Received: 18 August 2021 | Revised: 6 April 2023 | Accepted: 18 May 2023

DOI: 10.1111/1745-9125.12353

CRIMINOLOGY

ORIGINAL ARTICLE

Streetwork at the crossroads: An evaluation of a street gang outreach intervention and holistic appraisal of the research evidence

David M. Hureau[1] | **Anthony A. Braga**[2] | **Tracey Lloyd**[3] | **Christopher Winship**[4]

[1]School of Criminal Justice, University at Albany, SUNY

[2]Department of Criminology, University of Pennsylvania

[3]Center for Policing Equity

[4]Department of Sociology, Harvard University

Correspondence
David M. Hureau, School of Criminal Justice, University at Albany State University of New York, 135 Western Ave., Albany, NY 12222.
Email: dhureau@albany.edu

Additional supporting information can be found in the full text tab for this article in the Wiley Online Library at http://onlinelibrary.wiley.com/doi/10.1111/crim.2023.61.issue-4/issuetoc.

Generous funding support for this project was provided by The Boston Foundation. Hureau and Lloyd's participation was also partially supported by the Harvard University Multidisciplinary Program in Inequality & Social Policy (NSF-IGERT Grant #0333403).

The authors would like to thank the anonymous reviewers and the editorial team at *CRIMINOLOGY* for the thoughtful feedback that generated meaningful improvements to this article. This project benefited from a remarkable research team that included Baillie Aaron, Zinzi Bailey, Jaclyn Davis, Anthony Jack, Sonya Petri, and Joshua Wakeham.

Abstract

Spurred by the success of public health violence interventions, and accelerated by policy pressure to reduce violence without exacerbating overpolicing and mass incarceration, streetwork programs—those that provide anti-violence services by neighborhood-based workers who perform their work beyond the walls of parochial institutions—have positioned themselves as the most important non–law-enforcement violence prevention option available to urban policy makers. Yet despite their importance, the state of the field seems difficult to interpret for academics and practitioners alike. In this article, we make several contributions that bring forth new findings and deliver new perspectives on streetwork as a violence reduction strategy. First, we offer an extended analytic review of the streetwork evaluation literature that connects the study of contemporary public health violence interventions to a preceding tradition of criminologically inspired streetwork studies. Second, we present the results of an impact evaluation of StreetSafe Boston (SSB)—a multiyear streetwork intervention

Caroline Hostetter and Garrett Yursza Warfield from StreetSafe Boston also made important contributions to the quality of the data analyzed. From The Boston Foundation, we thank Paul Grogan, Robert Lewis Jr., and Jennifer Aaronson for their unwavering support of this research project. We also appreciate the help of former Boston Police Commissioner Ed Davis, former Superintendent Paul Fitzgerald, and former Boston Regional Intelligence Center Director David Carabin for their assistance in acquiring some of the data used in the evaluation. Finally, we would like to offer our most sincere thanks to the entire StreetSafe Boston team for its important work during the active years of the program and beyond.

Funding information
The Boston Foundation

that served 20 Boston gangs. We find that the SSB intervention had no detectable effect on violence among the gangs that it served. We conclude by offering a framework for understanding a field at multiple crossroads: past and present, proclaimed successes and failures, help and harm.

KEYWORDS
gang violence, streetwork, street outreach, violence prevention

1 | INTRODUCTION

Gun violence presents an unrelenting challenge to urban residents and policy makers. Despite historic declines in rates of violent crime, gun violence remains a persistent fact of life in neighborhoods in the United States, especially those affected by multigenerational legacies of racial segregation and concentrated disadvantage (Peterson & Krivo, 2010; Sampson, 2012; Sharkey, 2018). And apart from the cascading harms of gun violence—including the loss of loved ones, stress, trauma, depressed educational outcomes, and reduced economic investment—neighborhoods beleaguered by concentrated violence are typically also burdened by high rates of incarceration (Sampson & Loeffler, 2010), elevated levels of legal cynicism, and acrimonious relationships with police (Kirk & Papachristos, 2011). The policy dilemma is thus not merely responding to the complicated and historic problem of violence. It is also doing so in a way that does not reinforce patterns of concentrated punishment, further strains police–community relationships, and ideally provides material benefits for neighborhoods impacted by violence.

In this context, therefore, in the crucible of the two-sided problem of concentrated violence and concentrated punishment, policy makers have laid their hopes in "streetwork"—the provision of anti-violence services by neighborhood-based workers who perform their work on the street, beyond the walls of parochial institutions. Perhaps it is more accurate to assert that policy makers have *rediscovered* streetwork because not so long ago—in the mid-20th century—it was the preeminent policy instrument dedicated to addressing the problem of gang violence. Alongside mounting pressure to reduce violence without increasing incarceration, renewed interest in streetwork has been propelled by the ascendance of public health strategies to violence intervention that have adopted the approach. Following the launch of the model public health program, Chicago's Cure Violence circa 1999, streetwork programs proliferated across the United States and the globe, including dozens of Cure Violence replications and many more streetwork interventions inspired by the public health turn in violence prevention (Butts et al., 2015; Papachristos, 2011; Skogan et al.,

2009). In a policy domain still coming to grips with the harms associated with overpolicing and mass incarceration, streetwork programs—representing a stark alternative to punitive responses to violence—have positioned themselves as the most important non–law-enforcement violence prevention option available to urban policy makers (e.g., Gravel et al., 2021; Roman, 2021). And amidst sharp public pressure to redistribute public safety resources away from police, streetwork programs constitute the rare programmatic exemplar capable of (at least in theory) supplanting a core police function in addressing violence.

Yet despite the prevalence and heightened societal significance of streetworker approaches, the state of the field seems difficult to interpret for academics and practitioners alike. Depending on how one looks—or how one wants to see—the recent research literature into the efficacy of streetwork has offered evidence that such interventions are effective (e.g., Maguire et al., 2018), generate mixed results (e.g., Butts et al., 2015), or can even be harmful (e.g., Wilson & Chermak, 2011). Can this growing body of research evidence be interpreted in a way that provides clarity to this apparent haze of disconnected and conflicting empirical results? We believe so.

In this article, we make several complementary contributions that bring forth new findings and deliver new perspectives on streetwork as a violence reduction strategy. We first offer an extended analytic review of the streetwork evaluation literature that connects the study of contemporary public health violence interventions to a preceding tradition of criminologically inspired streetwork studies. Next, we present the results of a rigorous impact evaluation of StreetSafe Boston (SSB)—a multiyear streetwork intervention providing violence mediation and social service provision to active gang members associated with 20 Boston gangs. Using a quasi-experimental design with multiple matching techniques to compare trends in fatal and nonfatal shootings by gangs that received SSB services relative to trends in fatal and nonfatal shootings by matched comparison gangs, we found that the SSB intervention had no detectable effect on violence among the gangs that it served. Contemporary evaluations of streetwork initiatives have universally measured program impact at the area level. Here, we advance streetwork program evaluation methodology by assessing the programmatic impact of SSB at the *gang level*; among other benefits, this approach yields direct measures of program impact closely bound to program theory and activities, while avoiding the many practical and methodological challenges involved in developing area-level counterfactuals (see Roman et al., 2018).

Finally, we return to historic theoretical and program evaluation evidence to interpret our sobering findings alongside the results of other contemporary studies. We argue that by taking the long view and drawing on insights from a tradition of criminological and sociological streetwork research, the apparently mixed results of recent streetwork evaluations leave a different—and less positive—impression. We conclude by developing a new practical framework for understanding a field at multiple crossroads: past and present, success and failure, help and harm.

2 | GANG OUTREACH PROGRAMS IN HISTORICAL AND THEORETICAL CONTEXT

The roots of gang outreach programs extend at least as far back as the 19th century, and inspection of them reveals important turning points in the development of criminological theory and its intersection with public policy toward gangs. The forerunners of what would eventually come to be known as "gang outreach workers" or "streetworkers" were the 19th century "boys workers" and "settlement house workers," whose efforts were defined by the emergence of new institutions—especially the settlement house and the boy's club—dedicated toward meeting the

societal challenges generated by the rapid urbanization of the period. By the 20th century, such community work was increasingly influenced by social science, especially in large cities such as Chicago, where Jane Addams and other settlement workers brought their research and applied work into conversation with the empirical study of the city conducted by Park and Burgess—and especially their students Frederic Thrasher and, later, Clifford Shaw (Addams, 1909; Finestone, 1976; Park & Burgess, 1967/1925). These early Chicago School thinkers considered the gang to be a critical social force not only in the nascent social problem of juvenile delinquency but also in the overarching scientific study of neighborhoods (Bursik & Grasmick, 1993; Shaw & McKay, 1942; Thrasher, 2013/1936).

This scholarly view projected far beyond the university, however, proving consequential for its influence on policy and practice, as well as for its role in advancing criminological and sociological thought. By the mid-20th century, streetworker (or "detached worker") programs had established themselves as the default policy response to the social problems posed by street gangs (Finestone, 1976; Klein & Maxson, 2006). Guided by the Chicago idea that gangs could be transformed from crime and delinquency through the intervention of outreach workers who repurposed the proto-organizational structure of gangs toward pro-social ends benefiting young men and their broader communities (Thrasher, 2013), urban policy makers during the mid-century subscribed to a "transformational" approach toward street gangs (Klein & Maxson, 2006).

This transformational approach persisted even as researchers and policy makers, responding to societal shifts, defined new ends for streetworker programs. Whereas the pre–World War II streetworker served gangs to facilitate the adjustment of newcomers to the transitional areas of American cities, outreach workers in the 1960s were increasingly called on to leverage connections with street gangs in racially segregated areas of the city to reduce the likelihood of rioting, redress social isolation, provide opportunities, advance welfare state provision, and jumpstart community organization and empowerment efforts (Finestone, 1976; Hinton, 2016). This shift in focus was often adopted fluidly as organizational structures created to administer gang outreach programs later became recipients of the federal government's social welfare investment during the War on Poverty, with gang transformation understood as a key pathway for overarching community transformation (Dawley, 1992; Hinton, 2016; Woodsworth, 2016).[1] Through the 1960s, then, an academic and policy-making consensus had developed that understood gangs as products of their social conditions, regarded them to be socially important to the resolution of community problems, and considered them to be fundamentally *redeemable*, with the outreach worker serving as the primary mechanism for redemption. Under this consensus, streetworker programs proliferated across the United States's major cities during the middle of the 20th century (Cohen & Short, 1958; Klein, 1971; Kobrin, 1959; Miller, 1962; New York City Youth Board, 1960).

The prominent position afforded to streetworker programs in mid-20th century urban social policy may surprise the contemporary reader more familiar with the gang outreach programs that developed in response to the gang violence of the 1980s and 1990s. Streetworker programs fell from their favored status in the 1970s and early 1980s due to parallel developments in the domains

[1] The logic of gang transformation developed in the 1960s was less indebted to the Chicago School than to Lloyd Ohlin and Richard Cloward. Ohlin served on President Kennedy's (1961) Committee on Juvenile Delinquency and Youth Crime and helped set in motion a 1960s policy agenda that addressed urban Black social isolation via anti-delinquency programming embedded within community transformation (see Hinton, 2016). Despite key differences, Cloward and Ohlin's (1960) opportunity theory shared the view with the earlier Chicago School that the causes of crime were to be sought at the community level and that community institutions were critical to delinquency control.

of research and policy. In the academic realm, by the 1970s, the community-based sociology of the Chicago School—with its interest in gangs as important neighborhood social forces—was losing influence in academic and policy circles to criminological research concerned with predicting individual propensity for offending (Abbott, 1997; Klein & Maxson, 2006). In the policy domain, by the 1970s criminal justice policy—fueled by the "nothing works" findings of Robert Martinson (1974) and in step with broader trends in social policy—pivoted away from the idea of rehabilitating crime-involved individuals in favor of a novel punitive regime characterized by principles of deterrence, harsher penalties for crime, and increasing rates of incarceration (Garland, 2001; Travis et al., 2014; Western, 2006). Gangs themselves underwent a makeover in public and policy discourse as gang violence—especially in segregated and impoverished neighborhoods—became an object of concern for national law enforcement in addressing the crime and unrest at the heart of the nation's "urban problems" (Hinton, 2016). As violent crime rates climbed from the 1960s and into the 1980s, the idea of transforming gangs—through gang outreach and other social policy initiatives—was discarded in favor of punitive approaches emphasizing police and prosecution programs, suppression of violence, and intensive gang surveillance and supervision (Klein & Maxson, 2006; Maxson & Klein, 1983).

Streetworker programs would reemerge in the late 1980s and early 1990s in response to the significant increases in serious gang violence in many U.S. cities, but they would no longer be the favored policy response to addressing the problem of street gangs—that job had been usurped by law enforcement. As broad-based, gang-joining *prevention* programs expanded during the 1980s and 1990s, streetworker programs—practicing gang *intervention*—found a new niche as 1) community-based complements to law enforcement strategies and/or as 2) components of larger community-based "comprehensive" gang control efforts (Spergel, 1995). Streetworker programs embedded in such comprehensive efforts have typically taken a back seat to police gang suppression and have struggled to provide appropriate services for gang members, establish meaningful gang intervention alternatives to police enforcement, and develop streams of funding to support their work (Kennedy, 2011; Klein & Maxson, 2006; Spergel, 2007).

The most recent chapter in the theory and practice of streetworker programs has been authored by writers unaffiliated with sociology, criminology, or the tradition of street outreach informed by these disciplines. Led by the efforts of Cure Violence (formerly CeaseFire Chicago, established in the late 1990s), public health scholars and practitioners advancing the understanding of violence as a "public health issue," offered a new disciplinary perspective that promoted intervening in street violence as if it were infectious disease. The public health entry into the field of streetwork was momentous not only because it offered an alternative to tired policy debates regarding the nature of violent crime and how responsibility for it should be apportioned, but also because it generated new interest in street violence intervention, attracted new support from policy makers (especially those looking for novel policy alternatives to enforcement), engaged a large cohort of new researchers (especially those in medicine and public health), and mobilized new institutional streams of funding to support practice and research (see Hemenway, 2006; Papachristos, 2011).

Although sources of support for streetwork initiatives and the social problems they have addressed have shifted over time, these programs have been motivated by a common foundational presupposition: "Because gang members do not ordinarily respond well to standard agency programs inside the agency walls, it is necessary to take the programs to the gangs" (Klein, 1971, p. 46). But what does the historical record reveal about how effective such programs have been in taking their work to the street?

2.1 | Early/Classic Streetworker Program Evaluations

Four high-quality evaluations inform our understanding of the effectiveness of the early (pre-1970) streetworker programs (see table 1). The first was conducted by Walter Miller in Boston's Roxbury neighborhood between 1954 and 1957 and dubbed the "'Total Community' Gang Control Project" as it sought to reduce neighborhood adolescent delinquency by intervening at three ecological levels: the community, the family, and the gang. The project's main intervention was aimed at gangs, however, featuring seven professionally trained streetworkers assigned to 21 Roxbury gangs, with seven of the gangs receiving "intensive" attention from at least one streetworker. In line with the transformational approach to gang intervention at the time, streetworkers were directed "to contact, establish relations with, and attempt to change resident gangs" (Miller, 1962, p. 169). Using delinquency data from streetworker reports, as well as official agencies, Miller reported that the treatment gangs showed no improvement on a variety of delinquency measures when compared with counterpart control gangs. In fact, increases in delinquency were detected among several categories of offending—particularly increases in serious offending, among younger gang members, and among boys relative to their female counterparts (Miller, 1962).

The Chicago Youth Project was an initiative of the Chicago Boys' Clubs from 1960 through 1966, evaluated by the Institute for Social Research at the University of Michigan. The intervention followed the same logic of Miller's "total community" approach, but it placed greater emphasis on community organization and outreach to nongang youth and made greater use of data and research to refine its outreach efforts (Gold & Mattick, 1974; Klein, 1971; Spergel, 1995). The targeted group-based streetworker interventions did not yield meaningful results in the prevention of delinquency; in fact, individuals closest to their assigned streetworkers demonstrated the greatest delinquency increases. Although the project showed some promise in raising educational expectations, providing employment opportunities, and reengaging school dropouts, evaluation results suggested that youth living in the target intervention neighborhoods were slightly worse off on a variety of delinquency and pro-social indicators than were youth living in control neighborhoods (Gold & Mattick, 1974; Spergel, 1995).

The Los Angeles Group Guidance Project was sponsored by the Los Angeles County Probation Department and operated between 1961 and 1965 (see Klein, 1969, 1971). The Group Guidance Project employed a transformational streetworker approach to engage four majority-Black Los Angeles gangs, broken into 16 subgroups comprising approximately 800 members in total. Relying heavily on a streetworker-led group programming approach (which featured limited employment, educational, and community-organizing content), Klein found that the program was associated with a significant increase in delinquency among the gang members served. He argued that gang delinquency increased as a result of 1) increased gang cohesion brought about by large amounts of group-based programming and of 2) increased gang recruitment fueled by this group programming (Klein, 1969, 1971).

Klein's (1971) "Ladino Hills Project" of 1965–1966 was designed to redress the shortcomings of the Group Guidance Project. This Los Angeles project engaged a single Mexican gang for 18 months, with a 6-month follow-up period of data collection. Reasoning that the increased gang cohesion—and delinquency—produced by the Group Guidance Project was the result of group programming, the Ladino Hills project dispensed with all group programming in an effort to reduce gang cohesion and subsequent gang delinquency. In place of the group-based programming, Ladino Hills streetworkers, aided by research staff, outreached to individuals and preexisting cliques, promoting educational and employment opportunities to wean away members from the gang. Even though the project did not significantly affect rates of offending for gang

TABLE 1 Summary of Street Outreach Evaluations

Period	Intervention	Authors	Year	Overall Finding	Outcome of Interest	Implementation	Unit	Theory	Comparison	Considerations
Classic	Boston: Total Delinquency Control Project	Miller	1962	Null / Harmful	Delinquency	Full	Individual / Gang	Gang Transformation	Yes*	Study corrects prior results of two previously published papers that reported the project had produced limited reductions in delinquency
	Los Angeles: Group Guidance Project	Klein	1969 / 1971	Harmful	Delinquency	Full**	Individual / Gang	Gang Transformation	No	Pre/post design with steps taken to account for changes in age-related offending; those served most intensively showed greatest increases in delinquency
	Los Angeles: Ladino Hills Project	Klein	1971	Mixed / Some Beneficial	Delinquency	Full	Individual / Gang	Gang Transformation	No	Approach was developed in contradistinction to Group Guidance Project, with a specific aim of reducing gang cohesion through terminating group programming
	Chicago: Chicago Youth Project	Gold & Mattuck	1974	Null / Harmful	Delinquency	Full	Individual / Gang / Area	Gang Transformation / Opportunity Provision	Yes	Gang- and individual-level street intervention with educational and job placement in two Chicago neighborhoods did not reduce delinquency or improve pro-social outcomes compared to youth in comparison areas
Contemporary	Chicago: Little Village Project	Spergel	2007	Mixed / Beneficial	Arrest (Violence)	Mixed, Full for Streetwork & Suppression, Failed Comprehensive Coordination	Individual / Gang / Area	Comprehensive: Streetwork, Suppression, Services, Mobilization	Yes*	Program model used police suppression alongide streetwork, service provision, and community mobilization; evidence of program impact on violent and overall arrest for program participants

(Continues)

TABLE 1 (Continued)

Period	Intervention	Authors	Year	Overall Finding	Outcome of Interest	Implementation	Unit	Theory	Comparison	Considerations
Contemporary	Chicago: CeaseFire	Skogan et al.	2009	Null****	Violence	Mixed, Signs of Implementation Failure	Area	Public Health	Yes	Study found program effects on shooting outcomes for approximately half of treatment areas; funding / staffing challenges suggest uneven implementation across sites / time
	Newark: Ceasefire	Boyle et al.	2010	Null	Violence	Unknown***	Area	Public Health & Focused Deterrence	Yes	Used trauma center admissions as outcome of interest; blend of public health and focused deterrence approaches
	Pittsburgh: One Vision One Life	Wilson & Chermak	2011	Harmful	Violence	Signs of Implementation Failure	Area	Public Health & Strategic Problem Solving	Yes	Strong signs of implementation failure and programmatic drift to non-target populations; evidence program increased violence in target neighborhoods
	Baltimore: Safe Streets	Webster et al. & Buggs et al.	2013 / 2018 / 2022	Mixed / Harmful	Violence	Mixed, Signs of Implementation Failure	Area	Public Health	Yes	Strong fidelity to program model but implementation affected by staffing challenges; research design accounted for simultaneous violence prevention programs in target / comparison areas; revised analysis suggests more evidence of harm than benefit
	Brooklyn: Save Our Streets	Picard-Fritsche & Cerniglia	2013	Null****	Violence	Unknown***	Area	Public Health	Yes	Limited programmatic dosage (96 people, less than 1 hour per month of service) generates skepticism of large reported area-level decreases in violence

(Continues)

TABLE 1 (Continued)

Period	Intervention	Authors	Year	Overall Finding	Outcome of Interest	Implementation	Unit	Theory	Comparison	Considerations
Contemporary	Phoenix: Truce Project	Fox et al.	2015	Mixed / Harmful	Violence	Mixed	Area	Public Health	Yes	Program had success in reaching an appropriate target population but may not have implemented full CV model; Evaluation design lacked appropriate comparison neighborhood(s)
	Port of Spain: Project REASON	Maguire et al.	2018	Beneficial	Violence	Mixed	Area	Public Health	Yes	Program was associated with large decreases in violence across several outcomes measures; Very strong research design but program was weakly implemented, raising questions regarding the program's role in generating observed decreases

*Comparison groups were used for some but not all outcomes.
**Outreach and group programming were fully implemented, whereas other areas (e.g., employment) were less successful.
***Authors suggested full implementation but difficult to discern based on data/description provided.
****Authors reported/emphasized positive results.

members, the authors reported that overall crime by the gang was reduced by means of curtailing gang membership and gang joining.

2.2 | Contemporary Streetworker Program Evaluations

Although a variety of streetworker and gang outreach programs proliferated during the course of the late 1980s and 1990s, most of these programs went unevaluated. Irving Spergel's Chicago-based Little Village Gang Violence Reduction Project (GVRP) is the notable exception, representing a transitional moment in streetwork practice documented by thoughtful evaluation research (Spergel, 2007; Spergel & Grossman, 1997). Critical of narrowly focused "streetwork" interventions, GVRP instead aspired to be a comprehensive community gang strategy focused on enhancing community–institutional capacity for effectively responding to gang problems. Distinct from other programs covered in this review, GVRP's comprehensive approach actively involved law enforcement suppression—in fact, the program was sponsored and primarily administered by the Chicago Police Department. Through a team approach involving police, probation, outreach workers, and neighborhood organizations, the program integrated social intervention, suppression, opportunities provision, and community mobilization to influence the behavior of two almost exclusively Mexican gangs—representing approximately 200 members—in Chicago's Little Village neighborhood. GVRP's impact can be assessed along two distinct axes: programmatic impact and policy impact. Owing to the expansiveness of the intervention and the vastness of Spergel's analyses, GVRP's programmatic impact is challenging to summarize; still, the evaluation's most rigorous analysis demonstrated statistically significant reductions in violent arrests among individuals involved in the program when compared with comparable untreated populations. Accompanying analyses also reported decreases in self-reported violent offenses at the individual level but also *increases* in gang-level violence for the project's two focal gangs during the program period. With respect to its policy impact, GVRP served as the model for a line of 1990s Office of Juvenile Justice and Delinquency Prevention (OJJDP) "comprehensive" programs focused on gang prevention, intervention, and suppression. As Spergel and his colleagues lamented, however, the federally supported programs that would follow would not adequately invest in street outreach approaches and defaulted to following stand-alone suppression or service provision approaches (Spergel, 2007; Spergel et al., 2006). In retrospect, Spergel's comprehensive approach, despite its qualities, seems to represent a policy road not taken.

Most of the recent interest in the evaluation of streetwork programs has been driven by Cure Violence (formerly known as CeaseFire Chicago) and the various replication demonstrations it has inspired. Cure Violence has become the exemplar public health street violence intervention in the United States. The heart of its intervention is streetwork, but the streetworker function has become specialized in the program's model. One set of outreach workers (typically professionally trained) maintains contact with "at-risk" individuals and aims to broker services and pro-social opportunities, whereas (typically formerly street-involved) "violence interrupters" are freed to focus solely on the mediation of violent disputes and the prevention of retaliation. Other facets of Cure Violence's programming extend beyond the streetworker approach and reveal its public health underpinnings. Specifically, the program aims to promote broad-based, population-level shifts in attitudes toward the acceptability of the use of violence, akin to previous public health campaigns targeting issues such as smoking cessation and seatbelt use. Furthermore, the Cure Violence model prescribes the promotion of various community-level campaigns aimed at both community attitude change and the enrollment and mobilization of community members

in responding to high-profile violent events (Butts et al., 2015; Papachristos, 2011; Skogan et al., 2009; Wilson & Chermak, 2011).

Skogan et al.'s (2009) evaluation of CeaseFire Chicago, employing a neighborhood-level, quasi-experimental longitudinal analysis of shootings, reported that approximately half of CeaseFire's treatment neighborhoods experienced significantly fewer shootings (ranging from 16 percent to 34 percent) versus appropriate comparison neighborhoods. Additional analyses of the program's effects reveal similarly uneven indicators of programmatic success; CeaseFire demonstrated a positive impact on gun homicide in only one of seven treatment neighborhoods (and possible backfire effects in two others), achieved reductions in hot spots in three or four of seven neighborhoods, and positively influenced gang homicide networks in only two of eight neighborhoods. Although not emphasized in the report, several outcome measures seem to favor the comparison neighborhoods over the treatment neighborhoods. The research team's thorough evaluation noted implementation shortcomings due to volatility in program staffing and funding support, lack of data on program dosage and staff activity, and the challenges associated with detecting a program effect in the midst of a long-lasting crime drop in Chicago.

Newark's "Operation Ceasefire," evaluated by Boyle et al. (2010), integrated elements of group-based focused deterrence (e.g., see Braga et al., 2014) with Chicago CeaseFire's public health approach, using outreach workers and public education and organizing strategies to disrupt ongoing gun violence in a two-square-mile treatment area referred to as the "Ceasefire Zone." The evaluation examined the program's impact on gunshot wound admissions to Newark's Level 1 Trauma Center during a 3-year period, representing an important alternative to most streetwork evaluations, which use police data to measure program effects. Using a pre–post interrupted time-series design to assess impact in the Ceasefire Zone compared against a matched comparison neighborhood and citywide violence trends, the evaluation team concluded that the program did not significantly impact either trauma center admissions or gunshot wound prevalence in hot spot areas.

Pittsburgh's "One Vision One Life" was inspired by CeaseFire Chicago but deviated from its model due to a host of local and political factors beyond the scope of this review (see Wilson & Chermak, 2011). Operating in the mid-2000s, One Vision tasked former street-involved streetworkers with gathering meaningful intelligence into the nature of ongoing gang conflicts to 1) prevent the escalation of minor disputes from turning violent; 2) develop "behind-the-scenes" responses to all homicides in the program target area; and 3) connect high-risk individuals to pro-social and employment services. Evaluators found that One Vision rarely made effective use of program and crime data to organize responses to violent incidents and its streetworkers rarely engaged in "problem-solving" activities to prevent retaliation in the wake of a violent incident. The project was also characterized by implementation issues, most notably, that streetworkers were engaging nongang youth more than they were intervening with gang-involved individuals embedded in cycles of gun violence. Using a quasi-experimental, neighborhood-level, difference-in-difference analysis, One Vision evaluators found that the program had no effect on the incidence of homicide and was associated with a statistically significant *increase* in aggravated assaults and gun assaults in its target neighborhoods.

Baltimore's "Safe Streets" represented a rigorous CeaseFire (Cure Violence) replication that was originally piloted in several Baltimore neighborhoods between 2007 and 2010, and has continued through the present day. Among the contemporary interventions included in this review, Safe Streets is notable for having been rigorously evaluated at least three different times, using several different methods to estimate program effects. Overseen by the Baltimore City Health Department, in conjunction with the Chicago Project for Violence Prevention (the creators of

CeaseFire Chicago), Safe Streets received ample training and technical support to ensure that the program was implemented faithfully to the CeaseFire model. The program was implemented in four of Baltimore's highest violence neighborhoods, although a fifth program site experienced implementation failure and was later not included in evaluation results. An evaluation by Webster et al. (2013) made use of a longitudinal data set of homicides and nonfatal shootings to analyze the effect of Safe Streets on neighborhood violence across targeted neighborhoods. Only one program site (Cherry Hill) evidenced significant declines in both homicides and nonfatal shootings, although two others showed overall reductions in gun violence (driven by decreases in nonfatal shootings). A single program site (Madison-Eastend) experienced a statistically significant increase in homicides, accompanied by a significant decrease in nonfatal shootings.

Two ensuing evaluations of Baltimore's Safe Streets broadly confirmed the program's null or mixed results. Webster and colleagues (2018) built on their original evaluation by analyzing newer program sites (for a total of seven) and extending the intervention period to examine program impact on homicides and nonfatal shootings between 2007 and 2017. The investigators found no evidence of an aggregate program impact; results for nonfatal shootings were mixed and statistically insignificant across sites, whereas results for homicides were null but included both statistically significant decreases and increases. Recognizing the challenges involved in identifying appropriate comparison units for Safe Streets areas, the same research team employed synthetic control methodology to revisit earlier estimates of program effects (Buggs et al., 2022). This synthetic control analysis once again confirmed the lack of a positive program impact, revealed mixed outcomes across sites, but importantly revised earlier findings by concluding that the site-specific impacts suggested "more evidence of harm than benefit" (Buggs et al., 2022). Further dampening earlier optimism, the authors took advantage of their extended observation period to test the duration of program impacts, only to find that beneficial impacts reported in earlier Safe Streets evaluations had attenuated over time.

"Save Our Streets" (SOS) was another Cure Violence replication project evaluation, implemented during the course of 29 months in the Crown Heights neighborhood of Brooklyn, NY. Although the project closely followed the Cure Violence public health model, SOS operated at a much smaller scale, employing four outreach workers who served 96 clients—only 68 percent of whom were classified as "high risk" for involvement in gun violence (Picard-Fritsche & Cerniglia, 2013). Despite the modest scale of the SOS street outreach—although in keeping with the CeaseFire model, SOS conducted several community norm-changing campaigns—evaluators reported that SOS was associated with statistically significant violence reductions at the neighborhood level. Using an interrupted time-series method featuring matched comparison neighborhoods, evaluators credited SOS with bringing about an approximate 20 percent reduction in gun violence relative to comparison neighborhoods. We view these findings as optimistic. First, the reductions in gun crime were hypothesized to be the product of intervention with 96 individuals in a police district of more than 96,000 people.[2] Second, the evaluation design could not rule out alternate policing and social service interventions or other demographic and policy trends that may have influenced the results.

The Phoenix Truce Project was another Cure Violence replication project operational in the city's Hermoso Park neighborhood (Fox et al., 2015). The research team found that the program was fully implemented in terms of service provision—meaningfully reaching a heavily gang-involved population at high risk for violence—but lacked in some other dimensions of

[2] SOS streetworkers spent an average of 20 hours with their 96 participants during the 29-month evaluation period—another indicator of the modest dosage of the intervention (Picard-Fritsche & Cerniglia, 2013).

TABLE 2 Two Eras of Street Outreach Research

	Classic	Contemporary
Outcome	**Delinquency**	**Gun Violence**
Intervention theory	Gang transformation	Public health / violence interruption
Implementation	Full; Often aided by researchers	Mixed; often signs of implementation failure
Key findings	Null or increased delinquency	Mixed; proclaimed "successes"
Unit of analysis	Groups / individuals	Areas / neighborhoods

implementing the Cure Violence model (e.g., establishing a community board and engaging faith-based institutions). The team's impact evaluation revealed that the Truce Project reduced assaults and all violent incidents (both categories were defined expansively) in its treatment area relative to comparison neighborhoods, while unfortunately *increasing* shootings and shots-fired incidents—the outcome most closely tied to programmatic activity.

Maguire et al.'s (2018) evaluation of Port of Spain's (Trinidad) Project REASON stands out for both its methodological rigor and its unambiguously positive impact findings. The project was a Cure Violence adaptation driven by street outreach, community mobilization, public education, and institutional collaboration, and it was implemented in 16 high-violence communities in the Trinidadian capital. Using three distinct outcome measures (official reports of violence, police calls for service, and hospital admissions) and three quasi-experimental methods to test for impact (difference-in-difference, synthetic control, and interrupted time series), the evaluation team concluded that Project REASON generated large and statistically significant reductions in violence in its treatment communities. Representative of many of the challenges in evaluating programmatic impact in this research domain, Maguire et al.'s (2018) robust analyses suggest that *something* occurred to reduce violence in Port of Spain's high-violence communities during Project REASON's implementation—just not that it was Project REASON itself. The evaluation team found that the project was only partially implemented. Specifically, the authors found limited evidence that program staff were responsive to incidents of violence in the target communities; most importantly, they did not demonstrate clear connection with the people most likely to be involved in violence. During the course of the 2-year evaluation period, Project REASON outreach staff recruited only 64 participants and slightly more than 40 percent of these participants were classified as "high risk" for gun violence.

2.3 | Reassessing the Landscape of Streetworker Programs

Comparing the general features of evaluation efforts across the two time periods, here described as "classic" and "contemporary," yields some useful insights, as presented in table 1,2.

Although outreach programs have pursued a multitude of programmatic goals (Spergel, 1966), classic streetworker programs clearly were principally assessed on their ability to control gang delinquency (especially gang fighting), whereas contemporary programs measure success based on their ability to reduce gun violence. What is more, classic programs of the mid-20th century theorized that delinquency could be controlled by transforming gangs into productive units of social organization, whereas contemporary programs have theorized that gun violence can be reduced by means of violence interruption and community mobilization via norm-changing campaigns that seek to define violence as socially unacceptable. Whereas implementation failure seems to be a common problem for contemporary streetworker programs, the results of classic streetworker

evaluations were not obviously influenced by implementation failure. Klein (1971) argued that the presence of embedded researchers in these classic studies, as a practical matter, forced goal specification and strong implementation efforts on program administrators. Thus, although contemporary streetworker evaluations distinguish themselves for their technical sophistication, they are also characterized by greater distance from the on-the-ground work when compared with earlier evaluations. This distance has generated uncertainty regarding the quality of program implementation and has made it difficult to identify mechanisms, leaving several contemporary researchers struggling to explain the results produced by their econometric models.

The findings of early streetworker program evaluations are clear that such programs do not decrease—and can even increase—delinquency. Evaluations of contemporary programs concerned with violence prevention report mixed results, although our review raises the question of how much more encouraging the results from this period should be understood relative to those of the classic period. And finally, contemporary streetworker evaluations have uniformly tested program impacts at the area level. Although some early evaluations also employed neighborhood-level analyses, they also sought to document outcomes among the specific *groups* and *individuals* that served as the targets of their intervention.

This last point presents the greatest challenge for contemporary evaluations of streetworker programs, particularly those—like public health interventions—that seek to reduce neighborhood levels of violence: For reported neighborhood- or population-level decreases to be viewed as credible, these programs must first demonstrate behavioral change among the people or groups theorized to be driving neighborhood and population rates of violence. Our review of the evidence suggests that no contemporary streetworker program has taken this step to date, leaving the field with two overarching scholarly and practical shortcomings. First, by focusing exclusively on area-level outcomes, researchers have missed a crucial opportunity to develop knowledge on how such interventions influence the behavior of those most proximate to them. And second, policy makers and practitioners—at a time of intense need for community-based violence interventions—may have been oversold on the efficacy of certain intervention models, leading to an unnecessary isomorphism in the development of outreach and intervention practice over the last two decades.

3 | THE STREETSAFE BOSTON INTERVENTION

Operational from mid-2009 through 2014,[3] SSB was a programmatic initiative of The Boston Foundation (TBF) dedicated to reducing serious violence in Boston. Although SSB launched from a community-engaged planning process that identified five "Focus Areas" of concentrated gun violence where the intervention would be targeted, the program did not pursue the typical ambition of reducing citywide violence through place-based strategies.[4] Instead, SSB adopted a strategy of intervening upon 20 gangs (associated with these areas) by means of street outreach, mediation,

[3] The program was designed as a 4-year investment (mid-2009 through mid-2013) of TBF but was later extended an additional 18 months until programmatic activities could be brought under the umbrella of Boston's Centers for Youth and Families, with transitional funding provided by TBF. This article assesses the programmatic impact of the original 2009 through 2013 period.

[4] The selection of these Focus Areas was a political process rather than a data-driven one. Many citywide and neighborhood actors sought to influence TBF's decision of where to allocate scarce violence prevention resources (for more on the politics of violence prevention, see Vargas, 2016). In everyday practice among SSB staff, however, the Focus Areas had little meaning apart from their status as an administrative unit for the organization of work. That is, SSB organized supervisory duties, service opportunities, and meetings among the people serving gangs associated with a particular Focus Area—but intervention work itself was targeted at the gang level.

and the provision of social services. Theorizing that citywide levels of violence would decline if it could disrupt the ongoing cycles of retaliatory gun violence among a strategic cohort of the city's gangs, SSB carefully selected the gangs that it served. Drawing on the expertise of the Boston Police Department (BPD), partners from the city of Boston, researchers, and service practitioners with experience working with local gangs, TBF staff selected gangs into the intervention based on each gang's embeddedness in one of SSB's five Focus Areas, its historic level of violence, and its recent involvement in serious violence—including perceived risk for long-term engagement in violence.

The SSB intervention was guided by a theory of pragmatic and strategic problem-solving. Although such an approach has been uncommon among contemporary street outreach programs (Wilson & Chermak, 2011), it was familiar to SSB leadership, some of whom were influential street-workers during the 1990s. One cornerstone objective of this approach was to establish the capacity to provide concentrated street outreach resources *when* and *where* they were most needed. SSB consciously sought to create an "ideal" streetworker program that was singularly focused on addressing the gangs and conflicts that produced much of the city's violence.

The problem-solving capacity of SSB hinged on the strategic concentration of resources. Not only did SSB intervene upon 20 high-violence gangs, the program assigned each gang its own streetworker. This arrangement was theorized to have several problem-solving advantages. First, should a violent event occur, streetworkers could immediately diagnose the incident and its potential for continued retaliatory violence. The development of sound information regarding what happened in an incidence of serious violence is an underrecognized dimension of streetwork and a critical first step for ensuing mediation efforts. Second, unlike many area-based outreach programs, SSB streetworker embeddedness would afford a gang-specific social capital to be leveraged in service of immediate and long-term violence intervention.

StreetSafe Boston further combined the logic of targeted concentration of outreach services with sophisticated data monitoring and analysis. Streetworkers received real-time notification of shootings and violent incidents from the BPD and local hospitals, and the program supported a dedicated data manager who was afforded access to BPD data tracking gang-related shootings to monitor SSB's effectiveness at managing active conflicts and reaching its violence reduction targets. Taken together, SSB was logically organized to diagnose the sources of violent conflict, monitor these conflicts and the impact of their intervention efforts, and implement strategic mediations designed to prevent and disrupt cycles of retaliatory gun violence among a cohort of the city's most violent gangs.

StreetSafe's problem-solving approach comprised two complementary—if unevenly implemented—violence reduction strategies performed by two distinct sets of actors. The first was the organization's "streetworker strategy," which addressed the immediate and near-term dynamics of ongoing violent conflicts. SSB streetworkers—numbering ~20—were mostly formerly gang- and street-involved men in their 20s through their 40s, although some women and youth organizers also served in this position. Instead of merely responding to violent incidents, SSB streetworkers were expected to build connections to and knowledge of the gangs they served, establishing relationships with hard-to-reach individuals and subgroups while learning the gang's social history, including key events and its network of conflicts and alliances. From this posture, SSB streetworkers launched their short- and medium-term problem-solving work. In the short term, streetworkers responded to every gun violence incident involving their gang, most commonly to deescalate gatherings at the shooting scene or hospital emergency room and implementing safety plans in the wake of these events for the actors most likely to retaliate or be victimized. But in between violent events, SSB streetworkers leveraged their understanding of their gangs to *anticipate* and intervene in violence before it happened; streetworkers tracked

social media for signs of emerging tensions, "airlifted" clients out of risky situations, planned trips and events around the death anniversaries of gang members, and managed the risk associated with the places where gang members congregated. Thus, beyond activities that sought to prevent retaliation, SSB developed a set of practices dedicated to *protecting* the gang members they served, conceiving of their clients not merely as shooters but also as people uniquely vulnerable to victimization. And in further distinction to standard "violence interruption" work, SSB streetworkers used their intimate knowledge and group-specific problem-solving experience to build toward the long-term resolution of conflicts. Workers referred to this bridging of short-term problem-solving and long-term strategic work to stop cycles of violence as "bringing [gangs] to the table." These efforts did result in indirect negotiations to stop violence for a time, but more often they served to establish the relational and material conditions under which a pause in violence seemed possible from the perspective of those most invested in a given conflict.

SSB's second strategy, its "service delivery strategy," picked up at this point. Receiving clients through handoffs from streetworkers who had developed trusting relationships, the service delivery strategy sought to make long-term violence reduction possible by addressing the human capital, material, and social-emotional needs of individuals embedded within the gangs served by SSB, with a particular emphasis on providing transitional employment. Capitalizing on streetworker-negotiated or natural reprieves in violence, service delivery promoted the development and realization of life plans beyond consuming cycles of conflict, aimed to remove common barriers to work and independence (e.g., obtaining GEDs, driver's licenses, and social security cards), and shifted routine activities of gang members through paid educational and workforce trainings and employment. During the long run, SSB envisioned service delivery as a means of reducing attachment to the gang, ameliorating constraints that kept individuals in violent contexts, and offering a preview of life after gang conflict. In contrast to the larger streetworker group, SSB's service delivery team comprised four to five highly educated men and women with deep experience in the local human services field.

Although a detailed discussion of SSB's implementation is beyond the scope of the present article, the research team's broader analysis of program implementation revealed that SSB overwhelmingly succeeded in reaching its target population. Of the 533 gang members ever identified by a multistakeholder working group as being appropriate for street outreach services, SSB streetworkers reached 469 (88 percent), with the typical gang member receiving more than 100 hours of in-person contact during the intervention's final 3 years.[5] During this same time period, 245 (52 percent) of the 469 gang members ever on SSB's caseload received human services from SSB's service delivery team, including 118 (25 percent) who were placed in a paying job. Although this service uptake rate may seem low, it must be considered in proper context, specifically 1) that a primary rationale for streetwork is the exceptional difficulty of engaging violence- and gang-involved people in traditional "in-house" services and 2) that streetworkers were tasked with reaching a fixed list of people known to be exposed to violence, many of whom were likely interested in avoiding surveilling social institutions (see Brayne, 2014; Goffman, 2014). Thus, in a context in which many streetwork evaluations find obvious evidence of implementation failure through an inability to reach the appropriate population or administer a program with sufficient integrity and intensity (e.g., Goldstein, 1993; Wilson & Chermak, 2011), SSB represents a remarkable case

[5] Detailed record of streetworker time spent with clients was available only in the second, third, and fourth years of the intervention. With few exceptions, such measures of dosage are rare in contemporary streetwork evaluations and SSB workers spent more time with clients—in absolute and average-yearly terms—than did workers from comparable interventions (e.g., Fox et al., 2015; Klein, 1971).

of a coherent program that produced evidence of deep engagement with a population involved in violence.

4 | ESTIMATING STREETSAFE'S IMPACT ON GANG-LEVEL VIOLENCE

4.1 | Analytic Approach

We employed a quasi-experimental design to compare changes in serious gun violence among Boston gangs that experienced the StreetSafe Boston treatment to changes in serious gun violence among matched comparison gangs that did not receive the intervention (Shadish et al., 2002). We matched each SSB gang to one or more gangs with comparable pretreatment characteristics using a theoretically informed matching (TIM) technique alongside two additional approaches (propensity score matching and coarsened exact matching) to serve as checks on the robustness of our findings. We then ran multilevel negative binomial growth curve models including a difference-in-differences (DID) estimator to estimate SSB's effects on quarterly counts of fatal and nonfatal shootings for SSB gangs relative to the matched comparison gangs.

In a departure from prior streetwork evaluations, our approach relied on gang-level (rather than on neighborhood-level) measures of our key outcomes and covariates, ensuring that our outcome estimates were related to the intervention activities. In addition, our approach leveraged the research team's local knowledge of Boston gangs, historical data on Boston gang characteristics, and original data collection on all fatal and nonfatal shootings in Boston across a 7-year period. Below, we describe our data collection and processing efforts, discuss the matching methods we employed to identify comparison gangs, and introduce our growth curve models.

4.2 | Data

We developed measures of our outcome of interest—gang-level counts of shootings[6]—through a process that began with electronic records of BPD official reports of Homicide by Firearm and Assault and Battery by Means of a Deadly Weapon—Firearm (ABDW—Firearm) incidents between July 1, 2006 and June 30, 2013. These incident reports are generated by BPD detectives or police officers after an initial response to a request for police service and, in practical terms, represent all homicides with a firearm and nonfatal shooting events involving an injury. As described below, we conducted a series of systematic reviews of all citywide shootings in conjunction with BPD detectives and officers across 7 years and created a data set of shooting counts at the gang level.

Police incident data have well-known shortcomings and are biased by the absence of crimes not reported by citizens to the police and by police decisions not to record all crimes reported by citizens (see Black, 1970). Despite these limitations, these data are commonly used in the description of urban gun violence problems (e.g., Papachristos et al., 2015) and the evaluation

[6] Our final database required a decision whether to use gang-*motivated* or gang-*involved* shootings as the appropriate outcome of interest. Gang-*motivated* shootings refer to shootings determined by systematic review to be the product of ongoing gang feuds (as opposed to sudden personal/domestic disputes, drug disputes, etc.). By contrast, gang-*involved* shootings refer to incidents in which gang motives are not a proximate cause. We chose to use gang-motivated shootings as the outcome of interest because this measure was best aligned with SSB's theory of change and intervention activities—a decision that had no meaningful impact on the direction or significance of our results.

of gun violence reduction programs (e.g., Braga et al., 2019). Because homicide typically generates a cadaver, homicide incident reports—involving guns or other means—are regarded as the most reliable and valid data collected on crime. Similarly, nonfatal gun assault incidents that involve injuries are among the most likely crimes to generate police documentation because of various detection technologies, responses to emergency calls for service, and reports of gun injuries from hospitals due to mandatory reporting of gunshot wounds. Such nonfatal shooting data provide the advantage of analyzing a larger and more representative range of gun violence. In contrast to all-crime analyses of police data that frequently reveal serious biases and measurement errors, researchers have found that police reports of gang homicide in large U.S. cities 1) exhibit strong internal reliability; 2) are consistent with the principles of convergent–discriminant validity tests; and 3) demonstrate considerable external validity (Decker & Pyrooz, 2010). What is more, the validity of police-reported gang measures is higher in cities that had specialized policing units directed toward gang problems—such as the BPD's long-standing gang unit.

4.3 | Constructing the Shooting Data Set

We used a "crime incident review" process (see Klofas & Hipple, 2006) to determine whether any given shooting involved a gang member as a suspect, victim, or both. Between 2006 and 2013, the BPD convened separate quarterly shooting review meetings for each of the four policing districts that experience the bulk of gun violence in Boston and a fifth quarterly shooting review meeting for the remaining policing districts. Detectives and officers with detailed knowledge on gangs and gang violence problems (including district detectives, homicide investigators, and personnel from the gang and drug control units) were required to attend these meetings. Altogether, researchers from our team attended more than 35 meetings during the 7-year period.

Each meeting proceeded by BPD detectives and civilian analysts presenting the objective characteristics of each shooting event and the available gang intelligence on the event based on their electronic data systems. Meeting participants shared knowledge on circumstances of the shooting event, the relationships between victims and suspects, and—when the event involved gang members—details on the gangs involved in the shooting. The same two members of the research team attended each of these shooting review meetings and collected, coded, entered, and analyzed the qualitative insights on the nature of each shooting event. In conjunction with BPD analysts, the research team conducted yearly audits of the resulting shooting data set to ensure that all data were being collected consistently and that past incident information was updated as new information became available (such as an arrest for a homicide from an earlier period).

4.4 | Identifying Comparison Gangs

Because the gangs receiving SSB services were not selected randomly, we employed a quasi-experimental research design to match treated gangs to other gangs in the city that did not receive SSB services. We used multiple matching techniques to pair each treatment gang with one or more comparison gang(s) matched on theoretically relevant characteristics before the start of the SSB intervention. Our preferred approach, theoretically informed matching (TIM), leveraged the team's deep local knowledge to ensure strong matches; however, we also used propensity score

matching (PSM) and coarsened exact matching (CEM), a nonparametric alternative to PSM, to ensure our results were not dependent on our choice of matching methods. We then estimated the effect of SSB treatment on quarterly shooting counts by comparing outcomes across the treatment gangs with the matched comparison gangs using growth curve models.

4.5 | Matching Covariates

Contemporary street outreach evaluations have most often assessed programmatic impact using neighborhood-level covariates and outcomes, often due to data limitations that preclude the ability to evaluate impacts at the sub-neighborhood level. Our approach instead used gang-level measures, reflecting the logic of the SSB intervention. Although the common covariates used in neighborhood-level analyses (e.g., population, levels of pretreatment crime) are widely available in U.S. Census and city-level data repositories, such measures are not typically available at the gang level. We overcame this limitation by drawing on our long-standing research program into Boston gangs, which established gang-level covariates along which gangs in the city vary in their structure and composition, with a particular focus on the social factors that cause some gangs to be involved in more shootings than others. Following a body of criminological and sociological research, we identified seven pretreatment variables as necessary for our theoretically informed gang-matching process.

- *Pretreatment Shootings*: Boston gang violence is driven by an ongoing series of retaliatory conflicts (Kennedy et al., 1996). Gangs with higher levels of gun violence have an increased risk of persisting in their shooting behaviors over time (Papachristos et al., 2013). In this analysis, we employed counts of gang-motivated shootings from the aforementioned research team/BPD shooting data set. We sought to balance treatment and comparison gangs on the total number of shootings by and against each gang occurring over this 3-year pretreatment period between mid-2006 through SSB's implementation in mid-2009.
- *Gang Size*: Gangs with larger membership have an increased number of individuals at risk of being shot or committing shootings. We balanced SSB and comparison groups on pretreatment gang size using data from BPD's 2007 gang census.
- *Number of Active Conflicts*: Gangs with larger numbers of rivalries are at increased risk that one or more of these rivalries could turn into an active violent dispute that would generate a string of retaliatory shootings. Retaliation and retribution are perhaps the most frequently cited mechanisms of gang violence (Decker, 1996; Hughes & Short, 2005; Papachristos, 2009). We matched SSB groups to comparison groups with approximately equal numbers of active conflicts during the pretreatment period. We used a gang conflict map previously produced by the research team in collaboration with BPD to code the number of active conflicts for all major gangs in the city between 2007 and 2008.
- *Longevity*: We theorized that groups with deep historical roots were likely to have important differences from more recently formed gangs with regard to the structure of their conflict networks, leadership, influence of incarcerated individuals, and other important intangible factors. As such, we took advantage of an historic citywide gang conflict map produced by Kennedy et al. (1997) to establish our pretreatment measure of longevity. Gangs present in this conflict map in 1997 received a 1, whereas all others received a 0.
- *Housing Project Status*: Neighborhoods with housing projects experience increased levels of gang homicide relative to other city areas without housing projects (Smith, 2014). Moreover,

prior field-based research with Boston gangs (Hureau & Braga, 2018) revealed differences in access to guns between housing project gangs and those associated with "side streets." Gangs associated with a large housing development were assigned a 1, whereas all others received a 0.

- *Racialized Conflict Networks*: Gang violence is overwhelmingly intraracial, a social fact shaped by spatial and network adjacencies rather than by characteristics associated with essentialized notions of "race" (Gravel et al., 2018; Papachristos et al., 2013). Drawing from historic shooting data and neighborhood- and gang-level demographics, we assigned each gang to one of three dominant conflict networks (Black, Cape Verdean, and Latinx) and matched on this factor.
- *Neighborhood Disadvantage*: The degree of concentrated social disadvantage in a neighborhood is strongly correlated with the concentration of violent crime (Morenoff et al., 2001; Sampson & Wilson, 1995), as well as with gang violence in these areas (Papachristos & Kirk, 2015; Rosenfeld et al., 1999). Aligned with prior neighborhood research, we used U.S. Census block group data to develop an index measure of concentrated social disadvantage[7] in the areas surrounding all SSB treatment and comparison gangs. For those gangs whose primary turf spanned more than one block group, we used a spatially weighted mean of the connected block groups to calculate the disadvantage index.

To identify appropriate comparison gangs for each SSB gang, we used three distinct matching approaches: caliper propensity score (*p*-score) matching (Rosenbaum & Rubin, 1983, 1985), CEM (Blackwell et al., 2009; Iacus et al., 2012), and TIM (Rossi et al., 2003). These methods vary in their level of statistical sophistication, but the fundamental goal of each was the same: to minimize the extent to which SSB gangs differed from comparison gangs on the relevant pretreatment characteristics detailed above. With each matching method, we selected comparison gangs from a pool of 46 comparison gangs active in Boston during the entire 7-year evaluation period and did not receive SSB treatment.

As described in the next section, our three matching methods produced similar results. We present the results of all three matching techniques in our tables, while highlighting the theoretically informed matching approach. We focus our discussion on the TIM results because we believe this method produced the strongest matches, taking advantage of the ability to balance on statistical covariates, as well as nonobservable factors made possible by the local knowledge of the research team. We also present results from two other matching methods: the commonly used PSM approach, as well as CEM, a nonparametric alternative.

4.6 | Matching: PSM, CEM, and TIM Approaches

PSM (Rosenbaum & Rubin, 1983) functions by regressing the treatment variable (being an SSB gang) on the set of characteristics that influence selection into the treatment group. Based on these regression estimates, a propensity score (or *p*-score) is calculated for each treatment and potential comparison group reflecting how likely the group was to have been selected for treatment in the SSB program based on its characteristics. The *p*-score summarizes the information about the the-

[7] The concentrated disadvantage index is a standardized index composed of the percentage of residents who are Black, the percentage of residents receiving public assistance, the percentage of families living below the poverty line, the percentage of female-headed households with children, and the percentage of unemployed residents (see Morenoff et al., 2001; Sampson et al., 1997). Because of the high correlation of these variables, we conducted principal components factor analysis, which revealed that variables load on a single factor (which was retained as a standardized index variable).

oretically relevant covariates into a single measure. Each SSB gang can then be matched to one or more comparison groups based on this *p*-score.

We used radius matching with a caliper = .1. Radius matching selects all comparison groups within the propensity score caliper range, allowing for the inclusion of multiple comparison groups for a single SSB gang when multiple strong comparisons exist. If no comparison case appears within the caliper, however, no match is made (Caliendo & Kopeinig, 2008). Our caliper approach matched 16 SSB gangs to 38 comparison gangs. We dropped four SSB gangs from the analysis that had no comparison gangs within the .1 caliper range.

As an alternative to propensity score methods, we also used CEM. Unlike *p*-score methods, CEM guarantees improvement in the balance of each covariate ex ante rather than requiring postmatching balance checks and repeated specifications (Iacus et al., 2012). The CEM matching process is nonparametric, so we are not forced to make assumptions about the nature of the relationships between covariates and assignment to the SSB treatment. The basic premise of CEM is that each covariate is first "coarsened" into bins. For binary variables (e.g., housing project status), exact matching is possible: Project-based groups are in one bin, and nonproject groups are in another. For count variables, however, cutoff points are assigned within the distribution to create bins. For example, with neighborhood disadvantage, cutoffs could be set at the 33rd and 67th percentiles to create three groups with high, medium, and low levels of disadvantage.

After coarsening covariates, matches are made among gangs that appear in the same cell (i.e., set of bins). CEM matches all treatment gangs with the same coarsened characteristics to all comparison gangs with those same features. Thus, a single SSB gang could be matched to multiple comparison gangs, and multiple SSB gangs could be matched to a single comparison group. If an SSB gang has no exact matches, it is excluded from the analysis (i.e., assigned a weight of 0). The CEM approach produces more rigorous (and like-for-like) matches than does the *p*-score approach but at the cost of reducing the cases included in the final sample. To avoid excessive pruning, we matched on five of the seven key variables: pretreatment shootings, gang size, number of active conflicts, housing project status, and neighborhood disadvantage. Our CEM model generated 36 cells and matched 12 SSB gangs to 15 comparison gangs.

Finally, our TIM approach allowed for the local knowledge of the investigators to play an active role in the selection of comparison groups. In this approach, we sacrificed some degree of balance on any particular covariate to achieve matches that made the most intuitive sense. In this way, we sought to inject into the analysis balance on nonobservable characteristics of each gang in addition to balancing on the observable characteristics represented by our covariates. This approach not only satisfied our own qualitative sense of producing good matches, but it also achieved strong balance on observable covariates.

The TIM matching process aimed to match each SSB gang with a comparison gang that resembled the SSB group on both observable and nonobservable characteristics. Given the strong potential effect of pretreatment shootings on posttreatment violence, we took particular care to match pairs based on their approximate level of pretreatment shootings, followed by gang size, housing project status, race, and the remaining covariates. The matching process continued until each SSB gang was paired with an optimal match in the form of a Boston gang that did not receive SSB intervention. For one treatment gang, we could not establish a satisfactory match. This challenge stemmed from the fact that our pool of 46 comparison gangs contained a limited set of project gangs. In the end, we compromised by matching one higher violence, project-based comparison gang to two SSB gangs instead of to only one (weighting that comparison group twice as heavily).

Table 3 displays the covariate balance when comparing the SSB treatment gangs against 1) the total pool of potential Boston comparison gangs ($N = 46$, prematching); and 2) the subsets

TABLE 3 Covariate Balance Across Matching Strategies

	Prematching			Propensity Score, Caliper (PSM)			Coarsened Exact Matching (CEM)			Theoretically Informed Matching (TIM)		
Variables	SSB	All Other Boston Gangs	Absolute Standardized Difference	SSB	Matched Comparison	Standardized Difference	SSB	Matched Comparison	Standardized Difference	SSB	Matched Comparison	Standardized Difference
N=	20	46		16	38		12	15		20	19	
Pretreatment shootings	16.85	9.07	.79	16.56	13.91	.23	19.67	17.06	.23	16.85	14.75	.20
Size	45.95	26.59	.76	42.00	34.88	.26	44.58	42.65	.06	45.95	41.70	.15
Conflicts	3.45	2.65	.54	3.44	3.05	.25	3.67	3.50	.10	3.45	3.30	.09
Historicity/longevity	.75	.52	.49	.69	.54	.30	.67	.61	.11	.75	.65	.21
Housing project	.40	.28	.26	.38	.32	.12	.17	.17	.00	.40	.40	.00
Black network	.75	.74	.02	.75	.53	.46	.67	.79	−.27	.75	.70	.11
Latinx network	.05	.11	−.22	.06	.10	−.12	.00	.08	−.41	.05	.15	−.33
Cape Verdean network	.20	.13	.19	.19	.37	−.41	.33	.13	.49	.20	.15	.13
Neighborhood disadvantage	1.07	.78	.65	1.01	.92	.24	1.05	.71	.80	1.07	.80	.58

of matched comparison gangs associated with each of our three matching strategies (PSM, CEM, TIM). Within each matching strategy, the "Standardized Difference" column denotes the standardized difference in comparing means and prevalence across SSB and comparison groups. The standardized difference allows for comparison of the means across the SSB and comparison groups in units of the pooled standard deviation, a comparison of the relative balance of covariates that, unlike *t*-tests, is not influenced by sample size (Austin, 2009). No consensus exists, however, regarding the appropriate standard for what represents "imbalance;" some "large-*n*" medical studies have proposed a threshold as low as .1 (Austin, 2009), whereas others have followed Cohen's (1988) effect size index in regarding .2 as a threshold for imbalance. We regarded these thresholds as important indicators without considering them firm rules.

As the first column grouping in table 3 shows, consistent with program logic and goals, SSB gangs were more violent, larger, more embedded in conflict, and more likely to be historic than was the typical established Boston gang. The remaining columns display the postmatching means for the three distinct matching processes, demonstrating the quality of matches produced. Across each of the three strategies, we identified a set of comparison gangs for analysis that approximated the SSB gangs on relevant pretreatment characteristics.[8]

4.7 | Growth-Curve Regression Model Specification

We used a variation of a multilevel negative binomial regression model to analyze the quarterly change in gang-motivated shootings for treatment and comparison gangs across a 7-year observation period ($N = 28$ quarters). Specifically, we developed individual growth curve models to estimate gang-level changes in shooting incidents over the observation period (Gelman, 2005; Singer & Willet, 2003). We used a longitudinal negative binomial model in which we predicted within-unit variation at level 1 and between-unit variation at level 2 using level 1 intercepts and slopes as outcomes. Each gang was also allowed to have its own slope and intercept to model different starting levels of shootings and different rates of change. This result was consistent with the observed variation in shootings by gangs; that is, some gangs were highly active, and others were less active.

Our initial analysis proceeded by estimating the impact of the SSB treatment on total shootings involving treated gangs ($N = 20$) relative to total shootings involving comparison gangs via the DID estimator. We used quarterly counts of shootings by and against specific Boston gangs as the outcome of interest in our models. Because shootings by and against any particular gang were rare events, we aggregated shootings into quarterly counts to provide more stable estimates of any measurable impacts of SSB intervention on gang shooting behaviors. The DID estimator estimates the difference in a treated gang's postintervention outcome at time *t* compared with its preintervention outcome, relative to the same difference for the comparison gangs in the sample (see Card & Krueger, 1993). We created a DID estimator to estimate the distinct impacts of SSB on the treated Boston gangs. As such, our growth curve panel regression model was as follows:

[8] Readers may worry about imbalance in some matching strategies—for TIM, this applies to neighborhood disadvantage (ND). The TIM process forced a reckoning with such imbalances and their meaning in a small-*n* gang-level analysis. First, we note that ND is theoretically important *through its impact on gang-level violence* (already measured via pretreatment shootings). Second, the political selection of SSB gangs forced a trade-off between lower violence comparison gangs from disadvantaged areas or higher violence gangs from less (but still) disadvantaged areas. We chose to accept imbalance on ND to create the best like-for-like matches in terms of violence.

$$Y_{ij} = \alpha_i + \beta_{1i}(SSB) + \beta_{2i}(period) + \beta_{3i}(impact) + \beta_{4i}(trend) + \beta_{5i}(trend2) + \beta_{6i}(quarter2) + \beta_{7i}(quarter3) + \beta_{8i}(quarter4)$$

where the quarterly counts of total gang-involved shooting incidents over the 7-year study time period was our primary outcome measure (Y_{ij}; I = gang 1… N, and j = quarter 1 … 28). To estimate the effect of the SSB treatment, we included indicators of 1) whether a gang was in the treatment group (*SSB*) and 2) whether the quarter was postintervention (*period*), as well as a DID estimator (*impact*) that interacted these two variables. To account for secular linear and nonlinear quarterly trends in the dependent variable, we included a variable that was measured as the simple linear additive progression for each quarter during the course of the 7-year observation period (*trend*) and a variable that squared this simple linear additive progression for each quarter (*trend2*). We also controlled for seasonal variation in the quarterly counts of shootings using a series of quarter indicators (*quarter2*, *quarter3*, and *quarter4*).[9]

Stata 17 software was used to calculate the maximum likelihood estimate of the parameters for the DID estimator and to compute the associated probability values; this provided estimates of the effects of the SSB intervention on the treatment gangs relative to the comparison gangs. The parallel trends assumption, requiring that the difference between treatment gangs and comparison gangs is constant over time, is critical to the internal validity of the DID model (Angrist & Pischke, 2009). Our analyses suggest that the parallel trends assumption was met. A visual inspection of the estimates and 95 percent confidence intervals from a negative binomial panel regression of the total gang-motivated shootings on the interaction between quarterly dummy variables and gang treatment status revealed that pretreatment trends were not significantly different from zero (see supplement S1 in the online supporting information [10]). Furthermore, a joint F-test of interaction between gang treatment status and quarterly dummy variables during the preintervention quarters was not statistically significant ($F = 5.16$, $df = 11$, $p = .923$). For our analyses, following convention, the two-tailed .05 level of significance was selected as the benchmark to reject the null hypothesis of "no difference."

5 | RESULTS

5.1 | Trend and ATT Comparisons

Figure 1 displays citywide annual counts of total and gang shootings during the three pretreatment baseline years and four posttreatment years. Figure 1 shows that SSB was implemented during a period of declining gun violence in Boston,[11] with a steep decline occurring between the final baseline year (July 2008–June 2009) and SSB's first year of programming (July 2009–June 2010), when

[9] Quarter 1 (January, February, March) served as the reference category for this polychotomous dummy variable. = Yes, 0 = No). Quarter 4 represented whether the outcome included the sum of October, November, and December shootings (1 = Yes, 0 = No).

[10] Additional supporting information can be found in the full text tab for this article in the Wiley Online Library at http://onlinelibrary.wiley.com/doi/10.1111/crim.2023.61.issue-4/issuetoc.

[11] The number of gun homicides in Boston declined by 40 percent between 2006 (55) and 2013 (33). The BPD launched hot spots policing and focused deterrence programs to control serious violence in the city beginning in 2007–more than 2 years before the implementation of the SSB program. Quasi-experimental evaluations found these programs to be effective in reducing serious violence (see Braga et al., 2011, 2014). The impacts of these programs must be considered when observing both the secular declines in citywide gun violence and the treated and untreated gangs included in this evaluation.

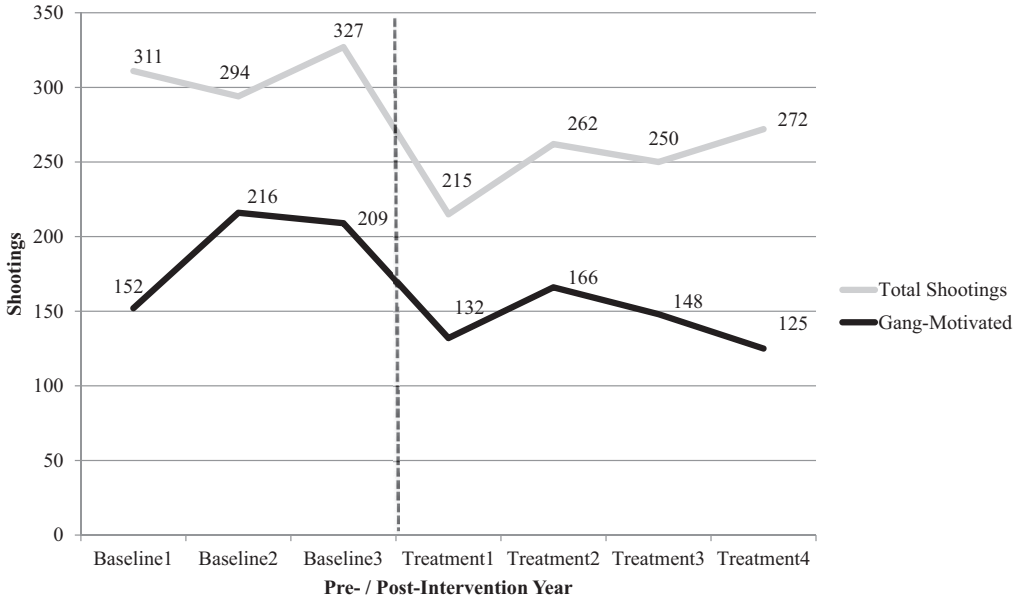

FIGURE 1 Citywide Shootings: Pretreatment Versus Intervention Period

the program was still ramping up its staffing and connection to the gangs it served. This general pattern of decline is evident among all shootings, as well as among gang shootings—the specific type that were the focus on SSB's intervention. In a simple pre–post comparison, mean yearly shootings decreased nearly 20 percent for all shootings (310.7 to 249.8) and decreased more than 25 percent for gang shootings (192.3 to 142.8). Figure 2 offers another representation of this overarching decline in violence, showing the pre–post quarterly shooting involvement of the 20 SSB gangs and 46 potential comparison gangs. Note that in the latter stage of the intervention period, the 20 SSB gangs generated slightly more shootings compared with the 46 potential comparison gangs combined.

Table 4 shows the average treatment effects on the treated gangs (ATTs) for each of the three matching strategies. These results represent the most interpretable estimations of SSB's program impact. Together, these ATT estimates indicate that the SSB treatment was associated with an *increase* in shootings among the gangs served relative to a matched and balanced set of comparison gangs, ranging from 3.65 additional shootings per gang (TIM) to 5.34 additional shootings per gang (PSM) during the 4-year intervention period. None of these estimates met the conventional threshold for statistical significance ($p \leq .05$) in a two-sample *t*-test with equal variances, although the ATT estimate for the PSM strategy did approach statistical significance.

These ATT results are similar in their direction and magnitude across the three strategies. Importantly, only two matching strategies (CEM and TIM) ever yield evidence of a beneficial program impact, and these results occur within a single intervention year, are small in magnitude, and are statistically insignificant. Viewed coarsely, the ATT results show a temporal pattern wherein the SSB treatment impacts appear disproportionately shaped by larger (but still statistically insignificant) increases in shootings in the first intervention year, as well as by larger increases in shootings (that approach statistical significance) in the final intervention year.

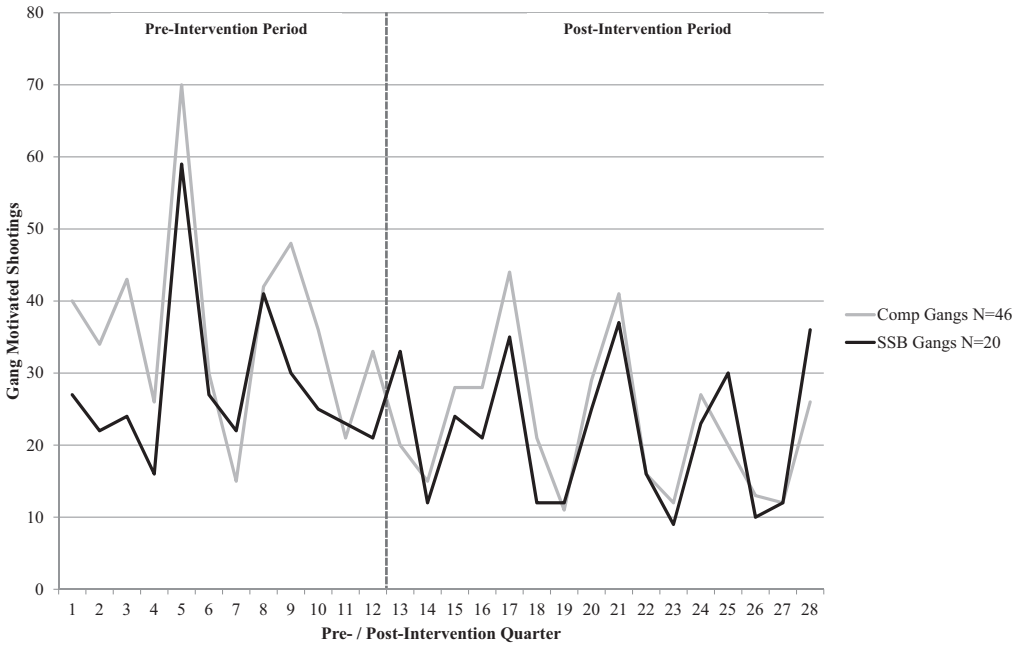

FIGURE 2 Gang-Motivated Shootings by Quarter: SSB Versus Comparison Gangs

TABLE 4 ATT by Matching Strategy and Program Intervention Year (Mean Shootings per Gang)

Matching Method	Y1	Y2	Y3	Y4	Total
Propensity Score (PSM)	1.67	.56	.91	2.20†	5.34†
Coarsened Exact Matching (CEM)	1.76	1.34	−.13	1.38	4.35
Theoretically Informed Matching (TIM)	1.30	−.10	.55	1.90†	3.65

†$p < .10$; *$p < .05$.

5.2 | Growth Curve Regression Model Results

Table 5 presents the results of the growth curve regression models using default standard errors. Controlling for the other covariates, the SSB intervention (*SSB Impact*) was associated with *increases* in quarterly gang-motivated shootings for the treated gangs relative to the comparison gangs across all three matching strategies, although these increases were not statistically significant at conventional thresholds. Furthermore, the SSB gang dummy variable (*SSB Gang*) was not statistically significant at $p < .05$ for all three matching approaches, suggesting that the matched gangs were comparable on the gun violence outcome measures controlling for the other covariates. Consistent with the aforementioned citywide decline in violence, the intervention period (*Period*) was consistently associated with significant decreases in shootings. Across the three matching strategies, the growth curve regression models revealed that Boston gang shootings had statistically significant seasonal variations; relative to January through March quarterly gang shooting counts (Quarter 1), April through June (*Quarter 2*) and July through September (*Quarter*

TABLE 5 StreetSafeBoston Impacts on Gang-Motivated Shootings: Growth Curve Regression Models by Matching Strategy

Variables	Unmatched	PSM	CEM	TIM
SSB impact (interaction)	.283 (.138)*	.304 (.179)†	.191 (.176)	.192 (.153)
SSB gang (1 = treated)	.598 (.165)**	.126 (.172)	.000 (.217)	.197 (.158)
Period (1 = intervention)	−.472 (.158)**	−.378 (.211)†	−.602 (.207)**	−.477 (.181)**
Trend	.020 (.020)	−.014 (.026)	−.009 (.026)	.029 (.023)
Trend-squared	−.001 (.001)	.000 (.001)	.001 (.001)	−.001 (.001)
Quarter 2	.303 (.101)**	.240 (.132)†	.417 (.133)**	.352 (.112)**
Quarter 3	.567 (.103)**	.565 (.133)**	.790 (.133)**	.683 (.115)**
Quarter 4	.006 (.112)	.061 (.143)	.231 (.147)	.087 (.125)
Constant	−.418 (.179)*	−.694 (.277)**	.285 (.251)	−.292 (.202)
p-score		1.41 (.344)**		
Log likelihood	−2,079.946	−1,188.781	−1,071.158	−1,580.565
Wald X^2	97.0	64.4	69.8	66.2
Wald *df*	8	9	8	8
Observations	1,848	1,512	756	1,092
Number of gangs	66	54	27	39

Note. CEM = coarsened exact matching; PSM = propensity score matching; TIM = theoretically informed matching. Default standard errors are in parentheses. Quarter 1 is the reference category for the seasonal dummy variable.
†$p < .10$; *$p < .05$; **$p < .01$.

3) experienced higher counts of gang shootings.[12] Regarded holistically, we interpret these growth curve and ATT results as evidence of a null program effect, albeit one including the possibility of harm.

Supplement S2 in the online supporting information additionally presents the results of the primary TIM analysis specified as a negative binomial panel regression model with bootstrapped standard errors (1,000 repetitions), as well as a Poisson panel regression model with robust standard errors clustered by gang. Supplement S3 further presents the results of main analyses using only suspected shootings as the outcome of interest (as opposed to a measure combining shootings and victimization). These alternative specifications do not yield substantively different results and are in general alignment with our main findings.

6 | DISCUSSION AND CONCLUSION

StreetSafe Boston represented an ambitious attempt to address Boston's gun violence problem through a community-based—rather than a law-enforcement—intervention. Considering the far-reaching impacts of gun violence on historically marginalized communities, as well as the social costs associated with intensive anti-violence policing, it is discouraging when such a

[12] We also ran a simple model specification test to determine whether any of the observed findings were due to a placebo effect. This approach involved restricting the growth curve regression models to the baseline time period analysis and the estimation of the group (0 = comparison, 1 = SSB treatment) dummy variable only. The SSB treatment group variable was not statistically significant in any of the growth curve regression models. This finding suggested that a placebo effect was not present in our main models.

community-based intervention fails to achieve its desired outcomes. But when a null program effect also contains indications of possible harm, such results serve as a call for deeper reflection. In a research domain in which null and negative results have rarely been used to inform programming—and more often have been downplayed, relegated to the file drawer, or used as grounds to withdraw funding for intervention programs (Klein, 2011)—the challenge before us is to articulate what streetwork research and practice can learn from the SSB case. Viewed in isolation, this study represents another humble—but meaningful—brick in the wall of evaluation research into street outreach programs. But, when carefully considered in light of the extended literature, this article can also offer a moment of pause to appraise the structures that have been built in the practice and study of streetwork.

In terms of its logic and method, our evaluation bridges the divide between classic streetwork research (with its proximity to the phenomenon under study) and contemporary outreach evaluations (characterized by their use of advanced quantitative methods). One of our central contributions to this research area was accomplished by means of novel data that allowed us to assess program impact on the gun violence behaviors of the specific gangs served by the SSB intervention, rather than relying on diffuse area-level measures of program impact that may be only loosely related to intervention activity. The impact of the SSB program was assessed using both simple statistical analyses and more complex longitudinal panel regression designs. Although a conventional pre–post research design would have indicated that SSB had a positive programmatic impact, our slate of quasi-experimental analyses revealed that the SSB gangs were not better off for having received the intervention. Indeed, the similarity of findings across the different matching processes and statistical models used lends additional weight to our interpretation of a null program effect. We view these results as typifying the findings of the evaluation literature as whole: null, with some disconcerting features. As such, any generative discussion of the future of street outreach—as well as its past and present—requires grappling with the field's long-standing record of producing null, and sometimes iatrogenic, effects.

When streetwork programs do not yield beneficial results, the finger is most often pointed at program implementation failure. This fingerpointing especially occurs with contemporary public health interventions, which seem to be particularly concerned with fidelity to specific program models (see Butts et al., 2015), although others have argued that the null-to-harmful effects generated by classic streetwork programs were also the result of implementation failure (see Goldstein, 1993). But the SSB results that we observed were unlikely the products of implementation failure. Our process evaluation revealed that SSB was well implemented; SSB overwhelmingly reached its target population and developed broad and deep relationships with the gangs and individuals it served. Despite the blame apportioned to implementation failure in contemporary evaluation research, our SSB result is far from an isolate. In summarizing the classic streetwork evaluation research nearly a half-century ago, Klein dismissed the implementation failure explanation out of hand, concluding:

> *Failure to reduce gang delinquency cannot be explained by lack of program implementation*. In fact, the very existence of research evaluations, putting action staffs "under the gun," probably led to persistent action. There can be no escaping the conclusion that detached work programs, as constituted in the recent past, are not effective in the reduction of gangs or the "violent" activities associated with gangs. They may inadvertently contribute to gang violence. (Klein, 1971, p. 51, emphasis in original)

In the contemporary context, wherein programs have obviously failed to reach the target population (e.g., Wilson & Chermak, 2011) or serve entire target neighborhoods (e.g., Webster et al., 2013), implementation failure remains a real problem for street outreach programs. The larger question, however, is how much weight to assign to implementation failure when streetwork programs produce undesired results. The case of SSB, when understood in the context of classic streetwork literature, suggests that failed implementation is far from the whole story. But if not implementation failure, where else can explanation be found?

Criminological program evaluation has advanced two additional general explanations for null and harmful program effects: 1) measurement failure, which occurs when a research design is not rigorous enough to detect the "true" impact of a program; and 2) theory failure, which occurs when the logic of intervention was unsound from the outset (Ekblom & Pease, 1995; Rosenbaum, 1986). Although, as we have argued, the problem of measurement failure looms large in streetwork evaluation research—principally because group and individual violence is not typically measured by program evaluators—*backfire* effects have most often been explained as a product of *theory* failure. Klein's (1969, 1971) argument that iatrogenic streetwork results flowed from theory failure—and not from implementation failure—was born from close field observation and reading of prior studies that regarded streetworker interventions as atheoretical, leaving the outreach workers themselves to serve as the intervention. This undirected streetworker activity resulted in problematic worker techniques—especially group-based programming—that increased delinquent acts by strengthening group identification and cohesion (Klein, 1969, 1971).

Although often overlooked by researchers working in the public health tradition, Klein's thinking has served as the principal source of explanation when iatrogenic effects are uncovered by criminologists (e.g., Fox et al., 2015; Wilson & Chermak, 2011). But what has been underappreciated, even among scholars familiar with this research, is its theoretic compatibility with a larger criminological tradition concerned with measuring and theorizing the harmful impacts of well-meaning crime prevention programs (e.g., McCord, 2003). By taking stock of the common scope conditions associated with the presence of iatrogenic effects, this line of scholarship provides a way of seeing how the seemingly idiosyncratic backfire effects associated with streetwork programs may be part of a more general criminological phenomenon.

Joan McCord's interest in the potential harms associated with social programs stemmed from her foundational research into the pre–World War II Cambridge-Somerville Youth Study (CSYS, see McCord, 1978; Powers & Witmer, 1951). Offering counseling, home visiting, tutoring, and recreational opportunities to preteen boys, the boys in the CSYS were matched on relevant pretreatment characteristics and assigned randomly to treatment and untreated comparison groups (McCord & McCord, 1959). Following initial null results (Powers & Witmer, 1951), McCord's follow-up with the matched pairs of boys in the 1970s and 1980s revealed a slate of adverse impacts associated with treatment, including mortality, involvement in serious crime, and substance abuse and mental health indicators (McCord, 1978, 1981, 1992). Later research by McCord (2003) and others (Cecile & Born, 2009; Rhule, 2005) dedicated to reviewing unanticipated negative impacts of crime prevention programs found that backfire effects are most often associated with intervention strategies that group delinquent peers together. Welsh and Rocque's (2014) meta-review of individual-based crime prevention programs showed that even though harmful program effects are rare (representing 3.4 percent of 645 effect sizes reviewed), these iatrogenic effects were 1) most often reported in unpublished studies and 2) explained by the grouping of delinquent peers and implementation/theory failure. For street outreach programs, the upshot of this body of research is that the theoretical and empirical case regarding the

pitfalls of group-based programming rests on a much broader base than a handful of midcentury streetwork evaluations—persuasive as these studies may be.

But, in the present-day context, could group-based outreach work increase gang cohesion, resulting in null and backfire effects? After all, the empirical results supporting such an interpretation were mostly produced during a bygone era, and the theoretic premise that increased cohesion could increase negative outcomes was also born from a now dusty period of delinquency research concerned with small-group processes (e.g., Short & Strodtbeck, 1965). What is more, many of the classic empirical illustrations of the processes by which streetworkers increased gang cohesion (e.g., leading weekly club meetings, formalizing sports teams) appear quaint by present-day standards, when gangs have been associated with increased lethal violence, have greater access to deadly firearms, and have become the targets of intense state suppression (Klein & Maxson, 2006).

We argue that the work of this earlier generation of scholars should not be so quickly dismissed. Although contemporary violence researchers may not share the same concern with face-to-face small group processes, they do appreciate the importance of social networks in shaping patterns of serious gun violence (Papachristos, 2009; Papachristos et al., 2013). Recast in social network terms, the seemingly anachronistic group cohesion explanation for outreach worker backfire effects assumes new relevance. Whereas midcentury researchers regarded the streetworker as a unifier of the gang as a social group, attending to their problems of status, prestige, and threat, the social network perspective provides contemporary researchers a way to understand the outreach worker as a potential network broker (Burt, 1992), capable of knitting together subgroups within gangs and increasing the network density associated within and across gang networks. Thus, today's outreach worker's peril is not that of making a *closer* gang but that of making a more *connected* social network—one in which any given violent event reverberates with enhanced consequence.

Although our process evaluation did not uncover evidence of SSB streetworkers engaging in the yesteryear practices of leading gang meetings and increasing gang *esprit de corps*, it did uncover multiple pathways for facilitating connection among SSB gangs: providing introductions between younger and older gang members of the same gang that did not know one another, bringing disparate gang members together in the wake of violent events in the name of safety, and unintentionally enhancing cross-gang "collaborations" by bringing nonfeuding gangs together for services and recreation. Seeing the problem of gang cohesion through the lens of social network analysis does not merely underscore the enduring insights of prior criminological research but further suggests a theoretical and methodological path forward in the evaluation of streetwork. If the increased network density of social groupings served by streetworkers is indeed a mechanism for increased violence, then researchers—as well as practitioners—must move toward measuring and analyzing the social networks of those served by outreach intervention.[13]

Whereas our discussion has focused on the impacts of stand-alone programs, another line of explanation for the ineffectiveness of streetwork programs has emphasized their lack of integration with larger comprehensive violence strategies (Kennedy, 2011; Spergel, 1995, 2007). These explanations have argued that the problem of urban violence is enormous, cast doubt on the

[13] Decker et al. (2012) have also noted the application of social network logic to the analysis of gang cohesiveness, and Hughes (2013) has offered an important empirical analysis of cohesion (operationalized as network density) and its association with violence among midcentury Chicago gangs. Furthermore, it should be emphasized that concerns regarding group programming and cohesion/density also apply to programs that do not use gangs as units of intervention. As Cheng (2018) has demonstrated, network processes shape participant selection and recruitment among public-health–oriented street outreach interventions.

idea that any single program can durably influence violence, underscored the paltry leverage that outreach programs have over the structural forces that permit violence to flourish, and—as an alternative model—highlighted case studies in which streetworkers have played important roles in achieving citywide violence reductions (e.g., Braga et al., 2014). Even though such arguments are difficult to empirically assess, they draw attention to the extra-programmatic forces that shape the creation—and likely the efficacy—of streetwork initiatives, and they remind us that such programs are products of institutional fields of power and the urban political economy. Most importantly, these explanations point to extra-programmatic causal mechanisms that could influence program outcomes (especially at the area level), such as the density and integration of programs within existing networks of capacity for violence prevention. Given recent research showing that the mere presence of nonprofit organizations is associated with neighborhood crime declines (Sharkey et al., 2017), the role that streetwork organizations play in anchoring, spawning, and networking (anti-violence) organizations is a critical area for future investigation.

Our results and interpretations must be tempered by caution and humility, circumscribed by the limitations of our own analysis, as well as by the evidentiary base of the field as a whole. Despite the innovation offered by using gangs as the units of analysis, this decision resulted in fewer units available to analyze, presented uncertainty in the measurement of gang-level covariates, and introduced challenges in achieving balanced treatment and comparison groups. Moreover, the units (gangs) we analyzed were likely to have some degree of interdependence with respect to their violence involvement, thus representing a violation of the stable unit value assumption (SUTVA; Rubin, 1980). Within our analysis of SSB, the chief concern is that the violence outcomes of comparison gangs were affected by the treatment afforded to SSB gangs, thus, biasing our DID estimates in a manner that would understate the true impact of the SSB intervention. Representing a rift in the epistemology of evaluation efforts (which assume the independence of units) and maturing research into networked gang violence (which assumes interdependence), violations of SUTVA are a common challenge for gang violence evaluation efforts at all levels of analysis, but they manifest most obviously at the gang level where spillover effects have been documented and have even served as an intended outcome (e.g., Braga et al., 2013). Although interdependence between SSB and comparison gangs was likely at play, the limited evidence we have to assess this problem indicates that a diffusion of SSB treatment benefits to comparison gangs was not responsible for the null result we observed.[14]

Scholars studying gang violence intervention share with us the challenge of measuring the impact of programs, as well as of developing consensus regarding evidentiary standards in communication with fellow researchers and our publics. As the matter of diffusion makes clear, researchers face a host of difficulties in *how* to appropriately measure gang violence outcomes, even in the uncommon scenarios when high-quality violence data are available. But, especially if null program effects continue to accumulate in the evaluation literature, researchers will also likely confront pressure regarding *what* they measure—or more precisely, *what else* should be measured to assess program impacts, apart from violence. A useful example is Pyrooz et al.'s (2019) effort to assess reductions in gang attachment as an evaluation outcome, but scholars can and should consider other sorts of "pro-social" outcomes such as employment, social integration, or even changes to network density within gangs as possibilities. Such efforts to capture the impacts of streetwork beyond violence have intrinsic value and may better reflect the work of its practitioners, but they can also be effectively combined with violence measures to assess the plausibility

[14] See supplement S4 in the online supporting information. Making use of gang conflict data, this analysis matched a subset of SSB treatment gangs with appropriate comparison gangs, both of which had no prior conflicts with SSB treatment gangs.

of overarching program impact estimates. However the research community responds to the need to develop a new paradigm of streetwork evaluation, our position is that both research and practice would be enlightened by research designs that more closely measure the behavioral impacts on the people, groups, and networks nearest to violence, and would further benefit from moving away from "black box" designs concerned principally with aggregate violence impacts at the city and area levels. We understand the pressures that evaluators face to provide an answer for the "does it work?" question, but we remain pessimistic regarding the quality and utility of the answers provided to this question until researchers dedicate more effort toward returning to the matter of *how* contemporary streetwork programs function.

Regarding the future of street outreach policy and practice, unlike others before us, we believe the claim that streetwork programs should be eliminated from the portfolio of violence prevention policy options altogether is wrong (e.g., Klein, 1971). Instead, policy makers must understand that such programs cannot be depended on to immediately reduce violence—and can produce harm—despite their persuasive policy framing (Butts et al., 2015; Papachristos, 2011). Observers that have held up the *efficacy* of streetwork models as an argument for immediately reallocating social resources dedicated to responding to violence are—for the moment, at least—overly optimistic. For existing streetwork practitioners, our message must be more urgent. Although outreach workers can and do positively influence the lives of the people they serve, everyday street outreach practices often risk increasing client exposure to violence. The existing evidence—built on a broad base of criminological study—persuades that leaders in the field of streetwork should commit to moving programs away from risky group-based activities and streetworker practices that facilitate connection within and across networks exposed to violence. Such a move will be difficult because such techniques comprise much of the commonsense core of outreach practice, and even those violence intervention models that favor individual-level and evidence-based theories of change (such as cognitive behavioral therapy and transitional employment) frequently deliver these services in group settings on grounds of efficiency. How all variants of outreach programming—from the street corner through clinical service provision—come to terms with matters of groupness and network density represents one of the field's most pressing problems of practice. As streetwork is called on to play a greater role in the provision of public safety, its new generation of practitioners should be thoughtfully ushered into a profession that embraces a duty to do no harm and understand the specific practices that make harm more likely.

The crossroads confronting streetwork is defined not only by a necessary reckoning with its neglected past but also by a necessary articulation of what it can and should do for public safety within a moment of reimagination. Streetwork programs answer this latter question in different ways. Entering a new decade, the once-dominant public health paradigm of streetwork rooted in violence interruption is making way for new forms dedicated to individual and community transformation—several of which would feel familiar to streetworkers of the mid-20th century. Recent outreach efforts in cities like Chicago have promoted community transformation not through the transformation of gangs but through the untapped potential of outreach organizations to organize themselves and their communities, building power to demand remedy to long-standing structural conditions that enable intolerable levels of violence and other acute inequalities (Barton et al., 2020; Buckley, 2019). In contrast to this expansive set of responsibilities, other violence prevention programs seem to be further shrinking the streetworker role. Focused on reducing individual-level risk for violence, these initiatives do not even assign the streetworker primary responsibility for violence reduction—instead, that job is performed via individual transformation in clinical and training settings.

Two common denominators lie beneath diversifying streetwork approaches. The first is tacit consensus regarding the field's core capacity: not immediate violence reduction but the essential ability to reach and engage those left behind by nearly all nonpunitive social institutions. And the second is the workers themselves; in Chicago (and elsewhere), violence prevention organizations in the city are stocked with former Cure Violence staff. A side effect of Cure Violence's political success has been the proliferation of diverse violence prevention professionals enlarging and remaking the field, representing a momentous and underrecognized shift in the violence prevention landscape. In contrast to the 1990s, when streetwork programs typically existed in an institutional context in which they were the sole alternative to law enforcement violence suppression efforts, the contemporary period is characterized by increasing nonprofit density in urban space and interlocking social institutions aspiring to establish "thick public safety" (Western, 2019). What was once a marginal organizational niche has now become a much larger institutional field, and this new contextual reality requires reorientation for urban policy makers and streetwork practitioners alike. Policy makers now have more options for non–law-enforcement investments in violence prevention, and they must balance streetwork's uneven track record at generating violence reductions against its rare capacity for bringing the most harmed and marginalized into supportive relationships with their communities. For streetwork practitioners, those interpreting the field's tradition and optimistically ushering in its next era, the responsibility to better realize the profession's potential is rooted not simply in the past but also within the growing community of violence prevention practice that they have helped construct.

REFERENCES

Abbott, A. (1997). Of time and space: The contemporary relevance of the Chicago school. *Social Forces*, *75*(4), 1149–1182.

Addams, J. (1909). *The spirit of youth and the city streets*. MacMillan.

Angrist, J., & Pischke, J. S. (2009). *Mostly harmless econometrics*. Princeton University Press.

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107.

Barton, C., Charles, B., Chavis, L., Givens, A., & Nass, D. (2020, May 6). When the shootings don't stop. *The Trace*.

Black, D. J. (1970). Production of crime rates. *American Sociological Review*, *35*(4), 733–748.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). Cem: Coarsened exact matching in Stata. *The Stata Journal*, *9*(4), 524–546.

Boyle, D. J., Lanterman, J. L., Pascarella, J. E., & Cheng, C. (2010). The impact of Newark's Operation Ceasefire on trauma center gunshot wound admissions. *Justice Research and Policy*, *12*(2), 105–123.

Braga, A. A., Apel, R., & Welsh, B. C. (2013). The spillover effects of focused deterrence on gang violence. *Evaluation Review*, *37*(3–4), 314–342.

Braga, A. A., Hureau, D. M., & Papachristos, A. V. (2011). An ex-post-facto evaluation framework for place-based police interventions. *Evaluation Review*, *35*(6), 592–626.

Braga, A. A., Hureau, D. M., & Papachristos, A. V. (2014). Deterring gang-involved gun violence. *Journal of Quantitative Criminology*, *30*(1), 113–139.

Braga, A. A., Zimmerman, G., Barao, L. M., Farrell, C., Brunson, R. K., & Papachristos, A. V. (2019). Street gangs, gun violence, and focused deterrence. *Journal of Research in Crime and Delinquency*, *56*(4), 524–562.

Brayne, S. (2014). Surveillance and system avoidance. *American Sociological Review*, *79*(3), 367–391.

Buckley, M. (2019, May 22). Ex-gang members and concerned neighbors say they need help from new mayor to fight violence in Chicago. *Chicago Tribune*.

Buggs, S. A., Webster, D. W., & Crifasi, C. K. (2022). Using synthetic control methodology to estimate effects of a *Cure Violence* intervention in Baltimore, Maryland. *Injury Prevention*, *28*, 61–67.

Bursik, R. J., & Grasmick, H. G. (1993). *Neighborhoods and crime: The dimensions of effective community control*. Lexington Books.

Burt, R. S. (1992). *Structural holes*. Harvard University Press.

Butts, J. A., Roman, C. G., Bostwick, L., & Porter, J. R. (2015). Cure Violence: A public health model to reduce gun violence. *Annual Review of Public Health*, *36*, 39–53.
Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.
Card, D., & Krueger, A. B. (1993). *Minimum wages and unemployment* (National Bureau of Economic Research Working Paper Series, w4509).
Cecile, M., & Born, M. (2009). Intervention in juvenile delinquency: Danger of iatrogenic effects? *Children and Youth Services Review*, *31*(12), 1217–1221.
Cheng, T. (2018). Recruitment through rule breaking. *City & Community*, *17*(1), 150–169.
Cohen, A. K., & Short, J. F. (1958). Research in delinquent subcultures. *Journal of Social Issues*, *14*(3), 20–37.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.
Cloward, R. A., & Ohlin, L. (1960). *Delinquency and opportunity*. Free Press.
Dawley, D. (1992). *A nation of lords*. Waveland Press.
Decker, S. H. (1996). Collective and normative features of gang violence. *Justice Quarterly*, *13*(2), 243–264.
Decker, S. H., Melde, C., & Pyrooz, D. C. (2012). What do we know about gangs and gang members and where do we go from here? *Justice Quarterly*, *30*(3), 369–402.
Decker, S. H., & Pyrooz, D. C. (2010). On the validity and reliability of gang homicide: A comparison of disparate sources. *Homicide Studies*, *14*(4), 359–376.
Ekblom, P., & Pease, K. (1995). Evaluating crime prevention. In M. Tonry & D. P. Farrington (Eds.), *Building a safer society*. University of Chicago Press.
Finestone, H. (1976). *Victims of change*. Greenwood Press.
Fox, A. M., Katz, C. M., Choate, D. E., & Hedberg, E. C. (2015). Evaluation of the Phoenix TRUCE project: A replication of Chicago CeaseFire. *Justice Quarterly*, *32*(1), 85–115.
Garland, D. (2001). *The culture of control*. University of Chicago Press.
Gelman, A. (2005). Analysis of variance. *The Annals of Statistics*, *33*(1), 1–53.
Goffman, A. (2014). *On the run*. University of Chicago Press.
Gold, M., & Mattick, H. W. (1974). *Experiment in the streets: The Chicago youth development project*. Institute for Social Research, University of Michigan.
Goldstein, A. P. (1993). Gang intervention. In A. P. Goldstein & C. R. Huff (Eds.), *The gang intervention handbook*. Research Press.
Gravel, J., Allison, B., West-Fagan, J., McBride, M., & Tita, G. E. (2018). Birds of a feather fight together. *Journal of Quantitative Criminology*, *34*(1), 189–219.
Gravel, J., Valasik, M., & Reid, S. E. (2021). Reimagining gang research without the police: Moving the field forward in the era of "defund the police." *Journal of Aggression, Conflict and Peace Research*, *13*(2/3), 73–82.
Hemenway, D. (2006). *Private guns, public health*. University of Michigan Press.
Hinton, E. (2016). *From the war on poverty to the war on crime*. Harvard University Press.
Hughes, L. A. (2013). Group cohesiveness, gang member prestige, delinquency, and violence in Chicago, 1959–1962. *Criminology*, *51*(3), 795–832.
Hughes, L. A., & Short, J. F. (2005). Disputes involving youth street gang members: Microsocial contexts. *Criminology*, *43*(1), 43–76.
Hureau, D. M., & Braga, A. A. (2018). The trade in tools: The market for illicit guns in high-risk networks. *Criminology*, *56*(3), 510–545.
Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: coarsened exact matching. *Political Analysis*, *20*(1), 1–24.
Kennedy, D. M. (2011). Whither streetwork? *Criminology & Public Policy*, *10*(4), 1045–1051.
Kennedy, D. M., Braga, A. A., & Piehl, A. M. (1997). The (un)known universe: Mapping gangs and gang violence in Boston. In D. L. Weisburd & T. McEwen (Eds.), *Crime mapping and crime prevention*. Criminal Justice Press.
Kennedy, D. M., Piehl, A. M., & Braga, A. A. (1996). Youth violence in Boston. *Law & Contemporary Problems*, *59*(1), 147–196.
Kirk, D. S., & Papachristos, A. V. (2011). Cultural mechanisms and the persistence of neighborhood violence. *American Journal of Sociology*, *116*(4), 1190–1233.
Klein, M. W. (1969). Gang cohesiveness, delinquency, and a street-work program. *Journal of Research in Crime and Delinquency*, *6*(2), 135–166.

Klein, M. W. (1971). *Street gangs and streetworkers*. Prentice-Hall.
Klein, M. W. (2011). Comprehensive gang and violence reduction programs. *Criminology & Public Policy*, *10*(4), 1037–1044.
Klein, M. W., & Maxson, C. L. (2006). *Street gangs*. Oxford University Press.
Klofas, J., & Hipple, N. K. (2006). *Crime incident reviews*. U.S. Department of Justice.
Kobrin, S. (1959). The Chicago Area Project—A 25-year assessment. *The ANNALS of the American Academy of Political and Social Science*, *322*(1), 19–29.
Maguire, E. R., Oakley, M. T., & Corsaro, N. (2018). *Evaluating Cure Violence in Trinidad & Tobago*. Inter-American Development Bank.
Martinson, R. (1974, Spring). *What works?* The Public Interest.
Maxson, C. L., & Klein, M. W. (1983). Gangs: Why we couldn't stay away. In J. R. Kluegel (Ed.), *Evaluating juvenile justice*. Sage.
McCord, J. (1978). A thirty-year follow-up of treatment effects. *American Psychologist*, *33*, 284–289.
McCord, J. (1981). Consideration of some of effects of a counseling program. In S. E. Martin, L. B. Sechrest, & R. Redner (Eds.), *New directions in the rehabilitation of criminal offenders*. National Academies Press.
McCord, J. (1992). The Cambridge-Somerville study. In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior*. Guilford Press.
McCord, J. (2003). Cures that harm. *The ANNALS of the American Academy of Political and Social Science*, *587*, 16–30.
McCord, J., & McCord, W. (1959). A follow-up report on the Cambridge-Somerville youth study. *The ANNALS of the American Academy of Political and Social Science*, *322*(1), 89–96.
Miller, W. B. (1962). The impact of a "total-community" delinquency control project. *Social Problems*, *10*(2), 168–191.
Morenoff, J. D., Sampson, R. J., & Raudenbush, S. W. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology*, *39*(3), 517–558.
New York City Youth Board. (1960). *Reaching the fighting gang*.
Papachristos, A. V. (2009). Murder by structure. *American Journal of Sociology*, *115*, 74–128.
Papachristos, A. V. (2011). Too big to fail. *Criminology & Public Policy*, *10*(4), 1053–1061.
Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). The corner and the crew. *American Sociological Review*, *78*, 417–447.
Papachristos, A. V., Braga, A. A., Piza, E., & Grossman, L. (2015). The company you keep? *Criminology*, *53*(4), 624–649.
Papachristos, A. V., & Kirk, D. S. (2015). Changing the street dynamic: Evaluating Chicago's group violence reduction strategy. *Criminology & Public Policy*, *14*(3), 525–558.
Park, R. E., & Burgess, E. W. (1967/1925). *The city: Suggestions for investigation of human behavior in the urban environment*. University of Chicago Press.
Peterson, R. D., & Krivo, L. (2010). *Divergent social worlds*. Russell Sage.
Picard-Fritsche, S., & Cerniglia, L. (2013). *Testing a public health approach to gun violence*. Center for Court Innovation.
Powers, E., & Witmer, H. (1951). *An experiment in the prevention of delinquency*. Columbia University Press.
Pyrooz, D., Weltman, E., & Sanchez, J. (2019). Intervening in the lives of gang members in Denver: A pilot evaluation of the gang reduction initiative of Denver. *Justice Evaluation Journal*, *2*(2), 139–163.
Rhule, D. M. (2005). Take care to do no harm. *Professional Psychology*, *36*(6), 618–625.
Roman, C. G. (2021). An evaluator's reflections and lessons learned about gang intervention strategies: An agenda for research. *Journal of Aggression, Conflict and Peace Research*, *13*(2/3), 148–167.
Roman, C. G., Klein, H. J., & Wolff, K. T. (2018). Quasi-experimental designs for community level public health reduction interventions. *Journal of Experimental Criminology*, *14*(2), 155–185.
Rosenbaum, D. P. (1986). *Community crime prevention*. Sage.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.
Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.
Rosenfeld, R., Bray, T. M., & Egley, A. (1999). Facilitating violence: A comparison of gangmotivated, gang-affiliated, and non-gang youth homicides. *Journal of Quantitative Criminology*, *15*, 495–516.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systemic approach*. Sage.
Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. *Journal of the American Statistical Association*, *75*(371), 591–593.
Sampson, R. J. (2012). *Great American city*. University of Chicago Press.
Sampson, R. J., & Loeffler, C. (2010). Punishment's place. *Daedalus*, Summer, 20–31.
Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, *277*(5328), 918–924.
Sampson, R. J., & Wilson, W. J. (1995). Toward a theory of race, crime, and urban inequality. In J. Hagan and R. D. Peterson (Eds.) *Crime and Inequality*. Stanford University Press.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
Sharkey, P. (2018). *Uneasy peace*. W.W. Norton.
Sharkey, P., Torrats-Espinosa, G., & Takyar, D. (2017). Community and crime decline: The causal effect of local nonprofits on violent crime. *American Sociological Review*, *82*(6), 1214–1240.
Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. University of Chicago Press.
Short, J. F., & Strodtbeck, F. L. (1965). *Group process and gang delinquency*. University of Chicago Press.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.
Skogan, W. G., Hartnett, S. M., Bump, N., & Dubois, J. (2009). *Evaluation of CeaseFire-Chicago*. National Institute of Justice.
Smith, C. M. (2014). The influence of gentrification on gang homicides in Chicago neighborhoods, 1994 to 2005. *Crime & Delinquency*, *60*(4), 569–591.
Spergel, I. A. (1966). *Street gang work: Theory and practice*. Addison-Wesley.
Spergel, I. A. (1995). *The youth gang problem*. Oxford University Press.
Spergel, I. A. (2007). *Reducing youth gang violence: The Little Village Gang Project in Chicago*. AltaMira Press.
Spergel, I. A., & Grossman, S. F. (1997). The Little Village Project: A community approach to the gang problem. *Social Work*, *42*(5), 456–470.
Spergel, I. A., Grossman, S. F., & Wa, K. M. (2006). The comprehensive, community-wide gang program model: Success and failure. In J. F. Short & L. A. Huges (Eds.), *Studying youth gangs*. AltaMira Press.
Thrasher, F. (2013). *The gang* (2nd ed.). University of Chicago Press. (Original work published 1936).
Travis, J., Western, B., & Redburn, S. (Eds.). (2014). *The growth of incarceration in the United States*. National Academies Press.
Vargas, R. (2016). *Wounded city*. Oxford University Press.
Webster, D. W., Whitehill, J. M., Vernick, J. S., & Curriero, F. C. (2013). Effects of Baltimore's Safe Streets program on gun violence. *Journal of Urban Health*, *90*(1), 27–40.
Webster, D. W., Buggs, S. A., & Crifasi, C. K. (2018). *Estimating the effects of law enforcement and public health interventions intended to reduce firearm violence in Baltimore*. Johns Hopkins Center for Firearm Policy and Research. https://www.jhsph.edu/research/centers-and-institutes/johns-hopkins-center-for-gun-violence-prevention-and-policy/_archive-2019/_pdfs/JHSPH-Gun-Violence-in-Baltimore.pdf
Welsh, B. C., & Rocque, M. (2014). When crime prevention harms: A review of systematic reviews. *Journal of Experimental Criminology*, *10*(3), 245–266.
Western, B. (2006). *Punishment and inequality in America*. Russell Sage Foundation.
Western, B. (2019). *The challenge of criminal justice reform* [Square One Project White Paper]. https://squareonejustice.org/wp-content/uploads/2019/09/executive-session-pdf-TheChallengeofCriminalJustice Reform_BruceWestern_Final.pdf
Wilson, J. M., & Chermak, S. (2011). Community-driven violence reduction programs: Examining Pittsburgh's One Vision One Life. *Criminology & Public Policy*, *10*(4), 993–1027.
Woodsworth, M. (2016). *Battle for Bed-Stuy*. Harvard University Press.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hureau, D. M., Braga, A. A., Lloyd, T., & Winship, C. (2023). Streetwork at the crossroads: An evaluation of a street gang outreach intervention and holistic appraisal of the research evidence. *Criminology*, *61,* 758–794. https://doi.org/10.1111/1745-9125.12353

AUTHOR BIOGRAPHIES

David M. Hureau is an assistant professor at the University at Albany, SUNY. He studies violence and social inequality with a focus on community violence intervention and firearms policy.

Anthony A. Braga is the Jerry Lee Professor of Criminology and director of the Crime and Justice Policy Lab at the University of Pennsylvania.

Tracey Lloyd is the senior vice president of science at the Center for Policing Equity. She directs research at the intersection of racial equity and public safety, including studies of school discipline and accountability, police use of force, and public safety redesign.

Christopher Winship is the Diker-Tishman Professor of Sociology at Harvard University. His research interests include causal inference with nonexperimental data, pragmatism, and community policing.