived as a threat. They strengthen internal cohesion, solidify group identity, and elicit a violent response (Papachristos 2009). This research highlights how events can trigger cascades of violent acts that are driven by intergroup conflict and diffuse through networks. While Black's and Papachristos's theories largely focus on violence in situations that provide no legal resort, their arguments about collective liability and generalized violence reinforce the idea that acts of violence against police officers strengthen police solidarity and foreground group tensions between the police and black communities. The police respond by acting aggressively against the offending group and may temporarily alter or even permanently increase the use of force against minorities. Increasing the use of force ensures social control, (re)asserts authority, and retaliates against the offending group. In his ethnographic study of Chicago's Little Village neighborhood, Vargas (2016) documents this police use of street justice to get back at residents who disrespected or assaulted fellow officers. Along similar lines, Fassin (2013) reports about police engaging in retaliatory violence in the suburbs of Paris.[4]

But an increase in the use of force might not necessarily be based on a conscious or even retaliatory response to the event. Events might also foreground *implicit* racial stereotypes that portray blacks as violent and increase concerns about personal safety among officers. The priming hypothesis at the core of this argument states that racial cues such as relevant crime events can activate or deactivate often implicit racial predispositions (Gilliam and Iyengar 2000; Blair 2002; Fazio and Olson 2003; Quillian 2006;

[4] An *AJS* reviewer pointed to a similar situation from his/her own ethnographic research. This reviewer described how a group of officers visited the home of a fallen officer's partner to comfort him and talked about retaliating against gang members on the street.

Weisbuch, Pauker, and Ambady 2009). As a consequence, police officers might be more likely to interpret identical actions—implicitly or explicitly—as hostile or threatening and react to behavior that went unnoticed or without consequences *before* a certain event with the use of physical force *after* the event. This response builds on preexisting stereotypes and prejudices of African-Americans as violent criminals. Particularly in the United States, the public directly links crime to race and conflates violence with African-Americans. Stereotypes are widespread not only in the general public (Devine and Elliot 1995; Russell-Brown 1998; Quillian and Pager 2001) but also among police officers (Welch 2007).[5] Relevant events foreground these stereotypes and increase concern about personal safety among officers. The broader literature on the effect of events provides support for the idea that events can shape these stereotypes and prejudices. Using a quasi experiment from nine European countries, Legewie (2013) shows that terrorist attacks can have a profound short-term effect on citizens' perception of immigrants and argues that such a response is partly driven by increased perceptions of threat that are mediated by the social environment. Similarly, the literature on priming effects documents that exposure to news reports about violent crimes with an alleged black perpetrator increases negative attitudes about blacks (Gilliam and Iyengar 2000) and induces fear (Peffley, Shields, and Williams 1996). Such priming effects have been documented using different methods, settings, and samples (Mendelberg 2008) including police officers (Eberhardt et al. 2004; Graham and Lowery 2004).

My argument highlights the temporal embeddedness of police-citizen interactions and how events can increase perceptions of threat and trigger intergroup conflict. The focus on the temporal nature of events points to a mechanism for short-term (and potentially long-term) changes in intergroup relations and the rates of discriminatory behavior.[6] It builds on and ad-

[5] Other studies have documented similar, although less pronounced, stereotypes of Hispanics, portraying them as violence-prone street gang members (Carnevale and Stone 1995; Portillos, Mann, and Zatz 1998; Bender 2003). Nonetheless, surveys show that blacks are consistently rated as more prone to violence than any other racial or ethnic group, presumably building on a long history of stereotyped perceptions (Smith 1991).

[6] While it is difficult to determine the temporal duration of the response to events on theoretical grounds alone, previous studies might be a helpful reference. The findings from laboratory priming experiments show short-lived effects that disappear quickly after exposure to the stimulus. Some survey-based panel studies, however, suggest that priming effects persist over an extended period lasting several month (Althaus and Kim 2006). Other research indicates that the effect of major events on attitudes toward out-groups (Legewie 2013) or hate crimes (Disha, Cavendish, and King 2011; King and Sutton 2013) is intense but modest in duration that nonetheless goes far beyond the minutes or hours documented in some laboratory experiments. In the end, the temporal duration largely remains an important empirical question that is essential for our understanding of the role of events.

vances a small but growing line of research that uses events to better understand the mechanism that triggers the diffusion of intergroup conflict and violence (Papachristos 2009; Vargas 2014). The argument helps us to pinpoint not only *where* discrimination takes place but also *when* it happens. The analysis in this article cannot fully test the different possible mechanisms driving the increase in police use of force—retaliation, increase in perceived threat, implicit stereotyping. But the different processes likely work jointly. They all indicate a race-specific response to events that reflects racial bias with a particularly pronounced effect for events involving black suspects. As such, events are an overlooked contextual factor that shapes racial profiling, the police use of force, and discrimination more broadly.

DATA AND METHODS

The analyses presented in this article are based on about 3.9 million time- and geocoded police stops of pedestrians in New York City between 2006 and 2012.[7] The SQF program is regulated by both *Terry* constitutional standards and judicial guidelines based on *People v. DeBour*. The program allows police officers who reasonably suspect that a person has committed, is committing, or is about to commit a felony or a Penal Law misdemeanor to stop and question that person and frisk the person for weapons, if the officer suspects he or she is in danger of physical injury (Ridgeway 2007). Stops are well documented. They are recorded by the officer on the "Stop, Question and Frisk Report Worksheet" (UF-250 form), which includes information on the exact timing, geographical location, the circumstances that led to the stop, the stopped person, and the stop itself such as the use of physical force by the police officer (see appendix D in Ridgeway 2007 for a reproduction). The dependent variable used in the main analysis is a binary indicator for the use of physical force by the police officer in stop-and-frisk operations. Supplementary analyses focus on alternative measures based on a continuum for the use of force (for details, see below).

Over the period in which the police conducted a large number of pedestrian stops, two NYPD police officers were fatally shot by black suspects and three other officers were killed in two separate incidents by a Hispanic and a white suspect. The four events are highly significant incidences for the police with similar rituals in the aftermath. But the events are also markedly different, including one case in which the officer died several days later in the hospital. In the first case involving a black suspect, two police officers pulled over a vehicle with wrong license plates on July 9, 2007. The drivers opened fire as soon as the officers approached the car, fatally injuring

[7] In January 2006, New York City started using a citywide records management system to collect SQF data. Earlier records are not geocoded and lack other information.

Officer Russel Timoshenko and wounding the other. Officer Timoshenko died five days later in the hospital. Even before the officer passed away, the shooting initiated debates about gun control and was compared to the Newhall Incident in April 1970—a deadly shootout between two criminals and officers of the California Highway Patrol.[8] In the second case involving a black suspect, officers responded to a 911 call reporting an ongoing robbery in a residential building on December 12, 2011. When the robbers tried to escape, they were confronted by two more police officers arriving as backup. One of the robbers fired a gun, striking Detective Peter J. Figoski in the face. He died five hours later at Jamaica Hospital Medical Center. The two events involving nonblack suspects are distinct as well. In one case officers responded to a domestic violence call (March 13, 2011), and in the other two auxiliary police officers pursued a suspect after a shooting in a bar (March 14, 2007).[9] In three out of these four cases, the suspect was apprehended or in one case killed within minutes of the incident. In the fourth case, the suspects fled from New York City with the police on their tail.

The fact that these events coincided with a period during which the police conducted a large number of pedestrian stops provides a unique opportunity for a quasi experiment that examines the impact of extreme violence against police officers on racial profiling and the subsequent use of police force against citizens. The different nature of the events has potential implications for the response by the police. The Officer Timoshenko shooting in particular allows me to examine the response to a severe but initially nonlethal attack against police officers.

Estimation Strategy

In the following analysis, I use the events as an exogenous source of variation, together with the timing of the pedestrian stops, to estimate the effect of these events on the use of physical force by police officers. My design is based on two estimation strategies illustrated in figure 1. These strategies overcome many challenges of common regression methods that fail to account

[8] Based on searches of the newspaper database Lexis-Nexis for the officer's last name, news coverage was substantially higher after the initial incident compared with after the actual death. Within the police community, the shooting itself triggered many of the same rituals as in the other three cases, mainly because the condition of the officer was critical and the event itself gruesome and shocking. For these reasons, my analysis focuses on the timing of the shooting, which also allows me to examine the response to a severe but initially nonlethal attack against police officers.

[9] While there are many differences between regular and auxiliary officers at the NYPD, the event itself was gruesome and horrific. The aftermath involved the same rituals as the other incidents including a funeral that attracted thousands of officers uniting the regular and auxiliary police force.

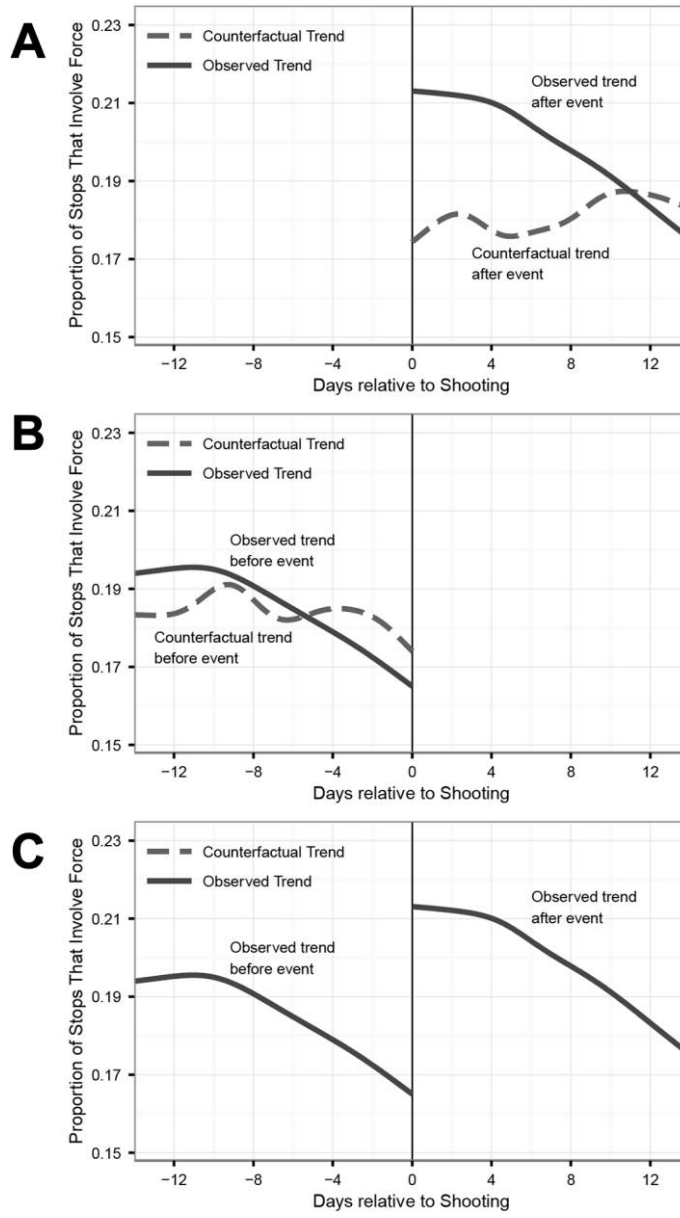


FIG. 1.—Illustration of research design. A, the main analysis compares the observed and counterfactual trend after the event; B, the plausibility of the key assumption can be evaluated with a comparison of the two trends before the event; C, a regression discontinuity design focuses on jumps or discontinuities in the proportion of stops that involve police force right around the event.

for the fact that racial disparities might reflect either racial bias or disproportionate minority crime. In the main analysis, I compare stops right after the shooting with matched stops before the event (fig. 1*A*). In particular, I match the stops after the events to similar stops before the events (i.e., same location and time of day) and use these matched stops as the counterfactual trend. The matching is exact for police precinct and suspect race and uses Mahalanobis distance for other covariates. The covariates include the geographical location of the stop in terms of the *X* and *Y* coordinates, the time of day, and the gender, age, and height of the stopped person. In additional specifications, I also match on the circumstances of the stop (for details on the matching procedure see app. A). The matched stops allow me to construct a counterfactual trend that captures what would have happened in the absence of the events. This matching-based approach compares the behavior of police officers in similar kind of situations before and after the event. It adjusts for factors such as the additional deployment of officers or increased SQF activity in certain neighborhoods as documented by Lacoe and Sharkey (2016). Using the matched sample, I begin my analyses with a set of logistic regression models that estimate the effect of the four events on the use of police force in the three days after the events. In the second step, I use a nonparametric approach to compare the observed and counterfactual trend in the proportion of stops that involve force (illustrated in fig. 1). To understand the construction of the counterfactual trend, imagine a simple scatterplot with the timing of the stop on the *X*-axis and the dependent variable on the *Y*-axis. For the observed stops, the *X*-axis reflects the actual timing of the stops. The matched stops from the control group, however, all took place before the event. Their place on the *X*-axis depends on the timing of the corresponding observed or treatment stop to capture what would have happened in the absence of the event. The lines shown in figure 1 are based on a nonparametric trend from a generalized additive model (Hastie and Tibshirani 1986) that estimates the probability that a stop involves force as a function of time and a number of control variables (see appendix A for more details). The causal effect can be estimated as the difference between the observed and counterfactual trend in the proportion of stops that involve the use of force (see illustration in fig. 1*A*).

The central assumption underlying this design is that, in the absence of the event, the two trends would have been the same. This assumption can be evaluated indirectly by comparing the observed and counterfactual trend before the event. This plausibility test is illustrated in figure 1*B*. It is based on a placebo treatment defined as a fictitious event two weeks before the actual event. Using this placebo treatment to determine the treatment and control group, I repeat the same matching procedure to construct an analogous counterfactual trend for the two weeks before the actual event. If the observed and counterfactual trends in the time interval before the event are

similar, the assumption that the two trends after the event would have been the same in the absence of the events is more plausible.

The second estimation strategy uses an RD design (Imbens and Lemieux 2008). In that case, the events define the cutoff points in the continuous variable "timing of pedestrian stop" and a jump or discontinuity in the proportion of stops that involve force right at the threshold provides evidence for a causal effect of the events (illustrated in fig. 1C). Accordingly, the design focuses on abrupt changes in the observed trend before and right after the event.

The two strategies are based on different assumptions and supplement each other. They focus on two patterns: (1) a comparison of the observed and counterfactual trend after the event (fig. 1*A*) and (2) a jump or discontinuity in the observed trend right around the event (fig. 1*C*). If both indicate an increase in the use of force, our confidence in the findings increases substantially. Appendix A discusses further details about the matching procedure, the regression models based on the matched sample, and the RD design.

Coding of Variables

The dependent variable used in this study measures the use of physical force by police officers in stop-and-frisk operations. The use of force is a continuum ranging from verbal commands to deadly force (Adams 1996; Garner et al. 2002; Durna 2011). Most police-citizen interactions do not involve the use of police force, and most of the police force used in encounters occurs at the lower level including verbal commands or light physical coercion (Adams 2015). Nonetheless, the disproportionate use of police force is a contentious political issue with serious consequences for the targeted citizens and for the relations between minority communities and law enforcement (e.g., Geller and Toch 1996*a*; Tyler et al. 2014). Indeed, previous research indicates that any negative encounters with law enforcement erode the sense of police legitimacy and increase critical perceptions of the police (Tyler and Wakslak 2004; Tyler et al. 2014). In this article, I measure the use of force with two different variables that are based on data provided by the police as part of the stop-and-frisk report form. The report form includes information on the use of physical force using the following categories: "Hands on Suspect," "Suspect on Ground," "Pointing Firearm at Suspect," "Handcuffing Suspect," "Suspect against Wall/Car," "Drawing Firearm," "Baton," "Pepper Spray," or "Other." All categories appear under the label "If Physical Force Was Used, Indicate Type" so that "Hands on Suspect" is clearly distinct from other questions, such as "Was Person Frisked?" Officers are trained to indicate multiple categories when different forms of force were used. In the main analysis, I follow common practice and use a binary indi-

cator for the use of physical force that is coded as one for all stops in which any form of force was used and as zero for all other stops. With the most common category "Hands on Suspect," this binary indicator includes lower levels of force so that the percentage of stops that involve force is relatively high (16.4% for whites, 22.2% for blacks, and 23.9% for Hispanics). Terrill and Mastrofski (2002) advocate for such a broad definition based on the argument that most force falls in this category and still has important repercussions for citizens including negative health outcomes (Harrell, Hall, and Taliaferro 2003; Geller et al. 2014). To supplement this measure, I also use a continuous variable that measures the amount of force used by officers. To construct this variable, I assign weights to the different categories listed above. The weights are based on a ranking of force categories by police officers in a survey conducted by Garner and Maxwell (2001). In the survey, 503 experienced officers were asked to rank 60 hypothetical types of force on a scale from 1 to 100 based on their personal experience and not departmental policy.[10] Using this ranking, the continuous variable is defined as the sum of the values for the different forms of force used in each stop. The results based on this continuous measure closely resemble the findings presented in the main text of this article. It is important to note that the use of police force and even the racial disparities do not necessarily imply racial bias. Force can be defensive or motivated by self-protection and is at the core of police activity (Harris 2009). If an officer thinks a suspect is armed, she or he will use force to pat down the suspect and perhaps handcuff him as a safety precaution. My approach focuses on changes in the rates of use of force and racial disparities in the response to events behavioral indications that measure disproportionate use of police force and may reflect bias.

The treatment indicator as the main independent variable is based on the timing of the stops and the different events. Stops that were conducted in the two weeks after the events are coded as the treatment group and stops in the year before the event as the control group. Other variables are the police precinct, the exact geocoded location of the stop, time of day, the circumstances of the stop, and the race, gender, age, and height of the stopped person.[11] I restrict my analysis to stops of blacks, whites, and (black or

[10] The categories used by Garner and Maxwell (2001) are not all the same as the categories used in the stop-and-frisk report sheet. To address this problem, I use broader categories to approximate the ranking. For example, the use of pepper spray in the stop-and-frisk data is matched to the broader category "police use of chemical agent" in Garner and Maxwell's data. Despite these imprecise matches, the resulting ranking makes sense intuitively with "Hands on Suspect" ranked as 15.9 and "Pointing Firearm at Suspect" substantially higher as 53.0.

[11] The circumstances that led to the stop include categories such as "carrying objects in plain view used in commission of crime," "fits description," "furtive movements," or

white) Hispanics as the three main ethnic/racial groups in New York City.[12] The number of stops for other groups (3.5% for Asians and 0.4% for Native Americans) is too small to support my analyses. Similar to many studies on racial profiling, police use of force, and other forms of discrimination (e.g., audit studies of employment), the analyses do not include information on officer's race. The NYPD has blocked several attempts to gain access to this information. The lack of officer's race is in line with the argument that police are a distinct social group and that events foreground tensions between the police and black youth. Nonetheless, officer race would be an important addition.

All variables were coded by police officers on the UF-250 form. The NYPD regularly trains officers on regulations and has implemented multiple layers of auditing to check the forms for complete and valid entries (Ridgeway 2007, p. 26). As reported by Ridgeway (2007), these layers mostly work effectively. Despite the rigorous standards enforced by the NYPD, a small percentage of cases has missing or implausible values on some of the variables (4.5%). These stops are dropped from all analyses. The training and auditing does not ensure that officers document all stops or report information correctly. While there is an incentive to report stops as a way to appear productive, officers might omit problematic stops or misreport certain aspects. Accordingly, officer-reported data might be biased to the extent that stops and the use of force are underreported (Black 1970). This concern is particularly pronounced for police brutality or retaliation in response to the events. Both factors, however, clearly suggest an underestimation of the potential effect and the use of physical force in general. To adversely affect my findings, officers not only would have to consistently record stops differently for minorities but also change this systematic misreporting after the events. Hence, the focus on a pre-post comparison rules out many of the problems in other studies. Any remaining bias would presumably entail an underestimation of our effects.

"suspicious bulge/object," which are included as binary variables in the regression analysis. The race categories are white, black, white Hispanic, black Hispanic, Asian/Pacific Islander, and American Indian/Alaskan Native.

[12] In supplementary regression models, I also separate the two Hispanic groups, which might be important considering the potential role of phenotype among Hispanics (e.g., Murguia and Telles 1996). While the use of police force is slightly different for the two groups, the findings reported here are essentially the same with some sample size problems when I separate the two groups. The distinct categories for white and black Hispanics are nonetheless important considering that they improve data quality particularly for comparisons between blacks and Hispanics.

RESULTS

Stop and Frisk in New York City

Over the last decade, stop-and-frisk operations first soared in New York City from approximately 160,750 in 2003 to about 684,000 in 2011 and then declined rapidly amid stark protests to about 45,000 stops in 2014. Table 1 shows some of the important descriptive statistics by racial group for all stops between 2006 and 2012. As reported widely, a disproportional number of stops target members of minority groups, with about 10% of stops involving whites, 54% African-Americans, and 32% Hispanics. The corresponding population shares in New York City are 45%, 25%, and 28%, respectively, so that minority groups are clearly overrepresented. These racial disparities in stop rates are troubling from a social perspective but do not necessarily reflect bias. At the same time, the success rate in terms of arrests, weapons, or contraband found tends to be higher among whites compared with minority groups, which might indicate that the police use more rigorous standards for stops of whites (Gelman et al. 2007; Persico 2009; Coviello and Persico 2013).

Here I focus on the use of police force in stop-and-frisk operations, which signifies a serious and potentially disturbing experience for the stopped person. Nearly a quarter of all stops involve the use of some physical force by the police, with a rate of 16.4% for whites, 22.2% for blacks, and 23.9% for Hispanics. As the number of stops increased over the years, the use of force declined from slightly above 30% in 2003 to 21.6% in 2011. Figure 2 shows the geographical distribution for the use of force across census tracts in 2011. In certain areas such as central Harlem or parts of the Bronx, some form of police force was used in over 50% of stops.

The Effect of Events on the Use of Police Force

I begin with a set of logistic regression models that estimate the effect of the four events on the use of police force against whites, blacks, and Hispanics in the three days after the event. These models are based on the matched sample and compare the stops after the event with similar stops before the event. They detrend the data by controlling for linear time, include the gender, age, and height of the stopped person, and are later extended by the Police Precinct as a fixed-effect term and the circumstances that led to the stop. Table 2 presents the results for the two events involving a black suspect, and table 3 for the events involving a Hispanic and white suspect. The findings reveal a race-specific pattern. The two shootings of NYPD police officers by black suspects in 2007 and 2011 increased the use of physical force by police officers against blacks substantially in the three days after the shooting

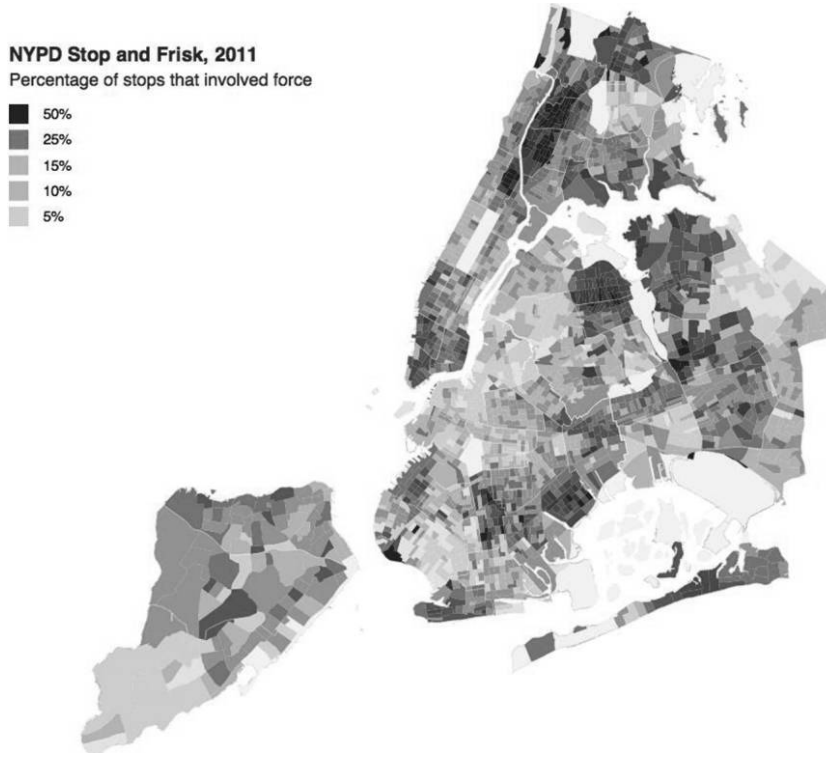


FIG. 2.—Police force in stop-and-frisk operations, 2011

(second and fifth model in table 2).[13] In particular, the proportion of stops that involve force against blacks increased by 16.0% (95% confidence interval [CI] = 7.3–25.5) from 21.9% to 25.4% in the first three days for the event in 2011 and by 13.3% (95% CI = 3.6–24.3) from 22.0% to 25.0% for the event in 2007 (estimates based on average marginal effects). These effect sizes are likely conservative considering that they are based on police-reported instances of police use of force. The use of force against whites and Hispanics, however, remained unchanged (other models in table 2).[14] The difference in the effect size for stops of African-Americans compared with the other groups

[13] The effect size varies somewhat depending on the specification of the matching procedure and the regression models. The two estimates, however, are substantially large and statistically significant in almost all specifications. App. C reports results from a range of different model specifications.

[14] The effect on the use of force against whites after the December 2011 shooting deviates from this pattern. The estimate is large and negative but not statistically significant for most specifications. This negative point estimate might reflect a decrease in the use of

TABLE 2
THREE-DAY EFFECT OF TWO SHOOTINGS BY BLACK SUSPECTS ON USE OF POLICE FORCE

	SHOOTING IN DECEMBER 2011			SHOOTING IN JULY 2007		
	White	Black	Hispanic	White	Black	Hispanic
T-effect of event	−.30	.20***	−.10	−.01	.16**	.02
	(.16)	(.05)	(.08)	(.14)	(.06)	(.08)
Control variables:						
Time	−.00	−.00***	−.00***	.00***	.00*	.00
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Female	−.60***	−.44***	−.59***	−.26*	−.75***	−1.18***
	(.12)	(.05)	(.07)	(.12)	(.07)	(.10)
Height	−.05	.04**	.03	.06	.04*	.03
	(.04)	(.01)	(.02)	(.04)	(.02)	(.02)
Age	.13**	−.08***	−.06**	.01	−.15***	−.27***
	(.04)	(.01)	(.02)	(.04)	(.02)	(.02)
Age (square)	−.10***	−.03**	−.04**	−.04	−.03*	.03
	(.03)	(.01)	(.01)	(.03)	(.01)	(.02)
Constant	−1.64***	−1.53***	−1.46***	−1.34***	−1.15***	−1.18***
	(.06)	(.03)	(.03)	(.06)	(.03)	(.04)
N	10,051	52,731	31,472	8,027	35,673	20,534

NOTE.—Estimates are based on matched sample and reported in log odds. SEs are in parentheses.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

is large and highly significant (based on a pooled model for all three racial groups with interaction terms between race and the treatment indicator). The two events involving a Hispanic and white suspect did not increase the use of force against any of the groups (table 3). These findings indicate a race-specific pattern in the response to events with a pronounced increase in the use of force against blacks but not against other citizens after the two events involving black suspects and the lack of any effect for the murders involving a white and Hispanic suspect.

force against whites after the event, or it might simply be a consequence of the uncertainty around the estimates for whites with substantially smaller sample sizes compared to the other groups. The corresponding local regression curve presented in figure 4*B* helps us to reconcile this somewhat contradictory result. It shows that the counterfactual trend is consistently above the observed trend even before the event and does not follow the same gradient. After the event, it starts at a similar level as the observed trend but slopes up immediately. The observed trend itself does not show a drop in the proportion of stops that involve force after the event. Overall, this pattern does not correspond to what we would expect from a negative effect. It indicates that the counterfactual trend is not a good comparison in this case. These findings suggest that the negative but statistically insignificant point estimate for whites does not signify a decrease in the use of force but instead seems to be driven by uncertainty in the effect estimates. Estimates based on the RD design discussed below confirm this conclusion.

TABLE 3
THREE-DAY EFFECT OF SHOOTINGS BY WHITE AND HISPANIC SUSPECT ON USE OF POLICE FORCE

	SHOOTING IN MARCH 2011 (Hispanic Suspect)			SHOOTING IN MARCH 2007 (White Suspect)		
	White	Black	Hispanic	White	Black	Hispanic
T-effect of event	−.07	−.04	.02	.03	.04	.02
	(.14)	(.05)	(.06)	(.15)	(.06)	(.09)
Control variables:						
Time	.00	.00*	−.00	.00*	.00	.00**
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Female.	−.61***	−.56***	−.64***	−.67***	−.75***	−.73***
	(.13)	(.05)	(.06)	(.15)	(.07)	(.09)
Height	−.04	.05***	−.04*	.09*	.05**	.04
	(.04)	(.01)	(.02)	(.04)	(.02)	(.02)
Age	.06	−.12***	−.10***	−.19***	−.17***	−.22***
	(.04)	(.02)	(.02)	(.05)	(.02)	(.03)
Age (square)	−.11***	−.02*	−.06***	−.04	.00	.02
	(.03)	(.01)	(.02)	(.03)	(.01)	(.02)
Constant	−1.41***	−1.10***	−.85***	−1.46***	−1.27***	−1.33
	(.06)	(.02)	(.03)	(.06)	(.03)	(.04)
N	8,892	50,957	33,018	7,682	36,150	20,758

NOTE.—Estimates are based on matched sample and reported in log odds. SEs are in parentheses.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

The findings are robust to different model specification (see appendix C). They remain stable after adding Police Precinct as a fixed-effect term and the circumstances that led to the stop as additional control variables. The findings are in line with my argument about a race-specific response. From this perspective, the shootings by black suspects increase the use of police force against blacks through their influence on tensions between the police and black communities, perceptions of minority threat, and concerns about personal safety among officers. The events act as cues that explicitly or implicitly foreground stereotypes of African-Americans as hostile and violent and increase the perceived level of threat. As a consequence, police officers seem to be more inclined to use force when stopping blacks in situations that did not lead to the use of force before the events. The explanation is based on different processes that all focus on a racially biased response by police officers (see below for a discussion of alternative accounts). I cannot, however, conclusively distinguish the different mechanisms driving the increase in police use of force against blacks.

Temporal Pattern and Duration of Effect

To further develop these findings and consider the role of time, I turn to local regression models with a nonparametric trend for the timing of stops. Figures 3–5 present the results showing the observed and counterfactual trend for the use of police force before and after the four events. Figure 3 focuses on the findings from the two shootings by black suspects on March 14, 2007 (fig. 3*A*), and December 12, 2011 (fig. 3*B*). It shows the proportion of police stops that involved the use of force against blacks in the two weeks before and after the shootings together with the counterfactual trends based on similar stops before the events. Comparing the observed and counterfactual trends *after* the events (strategy described in fig. 1*A*) shows a considerable increase in the proportion of stops that involve force against blacks. Resembling the estimates from the logistic regression models, the use of force against blacks increased by 11.6% and 18.4% in the three days after the event (95% CI = 2.7%–20.9% and 9.6%–27.9%, respectively).[15] The local regression curves also reveal that the increase in the use of force was temporary and declined over time without an enduring lift in the base rate of force. The figures indicate that the increase lasted about 10 days for the event on December 12, 2011, and about 3.5 days for the shooting on March 14, 2007. Analysis based on logistic regression models confirms this finding.

Figure 4 presents the corresponding local regression curves for stops of whites and Hispanics, and figure 5 the results for the two events involving a white and Hispanic suspect. The figures provide no evidence for an increase in the use of force. They show no clear pattern in the observed and counterfactual trend after the event and no jump in the observed trend before and after the event. This finding confirm the results from the logistic regression models. It indicates that the two shootings by black suspects temporarily increased the use of force against blacks but not against whites and Hispanics. The two events involving nonblack suspects did not change the use of force against any of the groups.

Placebo Treatment and other Sensitivity Analyses

A causal interpretation of these findings relies on the assumption that the observed and counterfactual trends *after* the events would have been the same in the absence of the events. Figures 3, 4, and 5 include a comparison of these trends *before* the events, which allows me to evaluate this assumption (see fig. 1*B* for an illustration of this approach). The two lines are very similar *before* the events, indicating that the counterfactual trend is plausible.

[15] These estimates are based on a comparison of the predicted values from the nonparametric curves for the observed and counterfactual trend (solid and dashed line). The CIs are calculated with a simulation-based approach (King, Tomz, and Wittenberg 2000).

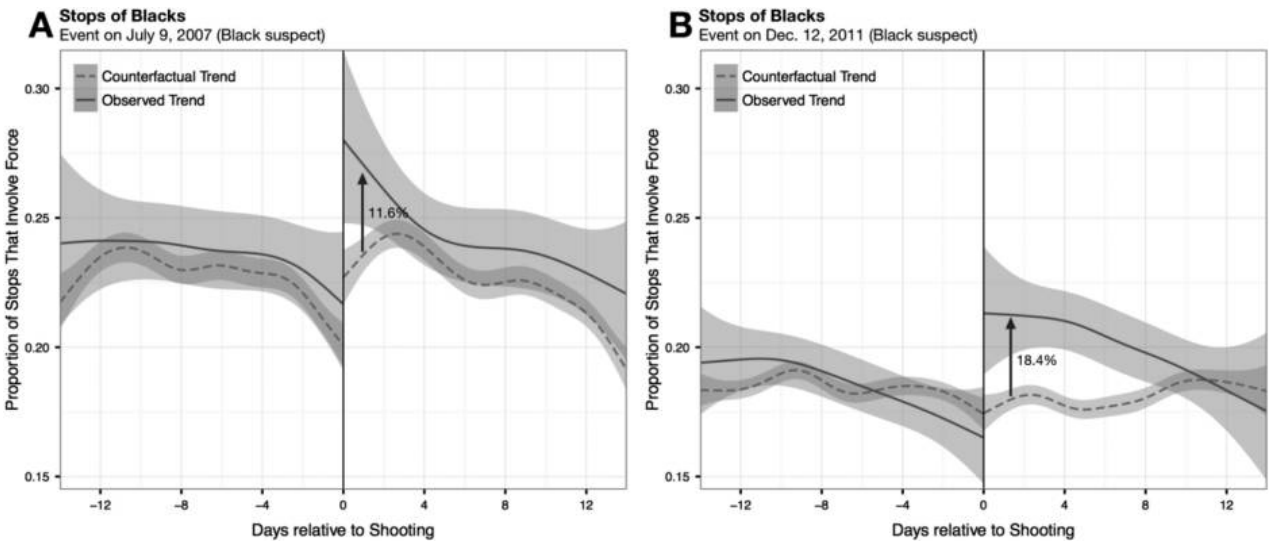


FIG. 3.—Effect of two shootings by black suspects on the use of police force against blacks. The lines show local regression estimates and 95% CIs. A comparison of the observed and counterfactual trends indicates that the use of force increased by 11.6% (95% CI, 2.7–20.9) and 18.4% (95% CI, 9.6–27.9) in the first three days after the events. The sample sizes for the observed trend are 15,380 (*A*) and 24,848 (*B*).

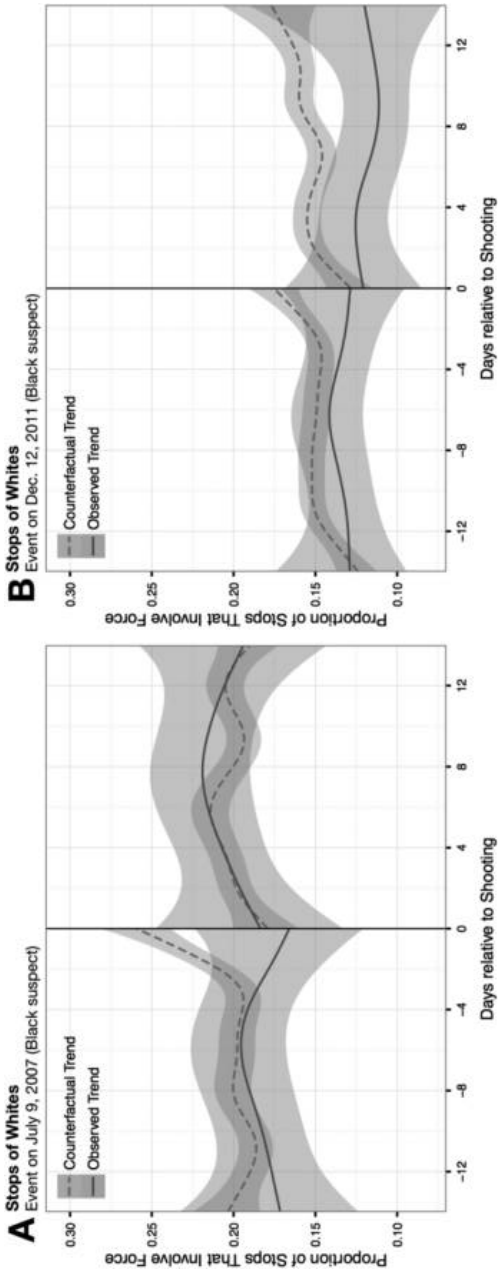
Appendix B includes additional sensitivity analysis based on fictitious event simulation (placebo treatments). The findings support the results presented here, indicating that chance alone is highly unlikely to account for the observed pattern.

Effect Heterogeneity across Events

The increase in the use of force against blacks after the shootings by black suspects resulted from two markedly different events. As described earlier, the 2007 event occurred during a vehicle stop and was initially nonlethal (the officer passed away five days later in the hospital). The fatal 2011 shooting, in contrast, occurred when officers responded to a residential robbery and the officer died shortly after the event. The different nature of the two events allows us to make two tentative conclusions. First, the initially nonlethal shooting of Officer Timoshenko in 2007 shows a clear increase in the use of police force against blacks, but this increase is less pronounced and the duration of the effect is shorter. The difference in the size of the effect might be explained by the fact that the officer died only several days later. Importantly, the finding also shows that the response to events is not confined to fatal attacks. The use of force against blacks increased right after the initial shooting, indicating that extreme acts of violence against a police officer can increase the use of force against blacks even when they are nonlethal. Second, the 2011 incident was entirely unrelated to pedestrian or vehicle stops so that the increase in the use of force does not seem to be confined to events related to stop-and-frisk operations. These conclusions are tentative considering the small number of events. Future research should establish when, where, and what kind of events trigger periods of increased racial disparities in police use of force.

Regression Discontinuity Design

The results based on the matched sample provide clear evidence that the two events involving black suspects increased the use of police force against blacks. The findings from an RD design confirm these results. First, a comparison of the observed trends before and after the two events in figure 3 (solid lines) shows a jump in the proportion of stops that involve force after the shootings. The figure resembles the graphical presentation of exploratory RD analysis with additional covariates that adjust for common temporal patterns such as time of day (Imbens and Lemieux 2008). Second and more formally, figure 6 presents results for RD estimates from a sharp RD design with the continuous variable time as the forcing or running variable and the timing of the event as the cutoff point. The estimates are based on



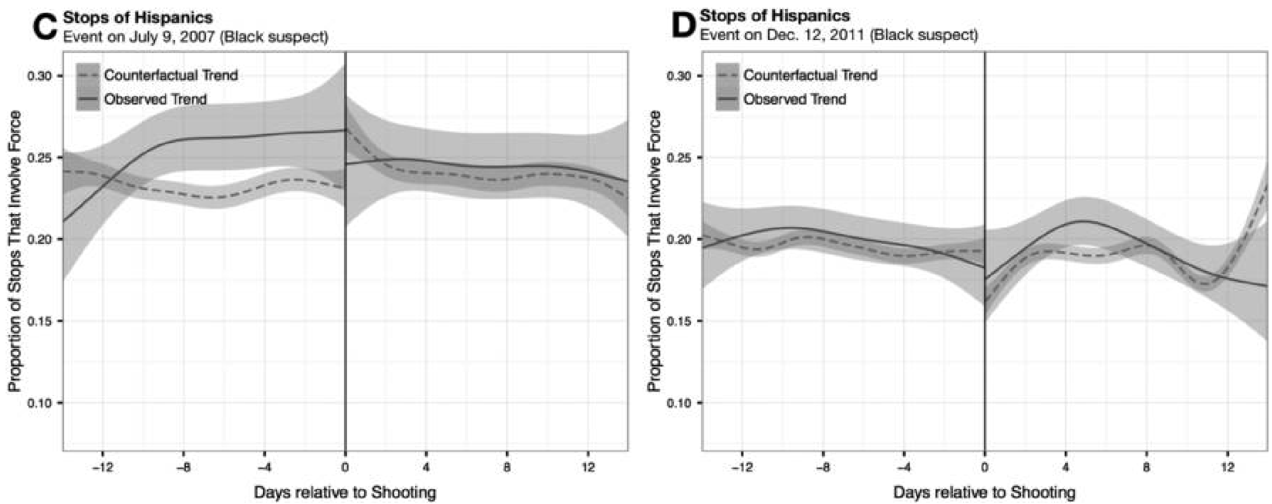


FIG. 4.—Effect of two shootings by black suspects on the use of police force against whites and Hispanics. The lines show local regression estimates and 95% CIs. The sample sizes for the observed trend are 3,434 (*A*) and 4,248 (*B*) for whites and 8,774 (*C*) and 14,778 (*D*) for Hispanics. Compared to figure 3, the substantially smaller sample size increases the uncertainty in the estimates.

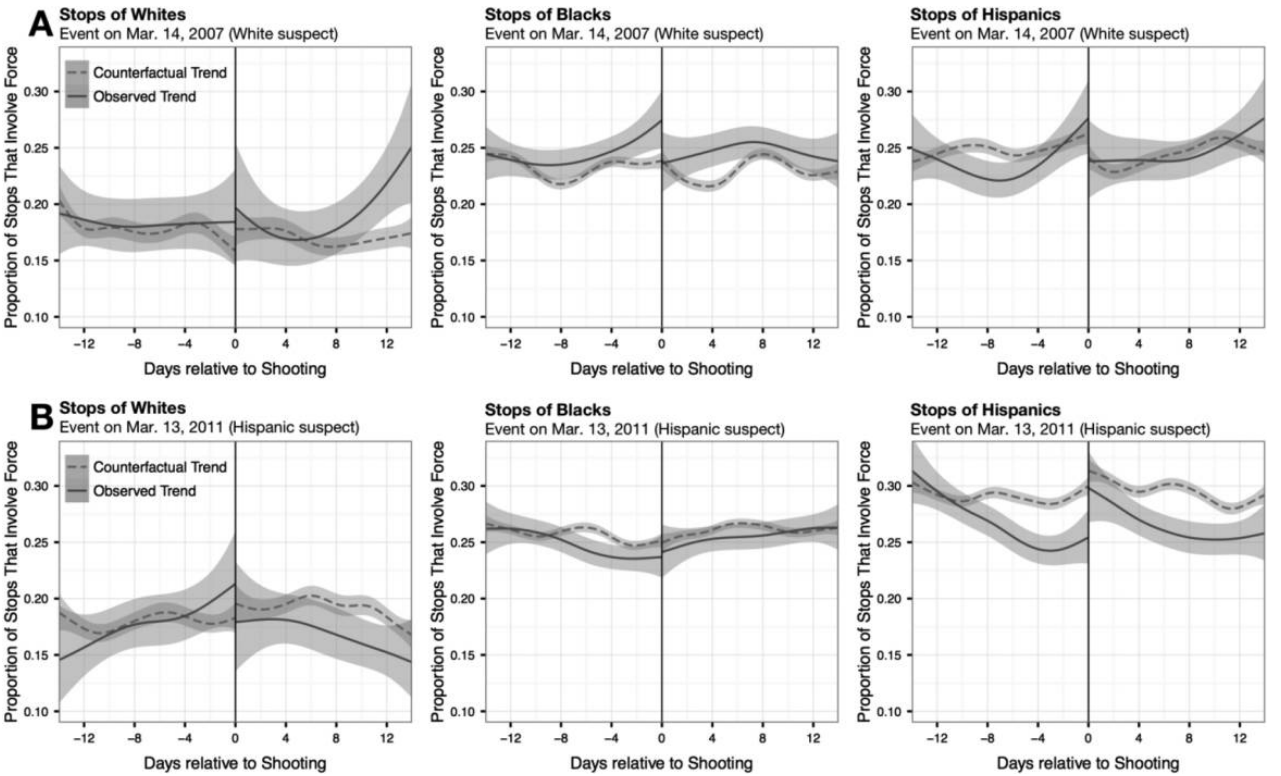


FIG. 5.—Effect of two homicides by a white and Hispanic suspect on the use of police force. The lines show local regression estimates and 95% CIs for the observed and counterfactual trend based on matched stops.

local linear regressions with triangular kernels that assign higher weights to observations that are closer to the cutoff (for details on RD estimates see Imbens and Lemieux [2008], Lee and Lemieux [2010], and Calonico, Cattaneo, and Titiunik [2015]). A key parameter for these kernel-based estimators is the bandwidth that determines the size of the kernel and therefore the range of data to the left and right of the cutoff point used in the estimation procedure. Figure 6 shows the effect sizes from over 300 RD estimates together with CIs (*Y*-axis) as a function of the bandwidth (*X*-axis). Vertical lines indicate the optimal or data-driven bandwidth based on the three most commonly used methods: Imbens and Kalyanaraman (2012; IK in the figure), Calonico, Cattaneo, and Titiunik (2014; CCT in the figure), and the cross-validation method (Ludwig and Miller 2007; CV in the figure). Irrespective of the bandwidth, the results show a clear increase in the use of police force against blacks after the two events involving black suspects. Particularly for the three most commonly used data-driven or optimal bandwidths, the estimates are large and statistically significant (note that the estimates are based on linear probability models and not logistic regressions as the previous results). For very small bandwidths (less than a quarter of the CCT optimal bandwidth), the uncertainty increases substantially and the effect estimate decreases for the event in July 2007. This decrease might reflect the fact that the news of the shooting has to spread among officers. The findings from the RD design provide further evidence for a causal interpretation of my findings. Using different estimation strategies based on different assumptions, the results consistently indicate that the use of police force against blacks increased substantially after the two events involving black suspects.

Alternative Explanations

The findings reveal a race-specific pattern in the response to the four events with a pronounced increase in the use of force against African-Americans but not other citizens and the lack of any effect for the murders involving a white and Hispanic suspect. This pattern is consistent across different estimation strategies and model specifications. The theoretical argument presented earlier suggests that relevant events increase the use of police force against minority groups through their influence on tensions between the police and black communities, perceptions of minority threat, and concerns about personal safety among officers. The response is driven by preexisting stereotypes about blacks as hostile and violent that explain the race-specific pattern in the response to events. These processes focus on a racially biased response by police officers. There are, however, alternative explanations of the observed pattern. One possibility is that the two shootings by black

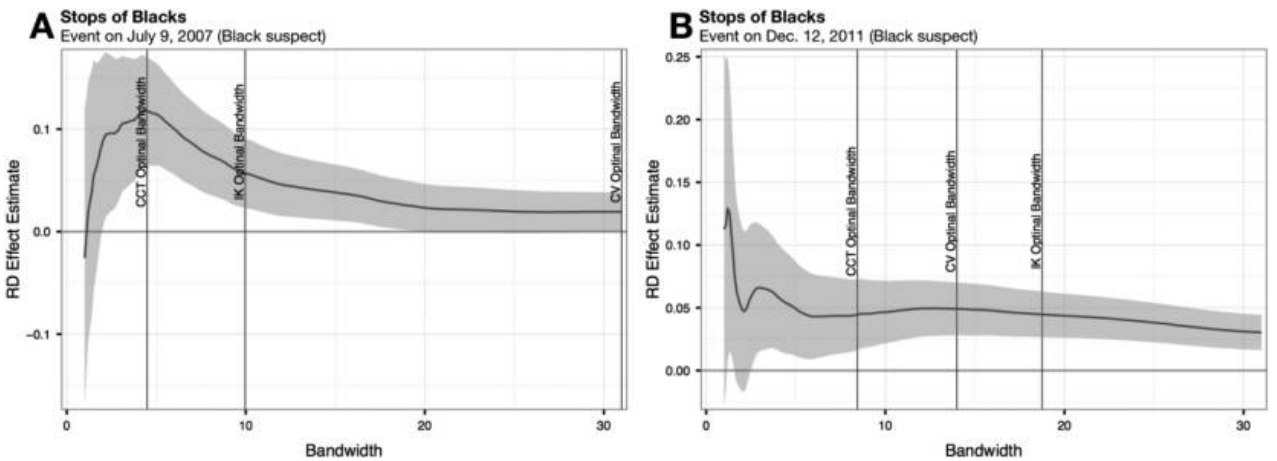


FIG. 6.—RD estimates as a function of bandwidth for the two shootings by black suspects on the use of police force against blacks. The figures present results from a series of local linear regression models based on an RD design. They show the effect size from over 300 RD estimates together with CIs (*Y*-axis) as a function of the bandwidth (*X*-axis). Vertical lines indicate the optimal or data-driven bandwidth based on the three most commonly used methods: Imbens and Kalyanaraman (2012; IK in the figure), Calonico, Cattaneo, and Titiunik (2014; CCT in the figure), and the cross-validation method (Ludwig and Miller 2007; CV in the figure).

suspects correspond to increases in drug- or gang-related activities, producing the need for police to exercise self-defense in black neighborhoods. From this perspective, the two shootings by black suspects were part of crime waves concentrated in black neighborhoods that also triggered the increased use of force against blacks. A factor that is causally prior to the events would have caused both the shootings and the increased use of force against blacks. To evaluate this possibility, I examine whether the increased use of force after the events coincided with an increase in drug and weapon seizures or arrests during stop-and-frisk operations in predominantly black neighborhoods. A crime wave in black neighborhoods would imply that the police not only exert more force but also that this activity leads to more arrests and the discovery of weapons and contraband. For this purpose, I restrict the sample to police precincts in which blacks are the majority and reestimate the same logistic regression models as above using arrest, weapon found, and contraband found as dependent variables. The results show that none of these three factors increased after the events. Instead, the three-day estimates show small and statistically insignificant effect estimates. For the shooting in 2011, the estimates are −0.074 for arrests ($P = .64$), 0.336 for weapon found ($P = .21$), and −0.21 for contraband found ($P = .45$) with similar estimates for the event in 2007 (except for the arrest rate, which increased significantly). Accordingly, the use of police force increased substantially after the two shootings by black suspects, but this increase did not coincide with a higher number of arrests or an increase in weapons and contraband found. This finding clearly contradicts the argument that the race-specific pattern is a consequence of a crime wave in black neighborhoods.

A second possibility is an organizational response. From this perspective, the increase in the use of force against blacks reflects a response by superiors in certain police precincts who direct their officers to exert more force on the street. Such a departmental or institutional mandate following a critical incident might explain the race-specific pattern if the organizational response is concentrated in predominantly black precincts and restricted to events involving black suspects. Such a response might be considered racially biased itself. It is in line with a response based on minority threat that is driven by interests of the general public or the police department as a whole. As such, it does not contradict the theoretical argument outlined above. Instead, it further specifies the mechanisms by focusing on an institutional response to minority threat and not a response driven by individual officers. This narrowed account of the race-specific pattern implies a larger increase in the use of force in predominantly black areas. To evaluate this implication, I reestimate the effect of the two shootings by black suspects with an additional interaction term between the treatment effect (stops after event) and the proportion of black residents first on the precinct and then the census tract level. The results from these multilevel logistic regression models are presented in

table 4. They show generally small and statistically insignificant estimates for the interaction terms with inconsistent signs across the two events, indicating that the increase in the use of force did not depend on the proportion of blacks in the neighborhood. Accordingly, the race-specific pattern in the response to the events does not seem to be driven solely by an organizational response that focuses on predominantly black neighborhoods.

Finally, the increase in the use of force might be the consequence of changes in the behavior of citizens and not police officers. After a highly publicized shooting of a police officer, fear and anxiety among African-Americans who are stopped by the police might increase when they believe the police engages in retaliation or when the citizen-police relations are strained. As a result, suspects might become defensive or more likely to try to escape from police custody, resulting in the use of force. Desmond, Papachristos, and Kirk's

TABLE 4
EFFECT OF THE TWO SHOOTINGS BY BLACK SUSPECTS ON THE USE OF FORCE AGAINST BLACKS BY PROPORTION OF BLACKS IN NEIGHBORHOOD

	SHOOTING IN DECEMBER 2011		SHOOTING IN JULY 2007	
	Precinct	Census Tract	Precinct	Census Tract
T-effect of event	.22***	.22***	.16*	.16***
	(.06)	(.06)	(.07)	(.07)
Proportion black (precinct)	−.00		.08***	
	(.01)		(.01)	
T × proportion black (precinct)	−.07		.01	
	(.05)		(.06)	
Proportion black (census tract)		−.04***		.05***
		(.01)		(.01)
T × proportion black (census tract)		−.08		.02
		(.05)		(.06)
Control variables:				
Time	−.00***	−.00***	.00*	.00*
	(.00)	(.00)	(.00)	(.00)
Female	−.44***	−.45***	−.75***	−.75***
	(.05)	(.05)	(.07)	(.07)
Height	.04**	.04**	.05**	.05**
	(.01)	(.01)	(.02)	(.02)
Age	−.08***	−.08***	−.14***	−.14***
	(.01)	(.02)	(.02)	(.02)
Age (square)	−.03**	−.03**	−.03*	−.03*
	(.01)	(.01)	(.01)	(.01)
Constant	−1.53***	−1.51***	−1.18***	−1.17***
	(.03)	(.03)	(.03)	(.03)
N	52,579	52,578	35,626	35,578

NOTE.—Estimates are based on matched sample and reported in log odds. SEs are in parentheses.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

(forthcoming) finding that widely publicized incidents of police brutality decrease crime reporting through 911 provides support for the idea that certain—although substantially different—events can alter how residents behave towards the police. To evaluate this possibility, I use additional data from the UF-250 form. This information includes indicators about suspects' behavior such as "Evasive, False or Inconsistent Response to Officer's Questions," "Changing Direction at Sight of Officer/Flight," and additional questions about the reasons why a person was frisked such as "inappropriate attire" and "verbal threats of violence by suspect." These indicators provide direct information about the behavior of stopped persons from thousands of police-citizen encounters after the events. Comparing the behavior rates before and after the events shows inconsistent, mostly small and statistically insignificant differences, indicating that changes in citizen behavior do not explain the increase in the use of police force.

CONCLUSION

This article argues that racial bias in policing and discrimination more broadly is not static but fluctuates, partly driven by significant events that provoke intergroup conflict and foreground racial stereotypes. My findings provide quasi-experimental evidence showing that incidents of extreme violence against police officers can lead to periods of substantially increased racial disparities in the use of police force. Using data from 3.9 million time and geocoded pedestrian stops, the findings show that, relative to similar stops before two shootings of NYPD police officers by black suspects in 2007 and 2011, the use of physical force by police officers against blacks increased substantially in the days after the shootings. The temporal duration of the effect is limited, but the relatively frequent nature of the events makes the consequences profound, particularly at a time of intense tensions between the police and black communities.[16] Indeed, previous research indicates that negative encounters with law enforcement erode the sense of police legitimacy and increase critical perceptions of the police (Tyler and Wakslak 2004; Tyler et al. 2014). Negative encounters far outweigh the influence of positive experiences (Skogan 2006). Timely and targeted interventions in reaction to certain events might be a way to counter an increase in the use of police force after relevant events.

Aside from the importance of these results for the ongoing debate about racial profiling and police use of force, this study has broader theoretical and empirical implications. First, the interpretation of my findings extends

[16] FBI statistics indicate that an average of 55.2 officers were feloniously killed per year between 1995 and 2013 and that thousands were assaulted (FBI Uniform Crime Reporting, Law Enforcement Officers Killed and Assaulted, 1993–2013).

beyond acts of extreme violence against police officers. It suggests a general set of processes where local events create intergroup conflict, foreground stereotypes, and trigger discriminatory responses. From this perspective, discriminatory behavior arises not only from static conditions but also from temporal sequences of events and responses. This process is applicable to all kinds of everyday interactions, both with the police and with other actors who might engage in discriminatory behavior (landlords, teachers, etc.). For instance, officers who have hostile interactions with local youth (or hear about hostile interactions that other officers have had) may respond more forcefully in their next interaction even if the interaction is with a different individual or group of individuals. Violent crimes or homicides that target particular ethnic or other social groups or threaten specific neighborhood communities may precipitate waves of discriminatory acts. A real estate agent might reject African-American applicants or steer them to black neighborhoods after a violent crime in a predominantly white or racially mixed community. Racial bias in hiring decisions might increase in low-wage labor markets that are particularly prone to discrimination.

This argument proposes an event-centered study of racial profiling, police use of force, and discrimination more broadly. Events are a largely neglected factor in the literature on discrimination. Similar to space and place, the findings indicate that events are an important contextual influence that frames subsequent interactions and triggers discriminatory behavior. This argument extends existing theories of discrimination and minority threat by a broader understanding of the social environment that nurtures racial bias. It contributes to a small but growing line of research in criminology that uses events to better understand the mechanisms that trigger the diffusion of intergroup violence (Papachristos 2009; Vargas 2014). The focus on events points to a mechanism that explains short-term (and potentially long-term) changes in the rates of discriminatory behavior. It extends the research agenda from *where* discrimination takes place to *when* it happens, helping us to understand the causal dynamics underlying racial profiling, disproportionate use of police force, and discrimination more broadly.

Second, the focus on events presents a complementary approach for the study of racial bias. Attributing bias in policing or other areas is challenging considering that unobserved factors might explain and partly justify different outcomes for racial and ethnic groups. Audit studies are a popular approach to rule out such confounding factors and have been applied repeatedly in research on employment and housing (Pager and Shepherd 2008). Such an experimental design, however, faces ethics problems in some areas such as the use of police force and has a number of other limitations (Heckman 1998; Pager 2007). The current study presents a complementary approach focused on temporal variations in certain behavior. The measurement of bias in terms of changes in the rates of use of force moves the dis-

crimination debate from trying to measure the subjective (perceived suspicion, selection of citizens for stops) to behavioral indicia that may reflect bias. At the same time, the design provides a methodological template for future research in criminology that examines how events trigger episodes of intergroup violence.

The findings from this study are also limited in several regards. First, this study lacks information on officer's race as a potentially important covariate. In fact, most research on racial profiling, police use of force, and other forms of discrimination (e.g., audit studies of employment) does not include information on the ascribed characteristics of the relevant actor (for an exception see Antonovics and Knight 2009). For the study at hand, an officer's race might be important to further evaluate my argument about racial bias insofar as the increase in the use of force should be more pronounced among white officers. Second, the events examined in this study substantially increased the use of police force against blacks, but the duration of the effects was modest. Research in the priming literature suggests that single primes temporarily increase prejudices but repeated primes can gradually lead to long-term changes (Althaus and Kim 2006). Accordingly, frequent exposure to certain events such as crimes that involve black suspects might not only trigger temporary periods of discrimination but also permanently raise the use of police force against minorities or other discriminatory acts. Future studies should establish whether frequent events can have such a long-term effect on discrimination.

APPENDIX A

Estimation Strategy

My research design is based on two estimation strategies illustrated above in figure 1. This appendix provides further technical details about the implementation of these methods.

Matching Procedure

My analyses use a matching procedure to construct a counterfactual trend that captures what would have happened in the absence of the event. All observations in the two weeks after the events are defined as the treatment group, and all observations in a certain time interval before the events as the control group. Accordingly, stop-and-frisk operations after the event are matched to similar stops before the event. These matched stops are used to construct the counterfactual trend (dashed line on the right-hand side of fig. 1*A*). For the plausibility test illustrated in figure 1*B*, I repeat the same procedure for the two weeks before the event. In that case, I use two weeks

before the event as the placebo treatment so that all observations in the two weeks after the placebo treatment (the two weeks before the actual event) are used as the treatment group and the observations in a certain time interval before the placebo treatment as the control group. For both the main analysis and the plausibility test, the matching procedure is based on exact matching for precinct and race and nearest-neighbor matching with replacement using the Mahalanobis distance (Zhao 2004) for additional covariates such as the location of stops. Mahalanobis distance clearly outperforms other distance measures such as propensity score because of the pairwise definition of distance, which is important for geographical matching. The actual implementation of the matching procedure relies on three important decisions.

First, what covariates are used in the matching procedure. Aside from police precinct and suspect race used for exact matching, the Mahalanobis distance is based on the geographical location of the stop in terms of the X and Y coordinates on the U.S. Plane Coordinate Systems for New York; the time of day of the stop translated to X and Y coordinates on a clock circle;[17] the gender, age, and height of the stopped person; and in additional specifications the circumstances of the stop. These variables are all influential predictors for the use of force in police stops and as such are important covariates. Second, the matching ratio determines the number of stops before the event (control group) that are matched to each stop after the event (treatment group). I conduct the same analyses for matching ratios of 1:5, 1:10, 1:15, and 1:25, and all lead to similar results. The main analyses reported in the text are based on a 1:25 matching ratio, and I report additional findings from a range of different model specifications in appendix C. Finally, the size of the time interval for the control group plays an important role for the outcome of the matching procedure. A larger time interval leads to extremely good matches in terms of geographical location and time of day of the stop as the most important predictors for the use of police force. Smaller time intervals reduce the temporal distance between the stops after the event and their matches but increase the imbalance in terms of the other covariates. For the analysis presented here, I use different time intervals ranging from one month to one year. In general, longer time intervals perform better both in terms of balance and in terms of the sensitivity analysis that compares the observed and counterfactual trend before the event. Accordingly, the main results discussed in this paper are based on a time inter-

[17] The coding of time of day as X and Y coordinates on the clock circle is necessary to properly define distance in time of day so that the difference between 11:55 and 12:05 is the same as the difference between 11:55 and 11:45.

val of one year and I report additional results from a range of different model specifications in appendix C, which strongly support the main findings.

To select the specification used in the main text of the paper, I compared the balance between the different matching specifications, used the placebo-treatment test based on the difference in the observed and counterfactual trend before the event, and discarded specifications that produce estimates with high uncertainty. Appendix C reports results from additional specifications. For the main specification, the balance between the treatment and control group is extremely high in the matched sample for stops of blacks and Hispanics and to a lesser extent for whites. For the December 2012 shooting, the average distance between treatment and control cases in the matched sample is 0.2 miles for blacks, which corresponds to the average length of a single north-south city block in Manhattan, and the average difference in daytime is below 60 minutes (e.g., a stop at 10:55 P.M. is matched to previous stops that, on average, took place between 9:55 and 11:55 P.M.). Accordingly, the pedestrian stops after the four shootings are compared to stops that took place in the same area, around the same time of the day, and of similar people (same race, similar age and height).

Regression Analysis

Using the matched sample, I begin my analysis with a set of logistic regression models that estimate the effect of the four events on the use of police force in the three days after the events. Formally, the models can be specified as

$$P(y_{ij} = 1) = \text{logit}^{-1}(\alpha + \delta D_{ij} + \boldsymbol{X}_{ij}\beta), \quad (1)$$

where i and j are indexes for stops and police precincts, respectively. The coefficient δ for the treatment indicator D_{ij} is the crucial statistic and represents the difference in the use of police force between the (matched) stops before and treatment stops for a certain time interval after the events. Under the assumptions discussed above, this difference can be interpreted causally as the average causal effect of the treatment on the outcome. $\boldsymbol{X}_{ij}$ represents a matrix of control variables, and β a vector of corresponding coefficients. The control variables include linear time as the date of the stop relative to the events, the gender, height, and age of the stopped person, and, in additional specifications, a precinct fixed-effect term and the circumstances of the stop. The date of the stop removes general temporal patterns in the use of force by police officers and therefore captures any trend we would have expected in the absence of the events. The precinct fixed-effect terms account for factors such as police leadership, the allocation of officers, the local crime rates, and other factors.

In the second step, I use a nonparametric approach to examine the trends for the proportion of stops that involve force before and after the four events. These models relax many of the assumptions of conventional regression models. They provide a convincing visual portrait of the findings and allow me to examine the temporal duration of the effect. In particular, I use generalized additive models (Hastie and Tibshirani 1986) and predict the probability that a stop involves force based on a nonparametric term for the time of the stop of the treatment unit, which is fitted separately for pre- and postevent data and for the treatment and control group (observed and counterfactual trend). Formally, the models can be specified as

$$P(y_{ij} = 1) = \text{logit}^{-1}\left[\alpha + \phi_d(\text{time}_{ij}) + \boldsymbol{X}_{ij}\beta\right], \tag{2}$$

where i and j are indices for stops and police precincts, respectively. The function $\phi_d(\text{time}_{ij})$ models a nonparametric relation between the outcome and the time of the stop. As indicated by the index d, the function is estimated separately for the treatment and control group. The difference between these two trends allows me to estimate the treatment effect indirectly by comparing the predicted values from the nonparametric curves for the observed and counterfactual trend (essentially, a comparison of the two trends illustrated in fig. 1*A*). The CIs are calculated using a simulation-based approach as outlined by King et al (2000). $\boldsymbol{X}_{ij}$ includes the same set of control variable as in the logistic regression models.

RD Design as an Alternative Estimation Strategy

An alternative estimation strategy relies on an RD design (Imbens and Lemieux 2008; Legewie 2012). In that case, the events define the cutoff points in the continuous variable "timing of pedestrian stop," and a jump or discontinuity in the proportion of stops that involve force right at the threshold provides evidence for a causal effect of the events (illustrated in fig. 1*C*). Several authors emphasize the virtues of such a design, which in some cases allows researchers to draw causal inference from nonexperimental data (Cook, Shadish, and Wong 2008; Cook and Wong 2008). Nonetheless, the design has some limitations for the application to the case at hand. First, an RD design provides no direct way to examine the temporal duration of effects, which is a central part of my research question.[18] Second, the exact

[18] To a certain extent, the guiding principle of the design even conflicts with such an attempt. In particular, the core idea of the design is to compare units that are immediately to the right and left of the cutoff point because bias increases as you move away from the threshold. Following this principle, the average causal effect of the treatment at the discontinuity point c is formally defined at the limit as $\lim_{x \downarrow c} E[Y_i|X_i = x] -$

cutoff point is essential for an RD design but not clearly defined in our case. While the time of the shootings is known, the news has to spread to police officers around the city before it can have any impact. Even though the news spreads rapidly, this process introduces some uncertainty about the exact cutoff point. These small differences can be consequential for estimates based on an RD design because RD estimates assign higher weights to units that are closer to the cutoff. These limitations of an RD design for the case at hand speak for the design outlined above. Nonetheless, the different estimation strategies are based on different assumptions and supplement each other. The central idea that discontinuities around the cutoff point can be used to estimate the causal effect provides further evidence for the causal interpretation of my findings.

$\lim_{x \downarrow c} E[Y_i|X_i = x]$ (Imbens and Lemieux 2008). Estimating the temporal duration of the effect, however, relies on moving away from the cutoff point. For the case at hand, the time of day is an important predictor for the use of police force and introduces major challenges without a clearly defined counterfactual trend.

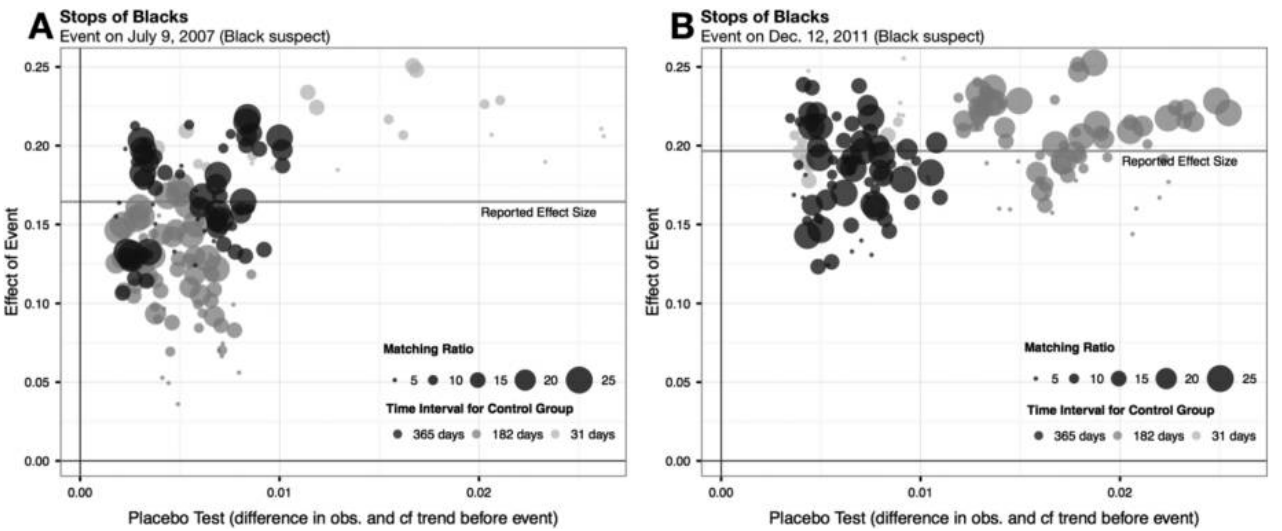


FIG. A1.—Sensitivity of results to 161 different model specifications for the two main effects observed in this study—the effect of the two fatal shootings by black suspects. The *X*-axis shows the placebo test based on the absolute difference between the observed and counterfactual trend in the two weeks before the events. The *Y*-axis shows the three-day effect estimates from a logistic regression. The horizontal lines show the effect size from the model reported in the main text.

APPENDIX B

Placebo Treatment and Other Sensitivity Analyses

The key assumption of my quasi-experimental research design is that the observed trend would have been the same as the counterfactual trend in the absence of the events (Keele and Minozzi 2013). Two sensitivity analyses can be used to evaluate the plausibility of this assumption. First, I construct a "placebo" treatment based on a comparison of the observed and counterfactual trend *before* the shootings, which is illustrated in figure 1*B*. For this purpose, I estimate the same regression models but replace the actual with a placebo treatment indicator based on fictitious events two weeks before the actual events. Accordingly, I repeat the matching and estimation procedure used in the main analysis for these fictitious events—for example, November 28, 2011, for the event on December 12, 2011. Figures 3, 4, and 5 include the resulting local regression curves—the observed and counterfactual trend for the two weeks *before* the events. As discussed above, the trends before the shootings closely resemble each other for blacks but not in all cases for whites and Hispanics. This difference is related to the substantially smaller sample size for whites and Hispanics. These findings provide strong support for the estimation strategy indicating that the counterfactual trend based on matched stops is plausible.

In a second sensitivity analysis, I simulate fictitious events and conduct the same analyses reported in this paper for these random events (for a similar strategy, see Legewie 2013). This simulation determines the likelihood of obtaining a result similar to the observed pattern by chance alone. Each of the 800 simulation runs involves three steps. First, I simulate four events by picking a random day and time between January 2007 and December 2012 to represent the two shootings by black suspect and the two murders by nonblacks. Second, I perform the same matching procedure as in the main analysis using the three days after the fictitious events as the treatment group and all stops in the year before the events as the control group. Third, I estimate the effect of the fictitious events over three days using the same specification of the regression model described above. The results of this simulation show that the probability of observing a pattern similar to the one reported here is 0.00 for random events. The pattern is defined as (1) significant effects on the use of force against blacks for the events involving black suspects that are at least as large as the observed effect for the shooting in July 2007 and (2) nonsignificant effects for the other groups and events. Just focusing on condition 1 shows that the probability of observing a substantial and statistically significant effect for the two events involving black suspects on stops of blacks is 0.013. The results from this simulation indicate that it is highly unlikely to observe a pattern similar to the one found

for the four attacks against police officers by chance alone, providing further support for the findings presented in this article.

APPENDIX C

Robustness to Model Specification

The sensitivity of results to model specifications is a pervasive problem across the social sciences and in quantitative research more broadly (Young 2009). To address this problem and demonstrate the robustness of my findings, figure A1 shows the results from 161 different model specifications for the two main effects observed in this study—the effect of the two fatal shootings by black suspects on the subsequent use of police force against blacks. The different model specifications vary in terms of the time interval for the control group (31, 182, or 365 days), the matching ratio (5, 10, 15, 20, or 25), and the set of variables used in the matching procedure and regression analysis. The different specifications reflect trade-offs between high-quality matches, uncertainty of the estimates, and different priorities in the estimation procedure. The *X*-axis shows the absolute difference between the observed and counterfactual trend in the two weeks *before* the events (essentially a summary of the information captured with the plausibility check illustrated in fig. 1*B*). Values close to zero suggest that the specification produces a credible counterfactual trend, indicating that the estimates of the causal effect are plausible. The *Y*-axis shows the three-day effect estimates based on the same logistic regression models reported in tables 2 and 3 in terms of log odds (partly with a different set of control variables). The effect sizes reported in the main text of the article are highlighted with horizontal lines. As is common in quantitative research (Young 2009), the size of the estimated effect (*Y*-axis) and the plausibility of the estimates (*X*-axis) vary substantially across different model specifications. More importantly, all specifications lead to positive effect estimates for the two events partly with substantially larger effect sizes compared to the one reported in the main text of the article. Accordingly, the findings presented here are highly robust to different model specifications.

REFERENCES

Adams, Kenneth. 1996. "Measuring the Prevalence of Police Abuse of Force." Pp. 52–93 in *Police Violence: Understanding and Controlling Police Abuse of Force*. New Haven, Conn.: Yale University Press.

———. 2015. "What We Know about Police Use of Force." Pp. 532–47 in *Critical Issues in Policing: Contemporary Readings*, edited by Roger G. Dunham and Geoffrey P. Alpert. Long Grove, Ill.: Waveland Press.

Alpert, Geoffrey P., and Roger G. Dunham. 1999. "The Force Factor: Measuring and Assessing Police Use of Force and Suspect Resistance." In *Use of Force by Police: Overview of National and Local Data Series*, edited by Kenneth Adams et al. Washington, D.C.: U.S. Department of Justice, National Institute of Justice and Bureau of Justice Statistics.

Althaus, Scott L., and Young Mie Kim. 2006. "Priming Effects in Complex Information Environments: Reassessing the Impact of News Discourse on Presidential Approval." *Journal of Politics* 68:960–76.

Anderson, Elijah. 1990. *Streetwise: Race, Class, and Change in an Urban Community*. Chicago: University of Chicago Press.

Antonovics, Kate, and Brian G. Knight. 2009. "A New Look at Racial Profiling: Evidence from the Boston Police Department." *Review of Economics and Statistics* 91:163–77.

Baumer, Eric P., Steven F. Messner, and Richard Rosenfeld. 2003. "Explaining Spatial Variation in Support for Capital Punishment: A Multilevel Analysis." *American Journal of Sociology* 108:844–75.

Bender, Steven. 2003. *Greasers and Gringos: Latinos, Law, and the American Imagination*. Critical America. New York: New York University Press.

Black, Donald J. 1970. "Production of Crime Rates." *American Sociological Review* 35:733–58.

———. 1976. *The Behavior of Law*. New York: Academic Press.

———. 1983. "Crime as Social Control." *American Sociological Review* 48:34–45.

Blair, Irene V. 2002. "The Malleability of Automatic Stereotypes and Prejudice." *Personality and Social Psychology Review* 6:242–61.

Blalock, Hubert M. 1967. *Toward a Theory of Minority-Group Relations*. New York: Capricorn Books.

Blumer, Herbert. 1958. "Race Prejudice as a Sense of Group Position." *Pacific Sociological Review* 1:3–7.

Bourgois, Philippe. 2003. *In Search of Respect: Selling Crack in El Barrio*. Cambridge: Cambridge University Press.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82:2295–326.

———. 2015. "rdrobust: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs." *R Journal* 7:38–51.

Carnevale, Anthony Patrick, and Susan Carol Stone. 1995. *The American Mosaic: An In-Depth Report on the Future of Diversity at Work*. New York: McGraw-Hill.

Chambliss, William J. 2001. *Power, Politics, and Crime*. Crime & Society. Boulder, Colo.: Westview Press.

Cole, David. 2000. *No Equal Justice: Race and Class in the American Criminal Justice System*. New York: New Press.

Cook, Thomas D., William R. Shadish, and Vivian C. Wong. 2008. "Three Conditions under Which Experiments and Observational Studies Produce Comparable Causal Estimates: New Findings from Within-Study Comparisons." *Journal of Policy Analysis and Management* 27:724–50.

Cook, Thomas D., and Vivian C. Wong. 2008. "Empirical Tests of the Validity of the Regression Discontinuity Design." *Annals of Economics and Statistics* 91–92:127–50.

Coviello, Decio, and Nicola Persico. 2013. "An Economic Analysis of Black-White Disparities in NYPD's Stop and Frisk Program." Working Paper No. 18803. National Bureau of Economic Research, Cambridge, Mass.

Desmond, Matthew, Andrew Papachristos, and David Kirk. 2016. "Police Brutality and Crime Reporting in the Inner City: The Beating of Frank Jude." *American Sociological Review* (forthcoming).

Devine, Patricia G., and Andrew J. Elliot. 1995. "Are Racial Stereotypes Really Fading? The Princeton Trilogy Revisited." *Personality and Social Psychology Bulletin* 21:1139–50.
Disha, Ilir, James C. Cavendish, and Ryan D. King. 2011. "Historical Events and Spaces of Hate: Hate Crimes against Arabs and Muslims in Post-9/11 America." *Social Problems* 58:21–46.
Durna, Tuncay. 2011. "Situational Determinants of Police Use of Force: Who the Suspect Is vs. What the Suspect Does." Technical Report. Geneva Centre for the Democratic Control of Armed Forces, Geneva.
Eberhardt, Jennifer L., Phillip Atiba Goff, Valerie J. Purdie, and Paul G. Davies. 2004. "Seeing Black: Race, Crime, and Visual Processing." *Journal of Personality and Social Psychology* 87:876–93.
Eitle, D., S. J. D'Alessio, and L. Stolzenberg. 2002. "Racial Threat and Social Control: A Test of The Political, Economic, and Threat of Black Crime Hypotheses." *Social Forces* 81:557–76.
Fagan, Jeffrey. 2002. "Law, Social Science, and Racial Profiling." *Justice Research and Policy* 4:103–30.
Fagan, Jeffrey, Greg Conyers, and Ian Ayres. 2014. *No Runs, Few Hits and Many Errors: Street Stops, Bias and Proactive Policing*. Berkeley: Conferences on Empirical Legal Studies.
Fagan, Jeffrey, and Garth Davies. 2000. "Street Stops and Broken Windows: Terry, Race and Disorder in New York City." *Fordham Urban Law Journal* 28:457–504.
Fagan, Jeffrey, Amanda Geller, Garth Davies, and Valerie West. 2010. "Street Stops and Broken Windows Revisited: The Demography and Logic of Proactive Policing in a Safe and Changing City." In *Race, Ethnicity, and Policing: New and Essential Readings*, edited by Stephen Rice and Michael White. New York: New York University Press.
Fagan, Jeffrey, Franklin E. Zimring, and June Kim. 1997. "Declining Homicide in New York City: A Tale of Two Trends." *Journal of Criminal Law and Criminology* 88: 1277–324.
Fassin, Didier. 2013. *Enforcing Order: An Ethnography of Urban Policing*. Cambridge: Polity Press.
Fazio, Russell H., and Michael A. Olson. 2003. "Implicit Measures in Social Cognition Research: Their Meaning and Use." *Annual Review of Psychology* 54:297–327.
Friedrich, Robert J. 1980. "Police Use of Force: Individuals, Situations, and Organizations." *ANNALS of the American Academy of Political and Social Science* 452:82–97.
Garner, Joel H., and Christopher D. Maxwell. 2001. *Understanding the Use of Force by and against the Police in Six Jurisdictions in the United States, 1996–1997 (ICPSR03172-v1)*. Ann Arbor, Mich.: Inter-University Consortium for Political and Social Research.
Garner, Joel H., Christopher D. Maxwell, and Cedrick G. Heraux. 2002. "Characteristics Associated with the Prevalence and Severity of Force Used by the Police." *Justice Quarterly* 19:705–46.
Gee, Gilbert C., and Ninez Ponce. 2010. "Associations between Racial Discrimination, Limited English Proficiency, and Health-Related Quality of Life among Six Asian Ethnic Groups in California." *American Journal of Public Health* 100:888–95.
Geller, Amanda, Jeffrey Fagan, Tom Tyler, and Bruce Link. 2014. "Aggressive Policing and the Mental Health of Young Urban Men." *American Journal of Public Health* 104:2321–7.
Geller, William A., and Hans Toch. 1996*a*. *Police Violence: Understanding and Controlling Police Abuse of Force*. New Haven, Conn.: Yale University Press.
———. 1996*b*. "Understanding and Controlling Police Abuse of Force." Pp. 292–328 in *Police Violence: Understanding and Controlling Police Abuse of Force*, edited by William A. Geller and Hans Toch. New Haven, Conn.: Yale University Press.

Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. 2007. "An Analysis of the New York City Police Department's 'Stop-and-Frisk' Policy in the Context of Claims of Racial Bias." *Journal of the American Statistical Association* 102:813–23.
Gilliam Jr., Franklin D., and Shanto Iyengar. 2000. "Prime Suspects: The Influence of Local Television News on the Viewing Public." *American Journal of Political Science* 44:560–73.
Goffman, Alice. 2014. *On the Run: Fugitive Life in an American City*. Chicago: University of Chicago Press.
Graham, Sandra, and Brian S. Lowery. 2004. "Priming Unconscious Racial Stereotypes about Adolescent Offenders." *Law and Human Behavior* 28:483–504.
Greene, Judith A. 1999. "Zero Tolerance: A Case Study of Police Policies and Practices in New York City." *Crime and Delinquency* 45:171–87.
Grogger, Jeffrey, and Greg Ridgeway. 2006. "Testing for Racial Profiling in Traffic Stops from behind a Veil of Darkness." *Journal of the American Statistical Association* 101:878–87.
Harrell, Jules P., Sadiki Hall, and James Taliaferro. 2003. "Physiological Responses to Racism and Discrimination: An Assessment of the Evidence." *American Journal of Public Health* 93:243–48.
Harris, Christopher J. 2009. "Police Use of Improper Force: A Systematic Review of the Evidence." *Victims and Offenders* 4:25–41.
Harris, David A. 1999. "The Stories, the Statistics, and the Law: Why Driving While Black Matters." *Minnesota Law Review* 84:265–326.
Hastie, Trevor, and Robert Tibshirani. 1986. "Generalized Additive Models." *Statistical Science* 1:297–310.
Heckman, James J. 1998. "Detecting Discrimination." *Journal of Economic Perspectives* 12:101–16.
Holmes, Malcolm D. 2000. "Minority Threat and Police Brutality: Determinants of Civil Rights Criminal Complaints in U.S. Municipalities." *Criminology* 38:343–68.
Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79:933–59.
Imbens, Guido, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142:615–35.
Jackson, Pamela I., and Leo Carroll. 1981. "Race and the War on Crime: The Sociopolitical Determinants of Municipal Police Expenditures in 90 Non-southern U.S. Cities." *American Sociological Review* 46:290–305.
Jacobs, Bruce A. 2004. "A Typology of Street Criminal Retaliation." *Journal of Research in Crime and Delinquency* 41:295–323.
Jacobs, David, and Jason T. Carmichael. 2002. "The Political Sociology of the Death Penalty: A Pooled Time-Series Analysis." *American Sociological Review* 67:109–31.
Jacobs, David, Jason T. Carmichael, and Stephanie L. Kent. 2005. "Vigilantism, Current Racial Threat, and Death Sentences." *American Sociological Review* 70:656–77.
Jacobs, David, and Robert M. O'Brien. 1998. "The Determinants of Deadly Force: A Structural Analysis of Police Violence." *American Journal of Sociology* 103:837–62.
Kane, Robert J. 2002. "The Social Ecology of Police Misconduct." *Criminology* 40: 867–96.
Keele, Luke, and William Minozzi. 2013. "How Much Is Minnesota Like Wisconsin? Assumptions and Counterfactuals in Causal Inference with Observational Data." *Political Analysis* 21:193–216.
King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44:341–55.
King, Ryan D., and Gretchen M. Sutton. 2013. "High Times for Hate Crimes: Explaining the Temporal Clustering of Hate-Motivated Offending." *Criminology* 51:871–94.

Lacoe, Johanna, and Patrick Sharkey. 2016. "Life in a Crime Scene: Stop, Question, and Frisk Activity in New York City Neighborhoods in the Aftermath of Homicides." *Sociological Science* 3:116–34.
Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48:281–355.
Legewie, Joscha. 2012. "Die Schätzung von Kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 64:123–53.
———. 2013. "Terrorist Events and Attitudes toward Immigrants: A Natural Experiment." *American Journal of Sociology* 118:1199–245.
Legewie, Joscha, and Jeffrey Fagan. 2016. "Group Threat, Police Officer Diversity and the Deadly Use of Police Force." Columbia Public Law Research Paper No. 14-512. http://ssrn.com/abstract=2778692.
Liska, Allen E. 1992. *Social Threat and Social Control*. Albany: SUNY Press.
Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122:159–208.
Lundman, Richard J., and Robert L. Kaufman. 2003. "Driving While Black: Effects of Race, Ethnicity, and Gender on Citizen Self-Reports of Traffic Stops and Police Actions." *Criminology* 41:195–220.
MacDonald, Heather. 2001. "The Myth of Racial Profiling." *City Journal* 11:14–27.
Mays, Vickie M., Susan D. Cochran, and Namdi W. Barnes. 2007. "Race, Race-Based Discrimination, and Health Outcomes among African Americans." *Annual Review of Psychology* 58:201–25.
Mendelberg, Tali. 2008. "Racial Priming Revived." *Perspectives on Politics* 6:109–23.
Murguia, Edward, and Edward E. Telles. 1996. "Phenotype and Schooling among Mexican Americans." *Sociology of Education* 69:276–89.
Pager, Devah. 2007. "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future." *Annals of the American Academy of Political and Social Science* 609:104–33.
Pager, Devah, and Hana Shepherd. 2008. "The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets." *Annual Review of Sociology* 34:181–209.
Papachristos, Andrew V. 2009. "Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide." *American Journal of Sociology* 115:74–128.
Peffley, Mark, Todd Shields, and Bruce Williams. 1996. "The Intersection of Race and Crime in Television News Stories: An Experimental Study." *Political Communication* 13:309–27.
Persico, Nicola. 2009. "Racial Profiling? Detecting Bias Using Statistical Evidence." *Annual Review of Economics* 1:229–54.
Pettit, Becky, and Bruce Western. 2004. "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration." *American Sociological Review* 69:151–69.
Plant, E. Ashby, and B. Michelle Peruche. 2005. "The Consequences of Race for Police Officers' Responses to Criminal Suspects." *Psychological Science* 16:180–83.
Portes, Alejandro and Julia Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology* 98 (6): 1320–50.
Portillos, Edwardo L., Coramae Richey Mann, and Marjorie S. Zatz. 1998. "Latinos, Gangs, and Drugs." Pp. 156–65 in *Images of Color, Images of Crime*. Los Angeles: Roxbury.
Quillian, Lincoln. 2006. "New Approaches to Understanding Racial Prejudice and Discrimination." *Annual Review of Sociology* 32:299–328.
Quillian, Lincoln, and Devah Pager. 2001. "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime." *American Journal of Sociology* 107:717–67.

Ridgeway, Greg. 2007. "Analysis of Racial Disparities in the New York Police Department's Stop, Question, and Frisk Practices." Technical report. RAND Corporation, Santa Monica, CA.
Rios, Victor M. 2011. *Punished: Policing the Lives of Black and Latino Boys*. New York: New York University Press.
Russell-Brown, Katheryn. 1998. *The Color of Crime: Racial Hoaxes, White Fear, Black Protectionism, Police Harassment, and Other Macroaggressions*. New York: New York University Press.
Sampson, Robert J., and Janet L. Lauritsen. 1997. "Racial and Ethnic Disparities in Crime and Criminal Justice in the United States." *Crime and Justice* 21:311–74.
Schuck, Amie M. 2004. "The Masking of Racial and Ethnic Disparity in Police Use of Physical Force: The Effects of Gender and Custody Status." *Journal of Criminal Justice* 32:557–64.
Skogan, Wesley G. 2006. "Asymmetry in the Impact of Encounters with Oolice." *Policing & Society* 16:99–126.
Skolnick, Jerome H., and James J. Fyfe. 1993. *Above the Law: Police and the Excessive Use of Force*. New York: Free Press New York.
Smith, Brad, and Malcolm D. Holmes. 2014. "Police Use of Excessive Force in Minority Communities: A Test of the Minority Threat, Place, and Community Accountability Hypotheses." *Social Problems* 61:83–104.
Smith, Douglas A. 1986. "The Neighborhood Context of Police Behavior." *Crime and Justice* 8:313–41.
Smith, Tom W. 1991. *What Americans Say about Jews*. New York: American Jewish Committee.
Stewart, Eric A., Eric P. Baumer, Rod K. Brunson, and Ronald L. Simons. 2009. "Neighborhood Racial Context and Perceptions of Police-Based Racial Discrimination Among Black Youth." *Criminology* 47:847–87.
Stuart, Forrest. 2011. "Race, Space, and the Regulation of Surplus Labor: Policing African Americans in Los Angeles's Skid Row." *Souls* 13:197–212.
Stults, Brian J., and Eric P. Baumer. 2007. "Racial Context and Police Force Size: Evaluating the Empirical Validity of the Minority Threat Perspective." *American Journal of Sociology* 113:507–46.
Terrill, William, and Stephen Mastrofski. 2002. "Situational and Officer-Based Determinants of Police Coercion." *Justice Quarterly* 19:215–48.
Terrill, William, and Michael D. Reisig. 2003. "Neighborhood Context and Police Use of Force." *Journal of Research in Crime and Delinquency* 40:291–321.
Thompson, Brian L., and James Daniel Lee. 2004. "Who Cares If Police Become Violent? Explaining Approval of Police Use of Force Using a National Sample." *Sociological Inquiry* 74:381–410.
Turk, Austin T. 1966. "Conflict and Criminality." *American Sociological Review* 31:338–52.
Tyler, Tom R., Jeffrey Fagan, and Amanda Geller. 2014. "Street Stops and Police Legitimacy: Teachable Moments in Young Urban Men's Legal Socialization." *Journal of Empirical Legal Studies* 11:751–85.
Tyler, Tom R., and Cheryl J. Wakslak. 2004. "Profiling and Police Legitimacy: Procedural Justice, Attributions of Motive, and Acceptance of Police Authority." *Criminology* 42:253–82.
Vargas, Robert. 2014. "Criminal Group Embeddedness and the Adverse Effects of Arresting a Gang's Leader: A Comparative Case Study." *Criminology* 52:143–68.
———. 2016. *Wounded City: Violent Turf Wars in a Chicago Barrio*. New York: Oxford University Press.
Venkatesh, Sudhir. 2002. *American Project: The Rise and Fall of a Modern Ghetto*. Cambridge, Mass.: Harvard University Press.

Weisbuch, Max, Kristin Pauker, and Nalini Ambady. 2009. "The Subtle Transmission of Race Bias via Televised Nonverbal Behavior." *Science* 326:1711–14.
Weitzer, Ronald. 1999. "Citizens' Perceptions of Police Misconduct: Race and Neighborhood Context." *Justice Quarterly* 16:819–46.
———. 2000. "Racialized Policing: Residents' Perceptions in Three Neighborhoods." *Law and Society Review* 34:129–55.
Weitzer, Ronald, and Steven A. Tuch. 2002. "Perceptions of Racial Profiling: Race, Class, and Personal Experience." *Criminology* 40:435–56.
Welch, Kelly. 2007. "Black Criminal Stereotypes and Racial Profiling." *Journal of Contemporary Criminal Justice* 23:276–88.
Werthman, Carl, and Irving Piliavin. 1967. "Gang Members and the Police." In *The Police: Six Sociological Essays*, edited by David Bordua. New York: Wiley.
Worden, R.E. 1996. "The Causes of Police Brutality: Theory and Evidence on the Police Use of Force." Pp. 23–51 in *Understanding and Controlling Police Abuse of Force*, edited by William A. Geller and Hans Toch. New Haven, Conn.: Yale University Press.
Worden, Robert E., Sarah J. McLean, and Andrew P. Wheeler. 2012. "Testing for Racial Profiling with the Veil-of-Darkness Method." *Police Quarterly* 15:92–111.
Young, Cristobal. 2009. "Model Uncertainty in Sociological Research: An Application to Religion and Economic Growth." *American Sociological Review* 74:380–97.
Zhao, Zhong. 2004. "Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence." *Review of Economics and Statistics* 86:91–107.
Zimring, Franklin E. 2013. *The City That Became Safe: New York's Lessons for Urban Crime and Its Control*. New York: Oxford University Press.