Reducing Gun Violence at Scale

Max Kapustin, Aaron Chalfin, Jeremy Biddle, Brian A. Wade
Natasha Khade, Cristina Layana, Ben Struhl
& Anthony A. Braga

May 27, 2026

Abstract

Baltimore's homicide rate fell by roughly 60% between 2022 and 2025, an exceptional decline among large U.S. cities. At the start of this period, Baltimore launched a strategy that concentrated police and social service resources on a small set of people thought to be driving group-involved gun violence. The approach—"focused deterrence"—has been implemented in some form by cities across the U.S. The strategy was introduced first in the Western police district, one of the highest-violence communities in the U.S. Relative to comparable Baltimore neighborhoods, we estimate that within 18 months shootings and homicides in the Western district fell by roughly one third and carjackings by about 40%, with no spillovers elsewhere in the city. These gains came without expanding overall enforcement: total arrests were flat even as arrests for serious violent crimes rose sharply, indicating that the strategy redirected police authority toward serious violence rather than widening the net of the justice system. Person-level and qualitative evidence point to deterrence, incapacitation, services, and community messengers' legitimacy as contributing channels, with no single mechanism explaining the bulk of the decline. The social value of the averted violence is roughly 35 times the program's first-year spending. Citywide, Baltimore's homicide rate over this period was about 25% below a synthetic counterfactual built from other large cities. The timing of Baltimore's homicide decline and the absence of a larger-than-expected drop in other violent crimes are consistent with the expansion of focused deterrence across the city and a broader shift toward a targeted, partnership-based response to group violence. Baltimore's experience offers an important blueprint for how cities can achieve reductions in gun violence at scale.

Kapustin: Department of Economics and Brooks School of Public Policy, Cornell University & NBER. Chalfin: Department of Criminology, University of Pennsylvania & NBER. Biddle, Wade, Khade, Struhl, Braga: Department of Criminology, University of Pennsylvania. Layana: Abt Associates. This research was made possible by the generous financial support of Bloomberg Philanthropies, the Abell Foundation, the Baltimore Community Foundation, Johns Hopkins University, Brown Advisory, the Goldseker Foundation, the Jacob and Hilda Blaustein Foundation, the Crane Foundation, the France-Merrick Foundation, the Krieger Fund, and Ravens Limited Partnership. Several of the authors have professional relationships with the City of Baltimore related to the program evaluated here. Through the University of Pennsylvania, Biddle, Braga, Khade, Struhl, and Wade have supported the City of Baltimore's continued implementation and expansion of the Group Violence Reduction Strategy (GVRS), through work supported by the funders listed above. Of these, Biddle, Khade, and

1 Introduction

Almost 21,000 Americans were victims of gun homicide in 2021 (Gramlich, 2023). The total number of gun assault victims that year, including non-fatal shooting victims, is likely closer to 100,000 (Cook et al., 2017). The most disadvantaged neighborhoods in American cities—with predominantly Black and Hispanic populations—bear the greatest economic, social, and psychological burdens from gun violence (Cullen and Levitt, 1999; Cook and Ludwig, 2000; Sharkey and Sampson, 2010). In some of these neighborhoods in recent years, young men faced a higher risk of gun homicide than did U.S. combat soldiers at the height of the Iraq war (Del Pozo et al., 2022).

Policing remains the primary means by which cities respond to gun violence. Hiring more officers and allocating them to higher-crime areas have been repeatedly shown to reduce crime and violence, including homicide.[1] Yet every interaction with a police officer carries risks, including the risk of injury to a citizen or an officer (Weisburst, 2019a) and an arrest. Some of these police actions undoubtedly yield considerable public safety benefits. But others may not, including many low-level arrests, for which the public safety benefits are likely to

Wade are currently compensated under that contract; Struhl is a party to the contract but draws no compensation from it. Braga is a principal investigator on the GVRS contracts but has received no compensation under them since 2022; his only related payments, for two summer months in 2021 and 2022, supported the initial problem analysis, assessment, and strategy development that preceded the launch of GVRS. We extend our deep appreciation to the dedicated individuals and organizations working to improve public safety in Baltimore: Mayor Brandon Scott, the late former Deputy Mayor Anthony Barksdale, and Assistant Deputy Mayor Sam Johnson; the Mayor's Office of Neighborhood Safety and Engagement (MONSE), including Director Stefanie Mavronis, former Director Shantay Jackson, GVRS Chief Terence Nash, and the MONSE GVRS team; the Baltimore Police Department (BPD), including Commissioner Richard Worley, former Commissioner Michael Harrison, the Operations Bureau, Colonel Robert Velte, the Crime Strategies and Intelligence Division, and especially the Group Violence Unit; the Baltimore City State's Attorney's Office, including State's Attorney Ivan Bates, Deputy State's Attorney Tom Donnelly, and the Major Investigations Unit; and Youth Advocate Programs, Inc., Roca, and Baltimore's community "moral voice" partners. All have been pivotal partners in this collaborative pursuit. We are also indebted to the BPD and Baltimore City Information & Technology for providing data access and supporting open science, and to the many BPD personnel who helped us better understand the data. We gratefully acknowledge Vaughn Crandall, Reygan Cunningham, Marina Gonzalez, David Kennedy, and David Muhammad for their strategic guidance and support. We thank Jens Ludwig, John MacDonald, and Emily Owens for very helpful comments. Finally, this work would not have been possible without research team members who contributed at various stages, including Ciara Tenney, Alexa Mason, and Deborah D'Orazi. The findings and conclusions reported here are solely the responsibility of the authors.

[1] See, e.g., Evans and Owens (2007); Chalfin and McCrary (2018); Mello (2019); Weisburst (2019b); Chalfin et al. (2022); Braga and Bond (2008); Ratcliffe (2004); Braga et al. (2019a). For summaries of this literature, see Nagin (2013), Chalfin and McCrary (2017), McCrary and Premkumar (2019) and, most recently, Chalfin (2025).

be especially low (Mas, 2006; Harcourt and Ludwig, 2006; Cho et al., 2024; Bacher-Hicks and de la Campa, 2021) and possibly negative.[2] The negative externalities of aggressive and indiscriminate enforcement actions such as abusive and disrespectful encounters (Tyler et al., 2015), unlawful stops and searches (Fagan and Davies, 2000), and excessive arrests (Vitale, 2021) are also borne by the same neighborhoods that suffer most from gun violence, leading them to be simultaneously over-policed and under-protected (Leovy, 2015).

How can a social planner reduce gun violence in the most affected communities while minimizing the costs of broad-based enforcement? A key insight is that gun violence is highly social, often occurring within groups of interconnected people (Papachristos et al., 2012; Green et al., 2017; Bruhn, 2021). These groups are relatively small and usually span a small number of communities within a city (Papachristos et al., 2013), and their composition is often known to the police and to the outreach organizations that operate in those communities. This shared visibility makes it possible to concentrate both law enforcement attention and social service delivery on the same small set of people, with the potential to generate a sizable multiplier effect. On the enforcement side, deterring group members from engaging in violence, or incapacitating them through arrest and detention when they do, can potentially reduce gun violence without the need for broad-based enforcement activity. With respect to social services, targeting outreach, transitional jobs, housing assistance, and other supports to the same group—and delivering them through community messengers whose legitimacy with group members often exceeds that of law enforcement—can offer credible alternatives to violence and reinforce the deterrent message. Such a "double dividend," in which crime and incarceration are both reduced (Durlauf and Nagin, 2011), has been a goal of criminal justice policymaking that has, thus far, remained elusive in national data characterized by stark tradeoffs, with more police presence and staffing leading to a modest gain in public safety but with considerably more low-level arrests (Chalfin et al., 2022; Jabri, 2021).

Strategies of this kind—commonly grouped under the label "focused deterrence"—have

[2] Prosecuting people for non-violent misdemeanor offenses and detaining them in jail for the duration of their case can increase criminal behavior (Leslie and Pope, 2017; Dobbie et al., 2018; Agan et al., 2023).

been implemented widely and studied extensively by criminologists, with mostly promising results (Braga et al., 2026).[3] Yet the existing evidence on these strategies' area-level effects comes from settings in which credible counterfactuals for the treated area are scarce, and it has had little to say about three questions central to evaluating the policy: whether such strategies change the broader composition of enforcement activity, which mechanisms—deterrence, incapacitation, the take-up of services, and the legitimacy of community messengers—account for any public safety gains, and how the strategies engage the people they target.

A smaller set of studies—including nine randomized controlled trials—test interventions that apply the principles of focused deterrence to people rather than to groups (e.g., Davis et al., 2025; Ariel et al., 2019; Ready et al., 2023; Aboaba et al., 2026). While most of the studies report promising effects, by randomizing across people these trials are designed to estimate only a partial equilibrium effect—of reducing one person's violence holding fixed the behavior of everyone else—rather than the general equilibrium effect of reducing violence across an entire network or an entire community. The two can differ: because shootings cluster in networks and propagate through retaliation (Papachristos et al., 2012; Green et al., 2017), preventing one shooting can also prevent the retaliatory shootings that it would otherwise have set off. Critically, like many social services interventions, the interventions are usually implemented at small scale, treating far too few people for an area-level analysis to be adequately powered. More broadly, even person-level interventions with well-identified effects—such as programs featuring cognitive behavioral therapy for high-risk youth and adults (Heller et al., 2017; Bhatt et al., 2024)—have seldom been shown to reduce violence at the scale of a community or city, leaving open whether gains measured one person at a time aggregate to population-scale reductions.

This paper seeks to determine whether a focused, group-based strategy can reduce gun violence at real-world scale—in an entire high-violence urban community, and possibly even an entire city. The strategy we study is not specific to its setting: it follows a focused

[3] We use the term "focused deterrence" to be consistent with the broader literature, but discuss in Section 4.2 its theory of change, which includes mechanisms beyond deterrence.

deterrence framework that cities across the U.S. have implemented in some form, though with widely varying fidelity—a point we return to in assessing what makes the approach replicable. Our setting has credible within-city counterfactuals, and we draw on administrative data of unusual granularity: person-level records of police enforcement actions and of service receipt, together with information identifying which subjects were targeted for communication and which were targeted for arrest. Interviews and focus groups with subjects in the treated community complement the administrative data. Together, these data let us speak not only to whether such a strategy reduces gun violence, but also to how it reshapes the broader composition of enforcement and to evidence on the roles that deterrence, incapacitation, services, and community legitimacy appear to play in producing its effects—roles that the area-level focused deterrence literature has had little ability to examine.

Our setting is Baltimore's Western police district, one of the highest-violence communities in the United States. Baltimore's annual homicide rate has been at or near the 95th percentile among U.S. cities with populations of at least 250,000 for most of the past six decades, with a particularly stark upward divergence beginning in 2015 (Figure 1). By 2021—the year before the strategy launched—its homicide rate of 58 per 100,000 residents was the second highest among large U.S. cities. The Western district of Baltimore is among the most disadvantaged communities in the U.S.—and one of the most widely recognized as such.[4] At the time of the intervention, more than one third of households in the Western district lived below the poverty line and 36% of the district's housing units were vacant, with many buildings boarded up and in visible disrepair. In the years before this strategy's adoption, the Western district's homicide rate was the fifth highest among 323 police districts across 27 large U.S. cities (Figure 2). When including non-fatal shootings, the number of people shot or killed annually in the Western district in those years amounted to 0.5% of its population and an even higher proportion among the community's young men.

[4] Much of West Baltimore, including the Western district, is the setting for David Simon's acclaimed series *The Wire* (2002–2008) and *The Corner* (2000), whose portrayals of concentrated poverty, vacancy, and gun violence brought the area national attention.

Against this backdrop, the Group Violence Reduction Strategy (GVRS) was launched in January 2022 in the Western district. GVRS is a collaboration between the Mayor's Office, the Baltimore Police Department (BPD), local and state prosecutors, and several non-profit and community-based organizations. At a high level, the strategy involves three parts: identification, notification, and action. First, a series of weekly meetings led by BPD identifies people at high risk of committing a shooting or being the victim of one. Then, a team of people—usually a combination of a detective, an outreach worker and a community member—notify the person, either during an individual visit or as part of a group "call-in," that they are at risk of being involved in a shooting and that they have become a focus of GVRS and law enforcement. GVRS then acts on two fronts: people are connected to outreach organizations that can refer them to services ranging from housing assistance to transitional jobs, while those suspected of having committed acts of gun violence are targeted for arrest and prosecution. In addition, BPD increased police presence in housing developments with a high concentration of GVRS subjects and worked with prosecutors to investigate and develop criminal cases against those continuing to engage in violence.

We estimate the effects of GVRS on public safety and enforcement activity in the Western district. Though the Western district has historically had the highest rates of homicide in Baltimore, other areas of the city have broadly similar levels of gun violence and therefore have the potential to be appropriate counterfactuals. Adopting the synthetic difference-in-differences estimator of Arkhangelsky et al. (2021) as our preferred estimator, with synthetic controls (Abadie et al., 2010) and difference-in-differences as alternatives in robustness checks, we construct counterfactuals from areas of Baltimore that had, at the time, not yet been exposed to GVRS. Since Baltimore has only nine police districts, we follow Kapustin et al. (2022) and improve the quality of our counterfactuals by estimating models at the level of smaller police beats or "posts."

We find that GVRS caused the rate of people being shot or killed in the Western district to decrease by approximately a third, and carjackings to decline by about 40%, over the 18 months after its introduction. There is no detectable change in property crimes or assaults more generally, consistent with the empirical regularity that such crimes are more common and less concentrated among a small group of offenders than homicides and shootings. The results are robust to differences in the set of donor areas from which counterfactuals are constructed and the estimation method used. Importantly, we find no evidence that these crimes spilled over to other parts of Baltimore, including areas adjacent to, or with pre-existing ties to targeted groups in, the Western district.

Despite generating such a large decline in gun violence, the strategy does not appear to have caused an increase in overall enforcement. We detect no effect on total arrests in the Western district, including arrests for non-violent offenses and drug crimes. In contrast, GVRS is estimated to have increased arrests for serious violent crimes by 81%, and to have increased the violent crime clearance rate from about 8% to 10%. Taken together, these patterns suggest that GVRS *redirected* the use of police authority in the Western district—toward serious violence and away from the high-volume routine enforcement that drives most exposure to the criminal justice system.

We complement this area-level analysis with a person-level analysis of GVRS subjects' subsequent arrests, matched to observationally similar controls drawn from the broader Baltimore population. The pattern of arrests differs sharply by referral type. Subjects referred for communication, rather than targeted for arrest, show no detectable change in arrest rates relative to controls in the period immediately after referral. Subjects targeted for arrest, by contrast, show an order-of-magnitude increase in arrest rates in the quarter of referral—driven principally by weapons and drug distribution charges, consistent with the strategy's focus on group-involved violence—and a modestly elevated arrest rate one quarter later, indicating that most arrest target subjects are released within the same quarter and remain visible to law enforcement.

Interviews and focus groups with GVRS subjects in the Western district complement the administrative data analysis. Participants describe a sharp increase in perceived risk of apprehension following GVRS contact, prompting changes in their day-to-day routines and social ties; a credible link between communication and follow-on enforcement that supported deterrence; meaningful uptake of services that helped reshape behavior and identity; and a role for community "moral voice" messengers in bridging the legitimacy gap between law enforcement and group members.

Finally, to assess whether these effects can plausibly account for Baltimore's recent homicide decline—a roughly 60% drop between 2022 and 2025 that outpaced most other major U.S. cities and has become the subject of national headlines—we conduct a cross-city analysis using monthly crime data from the Real-Time Crime Index. We find that Baltimore's homicide rate from 2022–2025 was approximately 25% below a synthetic counterfactual constructed from other large U.S. cities, while its non-fatal violent crime rate tracked the counterfactual closely. The Western district effect we estimate, mechanically scaled by the district's ∼16% share of Baltimore's homicides, would account for about a fifth of this 25% deviation.[5] The remainder is consistent with GVRS's subsequent expansion to additional districts and with a broader citywide reorientation toward focused, partnership-based responses to group-involved violence that may have followed. The cross-city analysis cannot, on its own, isolate the contribution of GVRS and this broader shift from other contemporaneous changes in Baltimore's policing or environment. Read together with the within-Baltimore evidence, however, the cross-city results are consistent with GVRS—and the citywide shift it may have catalyzed—accounting for a meaningful share of Baltimore's recent historic homicide decline.

2 Prior literature

2.1 Focused deterrence

The strategy of focused deterrence was first developed in Boston, Massachusetts in the mid-1990s as a response to ongoing youth gun violence that had its roots in the receding crack

[5] This scales our precisely estimated homicide and non-fatal shooting effect; the homicide-only effect is similar in magnitude but less precise. See Section 9.

cocaine epidemic (Kennedy et al., 1996, 2001). The so-called “Boston Gun Project” convened a working group of law enforcement, social service, community and academic partners to analyze the underlying conditions associated with the persistence of youth gun violence in the city. The analysis found that the bulk of youth homicides were driven by ongoing disputes among a very small number of gang members who were very well known to the criminal justice system (Kennedy, 1997). The inter-agency working group developed and implemented Operation Ceasefire, a focused, law enforcement-led strategy which had the goal of preventing gang members from continuing to engage in gun violence (Kennedy, 1997). When outbreaks of gang-involved shootings erupted, the working group concentrated its enforcement capacity on holding the offending groups accountable for their violent behavior, while street outreach workers and community members attempted to convince gang members to stop shooting and take advantage of available services and opportunities.

A hallmark of the Boston strategy was the direct communication of incentives and disincentives to the targeted audience through informal street conversations with specific gang members and formal “call-ins” of the offending groups (Kennedy, 1997). It was thought to be critically important for gang members to understand that shootings triggered the actions of the working group (i.e., demonstrating a “cause and effect” relationship between their engaging in gun violence and law enforcement scrutiny). Maintaining the credibility of these communications required robust enforcement responses by police, prosecutors, and corrections officials, as well as the availability of appropriate services and opportunities for individuals who wanted to step away from violent gang life. Large reductions in gun violence immediately followed the implementation of Operation Ceasefire in Boston (Braga et al., 2001).

The success of the Boston experience inspired several U.S. cities to adopt a similar focused deterrence approach to address violent group offending, including Los Angeles (Tita and Abrahamse, 2004), New Orleans (Corsaro and Engel, 2015), Oakland (Braga et al., 2019b), and Philadelphia (Roman et al., 2019). In addition to these group violence interventions, the principles of focused deterrence have also been used to reduce repeat offending by violent individuals (Papachristos et al., 2007; Ariel et al., 2019) and control disorderly street drug markets (Saunders et al., 2015).

The potential efficacy of a more tailored approach to law enforcement is supported by several theoretical perspectives. Most directly, deterrence strives to reduce crime by changing potential offenders' perceptions of official action and of the associated risk of sanctions (Cook, 1980; Nagin, 2013). As such, effective deterrence necessarily involves both advertising and persuasion (Zimring et al., 1973). Potential offenders need to know the punishment risks they face and need to believe that these risks are genuine.[6] As a result, focused deterrence places a premium on effective communication with potential offenders. Focused deterrence programs also attempt to create a more certain and swift sanction environment through the creative application of existing enforcement capacity and legal authority (Kennedy, 1997). When responding to a violent gang in a particular jurisdiction, authorities can "pull all available enforcement levers" such as warrant service, review and follow-up on open investigations, drug market disruption, civil code and traffic enforcement, monitoring closely and perhaps enhancing probation and parole conditions, more vigilant prosecutorial attention to all legal liabilities, and so forth (Braga and Kennedy, 2021).

Beyond attempting to maximize the deterrence value of law enforcement, these programs also try to control violence by mobilizing informal social control, enhancing police legitimacy, and reducing crime opportunities through situational crime prevention. Focused deterrence strategies emphasize the importance of engaging and enlisting community members. Community-based action in focused deterrence strategies helps remove the justifications used by offenders to explain away their responsibility for the targeted behavior. In call-ins and on the street, community members effectively invalidate the excuses for criminal behavior by challenging the norms and narratives that point to racism, poverty, injustice, and the like (Braga et al., 2001). Community members work with law enforcement and social service

[6] While theoretically obvious, deliberate communication with potential offenders has been neglected in practice (Kennedy, 1997).

agencies to (1) set basic rules for group-involved offenders such as "don't shoot guns" and (2) alert the conscience of these offenders by appealing to moral values inherent in taking the life of another, causing harm to their neighborhood, or the pain that would be experienced by their mothers if they were killed or sent to prison for a long time in a far-away location (Kennedy, 2011).

2.2 Evaluation evidence

An ongoing systematic review suggests focused deterrence programs reduce serious violence. The most recent iteration of the systematic review identified 50 controlled evaluations (41 quasi-experiments and nine randomized controlled trials) of focused deterrence programs (Braga et al., 2026). The overall meta-analysis suggested that focused deterrence was associated with a statistically significant 23% crime reduction in treatment groups relative to control groups. The largest impacts were produced by the gang and group-violence reduction programs. Meta-analysis of the nine randomized experiments suggested focused deterrence generated a smaller 16% crime reduction.

The randomized controlled trials have generally tested the efficacy of interventions featuring elements of focused deterrence applied to individual offenders as opposed to gangs or groups. The findings of four recently completed studies are briefly summarized here. Davis et al. (2025) cluster-randomize detained youth in the Cook County, IL Juvenile Detention Center to either receive or not receive focused deterrence-based youth outreach forums prior to release. They report a 20% reduction in returns to detention and an 18% reduction in total arrests for treated youth in the post-release period, including 43% and 40% reductions in arrests for violent and drug offenses, respectively. Ariel et al. (2019) compare re-arrest rates for "prolific offenders" in Sacramento, CA, some of whom were randomly assigned to be visited by police officers and delivered a deterrence message—including that they would be subject to regular, unannounced visits—and provided a referral to obtain services. They find that the treatment group had a lower re-arrest rate in the 12 month outcome period relative to the control group, a difference that is marginally statistically significant. Ready et al. (2023) test a similar police-led notification program targeting high-rate offenders in Melbourne, Australia, in which treated individuals were warned of enhanced enforcement consequences if they continued offending and offered social services if they sought to desist; the authors report statistically significant reductions in subsequent arrests for the treatment group relative to controls. Finally, Aboaba et al. (2026) study the effects of attending group meetings ("notification forums") that deliver a deterrence message, as some parolees in New York State were randomly required to do as part of their release requirements. They find that parolees assigned to attend notification forums were less likely to experience parole violations within six months relative to the control group, but there were no statistically significant differences in any or violent felony arrests. Furthermore, in the New York City site where some neighborhoods were randomly assigned to hold notification forums for parolees, they detect no difference in neighborhood-level crime rates between treatment and control neighborhoods, although the small sample size makes it difficult to rule out moderate-sized effects.

None of the interventions studied in the above randomized trials were designed to leverage the group dynamics that group-based focused deterrence sets out to exploit. Gun violence in these settings is sustained by retaliation cycling through tightly connected groups. Interrupting it most effectively likely requires reaching the groups as units rather than only some of their members. By engaging high-risk people individually—typically embedded in networks that otherwise go untreated—these interventions are unlikely to generate the multiplier effects, and the at-scale reductions in violence, that concentrating on entire groups can produce. The one randomized evaluation of a group-based focused deterrence program is Denley (2023), who uses a cluster-randomized design to evaluate a group violence intervention implemented in the West Midlands, UK and targeting organized crime groups. He finds that members of organized crime groups assigned to receive the intervention were 35% less likely to be arrested in the post-implementation period than members of control groups. While encouraging, the relevance of the program and its evaluation to U.S. urban gun violence is unclear, given substantial differences in the targeted population and institutional context.

3 Institutional setting

This section provides context on our setting: the city of Baltimore and its Western police district. We first describe the crime, arrest, and sociodemographic patterns in Baltimore and the Western district. Then we provide a brief summary of the social and political context at the time GVRS was launched.

3.1 Baltimore & the Western district

In 2021, 334 people were killed in homicides in Baltimore, the vast majority with guns. Another 724 people were victims of non-fatal shootings. Baltimore's homicide rate of 58 per 100,000 that year was over eight times higher than the U.S. homicide rate (6.9) and approximately double that of Chicago (29.6). Among U.S. cities with more than 250,000 residents, Baltimore's homicide rate was surpassed in 2021 by only that of St. Louis. And 2021 was not an isolated year: as Figure 1 shows, Baltimore's homicide rate has tracked at or near the 95th percentile of U.S. cities with more than 250,000 residents for most of the past six decades, with a sharp upward shift starting in 2015, the year Freddie Gray was killed by Baltimore police officers in the Western district.

The Western district, where GVRS was implemented, has long been the highest-violence area within Baltimore—and among the highest-violence places in the country. In 2021, 52 of Baltimore's homicide victims and 112 of its non-fatal shooting victims were in the Western district—approximately 16% and 15% of the citywide totals, respectively, despite the district comprising roughly 5% of the city's population (Table 1). Among 323 police districts in a set of 27 large U.S. cities, the Western district's homicide rate in the years preceding GVRS was the fifth highest, surpassed only by two districts in St. Louis, one in Louisville, and one in New Orleans (Figure 2). This pattern of intense, geographically concentrated violence in Baltimore was consistent in the seven years prior to the launch of GVRS in 2022. Figures 3 and 4 compare annual rates of reported crimes and arrests, respectively, in the Western district to those in the rest of Baltimore over this period. The Western district saw higher rates of reported crimes than the rest of the city. Yet this gap was greatest for acts of gun violence: rates of homicide and non-fatal shooting victims in the Western district were roughly three times higher than those in the rest of the city, compared to two times higher or less for other types of crimes. Enforcement activity, as measured by arrest rates, was also significantly higher in the Western district than in the rest of Baltimore.

Table 1 reports several sociodemographic characteristics for Baltimore and the Western district drawn from the 2017–2021 American Community Survey 5-year estimates. The Western district comprises approximately 5% of Baltimore's population. Relative to the rest of the city, people living in the Western district are much more likely to be Black (93% vs. 61%) and in households below the poverty line (36% vs. 20%). Median household income in the Western district is approximately half that of the city as a whole ($30,938 vs. $60,891), and, strikingly over 36% of housing units in the Western district are vacant.

3.2 Social and political context preceding GVRS

Baltimore has previously tried to implement several variants of focused deterrence to address persistently high rates of homicide during the late 1990s and again in the mid-2010s. Beginning in 1998, Baltimore developed and implemented a focused deterrence strategy, called Operation Safe Neighborhoods, which targeted violent groups in a drug market area in the Park Heights neighborhood. Despite some promising initial violence reductions, the strategy was dismantled due to political infighting, resistance to operational changes, and obstruction by some of the partnering agencies (Kennedy, 2011). Between 2014 and 2015, Baltimore implemented a precursor to GVRS in the Western district. Once again, after some promising initial reductions, the strategy was discontinued following upper management turnover, lack of governance and accountability structures, and inadequate development of social service capacity and community engagement (National Network for Safe Communities 2016). The 2015 death of Freddie Gray at the hands of Baltimore police officers in the Western district and subsequent social unrest effectively ended the pilot program.

Baltimore's near record level of gun violence in 2019, and debates over the city's inability to implement and sustain a violence reduction strategy to address it, were core issues in the mayoral election race that year. Then President of the City Council Brandon Scott pledged to fully support the adoption of a focused violence reduction strategy as part of his election platform. Immediately following his June 2020 election as the Democratic nominee in the citywide elections, presumptive Mayor-Elect Scott engaged the Crime and Justice Policy Lab at the University of Pennsylvania for technical assistance in the design and implementation of a such a strategy. In 2021, the Western district was selected as the initial implementation site based on data showing that it experienced the largest number of combined homicides and shootings between 2015 and 2020 among the city's nine police districts.

The University of Pennsylvania's Crime and Justice Policy Lab completed a problem analysis in mid-2021 of homicides and shootings in the Western district to guide initial operations. The analysis revealed a strong overlap between victim and offending populations: both were mostly Black (>95%), male (>80%), ages 16–34 (>70%), and well known to the criminal justice system (>81% had prior criminal justice involvement with a mean 12 prior arrests). Most homicides stemmed from interpersonal disputes that involved members of gangs, drug trafficking organizations, robbery crews and other criminally active groups (70%). At the time of the problem analysis, the Western district had 18 criminally active groups with an estimated 615–735 members representing about 2% of its population. Some groups were more engaged in violence than others; one group, operating out of the Poe Homes housing development, was thought to be responsible for at least 13 shootings in the 18 months between January 2020 and June 2021.

In 2021, the Mayor's Office, BPD, and community partners continued laying the groundwork for GVRS. The Mayor's Office of Neighborhood Safety and Engagement (MONSE) developed contracts with Roca, Inc. and Youth Advocate Programs, Inc. (YAP) to serve as the strategy's primary street outreach and service providers. Case management systems were adopted to monitor service uptake by treated individuals. Commitments from other local, state, and federal law enforcement agencies to support GVRS operations were secured. A specialized Group Violence Unit (GVU) was set up, staffed with 30 detectives, led by a lieutenant, and trained in the required investigation, suppression, and partnership work needed to implement GVRS. MONSE partnered with the Baltimore Community Mediation Center (BCMC) to support initial community engagement work in the Western district and elsewhere in the city. MONSE and BCMC held a series of listening sessions with residents on GVRS and recruited community members, such as clergy and local anti-violence activists, to serve as "community moral voice messengers" to support communications with targeted groups and individuals.[7]

4 Intervention

The implementation of GVRS in the Western district has been constantly evolving since its launch in 2022. This section first summarizes, at a high level, the three parts of the intervention—identification, notification, and action (the latter encompassing both supportive services and focused enforcement)—and then describes how this strategy might potentially reduce violence. Finally, we document how the implementation of GVRS in the Western district unfolded.

4.1 GVRS components

Identification BPD holds weekly meetings with other law enforcement agencies including federal law enforcement ("shooting reviews") to discuss shooting incidents in the Western district involving suspected group members. An output of these shooting reviews is the identification of group members or their associates who become the focus of GVRS activities. We refer to these individuals as *GVRS subjects*, and they can be further divided into two groups.

The first group of GVRS subjects are those thought to be at high risk of committing

[7] In 2022, the BCMC partnership ended as MONSE strengthened its staffing to continue community engagement work.

a shooting, or being a victim of one in the near future. This could include as a non-fatal shooting victim in a focal incident or as a relative or friend of a shooting victim, who could potentially be considering retaliation. These individuals are referred for communication (*communication referrals*).

The second group of GVRS subjects are those suspected of having committed or been involved in committing a recent shooting or another serious violent crime. This includes suspected offenders in a focal incident or in another recent violent crime. These individuals are targeted for investigation and, depending on the evidence uncovered, arrest (*arrest targets*).

Notification A separate weekly coordination meeting, held after the shooting review, is attended by BPD, MONSE, members of the community who can convey its "moral voice" and representatives of the two outreach organizations that are part of GVRS: Roca and YAP. This meeting is used to determine the best way to reach a communication referral. Most often, this is done via an individualized meeting ("custom notification") that includes a detective, an outreach worker, and a community member.[8] Less often, this is done as part of a group "call-in." In either case, the strategy calls for a direct contact with the person, rather than an indirect contact such as leaving behind a letter or a message with a friend or relative.

The message delivered by the GVRS team consists of four parts. First, the GVRS subject is told that they have been identified as being at high risk of being involved in a shooting. Second, they are told that engaging in violence will bring focused law enforcement attention upon them. Third, a local community member conveys to the subject their disapproval of gun violence and the harm it is inflicting on the community, and encourages the subject to take advantage of services. Finally, a Roca or YAP outreach worker discusses with the subject the services and opportunities available to them.

[8] In some cases, such as when the GVRS subject is part of an active and ongoing investigation or is directly an arrest target, a custom notification may involve only a detective. Referred to as "law enforcement only" custom notifications, these can happen following an arrest and while the GVRS subject is in custody.

Action Notification is paired with action on two fronts—supportive services and focused enforcement—calibrated to each subject's referral type and conduct. The two embody the strategy's "carrots and sticks": services offer credible alternatives to violence, while enforcement raises the cost of continuing it.

Supportive services. Through Roca and YAP, GVRS subjects—predominantly those referred for communication—are offered low-barrier, non-judgmental services and supports ranging from intensive mentorship and crisis management to counseling and cognitive behavioral therapy, employment and licensing assistance, education and vocational training, emergency relocation, microgrants, and stipends. A Roca or YAP outreach worker is typically introduced during the notification itself and continues to engage the subject afterward in an ongoing relationship that aims to provide a credible alternative path away from violence. Because these services are delivered through community-based organizations whose legitimacy with group members often exceeds that of law enforcement, they also reinforce the credibility of the strategy's broader message.

Focused enforcement. A key part of GVRS is enforcement activity. In some cases, enforcement is directed at a group suspected of carrying out shootings and results in coordinated arrests of multiple group members (a "group takedown"), usually following a prolonged investigation by BPD and a prosecuting agency such as the State's Attorney's Office. In other cases, enforcement is directed at individual subjects. In addition to efforts directed at groups or individuals, the GVU and other BPD units sometimes engaged in suppression activities designed to increase police presence in areas with a high concentration of GVRS subjects, such as certain housing developments where groups linked to shootings operate.

To coordinate these efforts and lend credibility to the focused attention described in communications to GVRS subjects, BPD holds monthly "strategic enforcement" meetings with other law enforcement partners, such as the State's Attorney's Office and federal agencies. The goal of these meetings is to discuss and establish investigative priorities, strengthen ongoing cases and determine the appropriate venue for prosecuting them (state vs. federal court).

4.2 Theory of change

The focused deterrence literature suggests—though does not firmly establish—that the primary mechanism through which the strategy works, consistent with its name, is deterrence. The idea is that through the use of both "carrots" and "sticks"—services meant to steer people away from violence and threats of enforcement if they fail to do so, respectively—people who would otherwise be at high risk of becoming involved in violence—mainly as offenders, but potentially also as victims—will take steps to reduce that risk.

Under the umbrella of deterrence, there are at least three related but distinct ways that this strategy can generate effects. First, the custom notifications and call-ins—with their message of greater law enforcement scrutiny—could have the intended effect on the people receiving them, deterring them from engaging in violence due to a heightened perception that they will be caught and punished if they do so. Second, suppression efforts that focus law enforcement attention on GVRS subjects and the places where they spend their time could also generate a deterrence effect. Third, both of the previous pathways for generating deterrence could create spillover effects on people connected to GVRS subjects.

How quickly any such deterrent effect can take hold depends on how offenders come to perceive the enforcement they face. Canonical models of deterrence assume potential offenders are well informed about the risk of sanctions, yet deterrence operates through *perceived* risk, which need not track actual enforcement closely (Nagin, 2013). In practice, offenders must infer the prevailing enforcement regime from experience—their own and their peers'—which can be slow and noisy, especially when the underlying probabilities are low. Detecting on one's own that the odds of arrest for a violent crime had risen would require observing many arrests across one's network. By communicating the change in enforcement directly to those at highest risk, GVRS aims to short circuit this learning: rather than waiting for offenders to infer a tougher environment from accumulated experience, it tells them, credibly and immediately, that the regime has changed. To the extent deterrence operates here, this acceleration of beliefs—and not only the underlying change in enforcement—may be central to it.

However, there are also mechanisms other than deterrence through which GVRS can affect violence. The first and most obvious of these is incapacitation. This follows from the fact that a key component of GVRS is enforcement activity against people suspected of having committed shootings, with the intention of arresting and prosecuting them. Indeed, 110 GVRS subjects were arrested as part of GVRS enforcement efforts during the 18-month study period (Table 3). The second mechanism is behavioral change that is due not to deterrence, but rather to the rehabilitative services offered to GVRS subjects, 90 of whom took them up. A growing literature suggests that some social services, particularly those featuring elements of cognitive behavioral therapy designed to teach people to recognize problematic thinking that can lead to harmful behavior, can substantially reduce violence involvement among participants (Heller et al., 2017; Blattman et al., 2017, 2023; Bhatt et al., 2024).

4.3 Implementation of GVRS in the Western district

Our description of how GVRS unfolded in the Western district draws on records that GVRS maintains for each subject and each contact event with them. These records contain each subject's initial referral date and type (communication referral or arrest target), subsequent direct communications with their date and subtype (full custom notifications including service providers, law enforcement only notifications, and call-ins); service connections through Roca and YAP, with the categories of services delivered (such as counseling and cognitive behavioral therapy, employment assistance, education and vocational training, emergency relocation, and stipends); and arrest events with their date and the type of arrest (group takedown or other GVRS-related arrest). Subjects can also be linked to specific groups or named takedowns.

On-the-ground implementation of GVRS began in January 2022 with efforts to communicate the strategy's message to subjects. Focused initially on a group operating in the Poe

Homes housing development, the intervention began with custom notifications directed at Poe Homes group members. This was followed by a takedown of a subset of Poe Homes group members in March 2022 that resultied in the indictment of 11 targeted individuals on a variety of violent, gun and drug charges. Later, in September 2022, eight members of the Poe Homes group participated in a formal call-in. Subsequent takedowns are listed in Appendix Table A.1.

Table 2 reports baseline characteristics of the 276 GVRS subjects whose first referral was in the Western district prior to July 2023. The data are broken out by initial referral type (179 communication referrals and 97 arrest targets). The subjects are nearly uniformly young Black men whose average age at the time of referral was 31. They have substantial prior involvement with the criminal justice system: 81% had at least one prior arrest in BPD records before their GVRS referral, 57% had at least one prior arrest specifically in the Western district, and the average subject with any prior arrest record had 6.3 prior arrests on file. Communication referrals and arrest targets are similar on most of these dimensions but differ sharply on prior victimization: communication referrals are two and a half times as likely as arrest targets to have ever been shot (30% vs. 12%) and five times as likely as arrest targets to have been shot in the year before referral (20% vs. 4%), reflecting the strategy's use of recent shooting victimization as a marker of elevated risk for inclusion in the communication referral pool.

Table 3 summarizes activities directed toward these subjects, overall and separately by initial referral type. Roughly 70% of all GVRS subjects, and 85% of the communication referrals, received some type of direct communication—almost always a custom notification rather than a call-in. Eight subjects participated in the only call-in to occur during the study period, the event involving members of the Poe Homes group in September 2022.

Service uptake was concentrated among communication referrals: of the 90 GVRS subjects who took up services through Roca or YAP, 81 were communication referrals. The most common service category was counseling and life coaching, including cognitive behavioral therapy (86 subjects). Smaller numbers received employment and licensing help (30), education and vocational training (24), and stipends or emergency relocation help (13).

Finally, about 40% of all GVRS subjects, and nearly all of the arrest targets, were arrested as part of GVRS enforcement. Most of these arrests occurred as part of group takedowns rather than individual targeted arrests.

5 Effects on crime, enforcement, and reporting in the Western district

This section estimates the effects of GVRS on rates of reported crime, arrests, and willingness to report crime to the police in the Western district. To do this, we estimate counterfactual outcomes for the Western district using observed outcomes from other areas of Baltimore that did not implement GVRS. Our willingness-to-report measure follows the approach of Ang et al. (2025), comparing 911 calls to acoustically-detected gunfire from ShotSpotter. We focus on effects during the 18-month period starting from the launch of GVRS in January 2022 through June 2023. The evaluation priod ends in mid-2023 because Baltimore's police district boundaries were redrawn in July 2023, complicating our ability to implement the estimation strategy we describe below in later time periods. We also look for evidence of spatial spillover effects on reported crime rates in areas adjacent or otherwise linked to the Western district.

5.1 Data

To conduct this analysis, we use four types of administrative data to build a panel dataset of reported crimes, arrests, 911 calls and incidents of gunshots detected by ShotSpotter for each area in Baltimore over time. Baltimore is divided into nine police districts, which are further divided into 126 "posts," which are smaller police beats within each district. To have greater flexibility in estimating counterfactual outcomes for the Western district, we aggregate data to the post by calendar quarter level.

Reported crimes We use publicly available data on crimes reported to BPD.[9] We aggregate these data into counts, at the post-quarter level, for four crimes that commonly involve the use or threat of firearms against a person: 1) homicides and non-fatal shootings, 2) aggravated assaults, 3) robberies, and 4) carjackings.[10] Due to potential under-reporting of some crimes in earlier years (including non-fatal shooting victims), we only use reported crime data from January 2015 onward.

Arrests We use non-public arrest data from BPD. Each record in these data is a criminal charge associated with an arrest made by BPD officers. Our goal is to first categorize each charge using its standardized Criminal Justice Information System (CJIS) code when available, or using a pair of free text charge description fields entered by BPD officers when the CJIS code is missing.[11] We then use these categorized charges to construct several arrest-level measures aggregated to the post-quarter level from January 2015 onward.

Our broadest measure is *total arrests*, the count of all custodial arrests in a post-quarter. We further divide arrests by whether they involve a charge for a *violent index crime*: homicide, aggravated assault, sexual assault, robbery, or carjacking. *Violent index crime arrests* are arrests with at least one charge for such an offense. *Non-violent index crime arrests* are arrests where none of the charges are for violent index crimes; these are arrests over which officers may have greater discretion. Among the non-violent index crime arrests, we further isolate *drug arrests*, defined as arrests with charges for drug possession or drug distribution. Finally, to capture the intensive margin of enforcement targeted at violent offending, we construct a *violent index crime clearance rate*, defined at the post-quarter level as the ratio of violent index crime arrests to reported violent index crimes.[12]

[9] https://data.baltimorecity.gov

[10] We count *victims* of homicides and non-fatal shootings, and *incidents* of aggravated assault, robbery, or carjacking. Victims of non-fatal shootings are often reported twice in the public crime data: once as non-fatal shooting victims, and once as aggravated assault victims. In these cases, to keep our crime categories mutually exclusive and avoid double-counting, we retain only the reported non-fatal shooting victims.

[11] For CJIS codes, see https://mdcourts.gov/district/chargedb.

[12] This is a "quasi-clearance rate" measure because individual arrests cannot be reliably linked to the specific reported crimes they resolve in these data, so we instead compute the ratio of arrests in a post-quarter to reported crimes in the same post-quarter.

There are 167,413 BPD arrests from January 2015 through June 2023. Geographic information in the arrest data is often incomplete: 90,586 (54%) of these arrests have no street address recorded—predominantly arrests served by the BPD Warrant Section (35%) and the Central District—and so we drop them from this analysis because they cannot be assigned to a post.

911 calls and ShotSpotter incidents We use non-public data from BPD on 911 calls and shooting incidents identified by ShotSpotter, an acoustic gunshot detection system deployed in some parts of Baltimore. Following Ang et al. (2025), we use these data to construct a proxy measure of the public's willingness to report crime to the police, computed at the post-quarter level as the ratio of 911 calls to ShotSpotter incidents.[13] The intuition behind this measure is that ShotSpotter detects the volume of gunfire in an area independent of the public's willingness to report; holding the underlying rate of gunfire fixed, an increase in willingness to report would push the calls-to-incidents ratio up (more public calls per acoustically-detected incident), while a decrease would push it down. ShotSpotter coverage in Baltimore is, however, incomplete: large portions of the Northern, Northwestern, Southern, and Southeastern districts register essentially no incidents over the entire January 2019 through June 2023 window, indicating that the system was not deployed in those areas (Appendix Figure A.3). For analyses involving ShotSpotter data, we therefore restrict the dataset to posts that recorded at least one ShotSpotter incident in every quarter of that window (25 posts).

5.2 Estimation strategy

We estimate the effects of GVRS in the Western district by constructing counterfactual outcomes from areas of Baltimore that did not receive the intervention. Adopting the notation of Arkhangelsky et al. (2021), let Y_{idt} be an outcome—for example the rate of homicide and shooting victims per capita—observed in post i in district d in quarter t. A binary measure

[13] ShotSpotter data are not available before January 2019, so this measure begins in January 2019. We restrict 911 calls to those made by the public, removing "on view" calls for service generated by officers in the field.

of whether GVRS was active in a district in a given quarter is $W_{dt} \in \{0, 1\}$. Consider estimating for the Western district d^* the average treatment effect on the treated (ATT) τ_{d^*} of GVRS, which launched in $t = T_{post}^{d^*}$. Our panel dataset consists of N posts observed for T quarters, with the first $N_{tr}^{d^*}$ posts being those in the Western district d^* (treated) and the remaining $N_{co} = N - N_{tr}^{d^*}$ being those in districts that did not have GVRS during the 18-month post-intervention period (control, or donors), and where $W_{d^*,t} = 1$ for $t = T_{post}^{d^*}, ..., T$ and $W_{dt} = 0$ otherwise.

Estimating τ_{d^*} in this setting is non-trivial. The Western district has both the highest levels of violence in Baltimore historically and experienced different trends in violence than the rest of the city in the years preceding GVRS's launch. Figure 3 presents rates of four crimes that are our focus throughout—homicides and non-fatal shootings, aggravated assaults, robberies, and carjackings—in the Western district and the rest of Baltimore, before and after GVRS launched. Even within the pre-treatment period, the gap between the Western district and the rest of Baltimore as a whole is far from constant: it widens and narrows in different ways across these four outcomes and across time, with especially sharp movement in robberies around the onset of the pandemic. A similar pattern is observed for arrest rates (Figure 4), with secular declines over many years across the city and notably sharp changes in drug enforcement around the onset of the pandemic.

The standard two-way fixed effects difference-in-differences (DID) estimator absorbs time-invariant level differences between treatment and donor units, but it assumes that their outcomes would have evolved in parallel in the post-treatment period in the absence of treatment. This parallel trends assumption is hard to defend in our setting given how clearly the Western district and the rest of Baltimore diverge on some outcomes in the pre-treatment period. The synthetic controls (SCM) estimator (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015) chooses donor unit weights to match the treated unit's pre-treatment outcome path but, because it omits unit fixed effects, must absorb level differences through the weights themselves. This can leave it sensitive to differential pre-treatment trends—

particularly when, as for several of our outcomes, the pre-treatment period spans distinct regimes separated by events like the COVID-19 pandemic.

We therefore use synthetic difference-in-differences (SDID) (Arkhangelsky et al., 2021) as our preferred estimator. SDID solves:

$$(\hat{\tau}^{sdid}_{d^*}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{idt} - \mu - \alpha_i - \beta_t - W_{dt}\tau\right)^2 \hat{\omega}^{sdid}_i \hat{\lambda}^{sdid}_t \right\}. \tag{1}$$

SDID combines features of both DID and SC. Like DID, it includes unit fixed effects α_i, making the estimator robust to time-invariant level differences between the treated unit and the donor pool. Like SC, it assigns donor weights $\hat{\omega}^{sdid}_i$ chosen to approximately match the treated unit's pre-treatment outcome trajectory, subject to constraints against extrapolation; because the unit fixed effects already absorb time-invariant level differences, the SDID weights need only match pre-treatment trends rather than levels.[14] Crucially, and unlike either DID or SCM, SDID introduces time weights $\hat{\lambda}^{sdid}_t$ that give greater emphasis to the pre-treatment quarters in which donor outcomes most resemble the post-treatment period. The combination of donor and time weights is intended to make the estimator more robust to deviations from parallel trends than either set of weights alone, including in settings where pre-treatment outcomes contain structural breaks.[15] Note that, while the Western district d^* contains multiple posts ($N^{d^*}_{tr} > 1$), these are averaged together to form a single treated unit ($Y_{d^*,t} = \frac{1}{N^{d^*}_{tr}} \sum_{i=1}^{N^{d^*}_{tr}} Y_{idt}$) prior to the estimation of donor weights, as is typically the case in applications of SC and SDID.

DID and SCM can each be viewed as special cases of SDID. Both impose uniform time weights ($\hat{\lambda}^{sdid}_t = 1/T_{pre}$). DID also imposes uniform donor weights ($\hat{\omega}^{sdid}_i = 1$) while retaining unit fixed effects. SC allows donor weights to vary but lacks unit fixed effects. We report

[14] The papers establishing the SCM method emphasize that donor weights are chosen to minimize differences in pre-treatment covariates with the treated unit, where the covariates can include, but are not limited to, pre-treatment outcomes. In practice, as Klößner et al. (2018) note, many researchers include the entire pre-treatment outcome path among the covariates, which renders other covariates irrelevant for which weights are chosen. Lacking detailed covariates on our donor units, we adopt this approach as well.

[15] We implement SDID using the `sdid` package in `Stata` (Pailañir and Clarke, 2023).

SCM and DID estimates alongside SDID as robustness checks in Section 5.4. As we show there, our headline conclusions are robust to estimator choice for nearly all outcomes; the only outcomes for which SDID and DID diverge meaningfully are robberies and drug arrests not co-charged with a violent index crime, both of which exhibit a pronounced structural break around the COVID-19 pandemic in Baltimore, as shown above. SDID's time weights load heavily on the post-pandemic pre-period, reconstructing a counterfactual that tracks the Western district's actual pre-treatment trajectory just before GVRS launched, while DID's uniform time weighting averages across two distinct regimes and inflates the implied pre-versus-post change. We view this as supporting our choice of SDID as our preferred estimator.

The SDID specification above delivers a single aggregate ATT averaged over the post-treatment window. To examine the dynamics of the effect quarter by quarter and to assess pre-treatment fit, we additionally report estimates following the event-time SDID extension of Ciccia (2024).[16] The standard event-study translation used in difference-in-differences—interacting treatment with a series of relative time dummies in a single regression—does not carry over directly to SDID. The time weights $\hat{\lambda}_t^{sdid}$ are estimated using only pre-treatment outcomes and are calibrated to make a single $\hat{\lambda}$-weighted pre-treatment average comparable to the post-treatment average, useful for estimating the aggregate $\hat{\tau}_{d^*}^{sdid}$. Putting period-specific treatment indicators on the right-hand side of equation 1 would tie the post-treatment coefficients to this single calibration, and would similarly tie any pre-treatment coefficients to the same $\hat{\lambda}$-weighted reference that the aggregate estimate uses.

We therefore construct period-by-period event-time estimates by holding the donor weights $\hat{\omega}_i^{sdid}$ and time weights $\hat{\lambda}_t^{sdid}$ from equation 1 fixed and substituting one quarter at a time as the post-treatment period in the SDID estimator. The period-r estimate is the $\hat{\omega}$-weighted gap between the Western district and the donor pool at quarter $T_{post}^{d^*}+r$, minus the $\hat{\lambda}$-weighted average of that gap over the pre-period.[17] For post-treatment event times $r \in \{0, 1, \ldots, 5\}$,

[16] We implement event-time SDID using the `sdid_event` package in `Stata`.

[17] Formally, $\hat{\tau}_{d^*,r}^{sdid} = \big(Y_{d^*,T_{post}^{d^*}+r} - \sum_i \hat{\omega}_i^{sdid} Y_{i,T_{post}^{d^*}+r}\big) - \sum_{t<T_{post}^{d^*}} \hat{\lambda}_t^{sdid} \big(Y_{d^*,t} - \sum_i \hat{\omega}_i^{sdid} Y_{i,t}\big)$. Neither set of

the construction decomposes the aggregate $\hat{\tau}^{sdid}_{d^*}$ exactly: the $\hat{\tau}^{sdid}_{d^*,r}$ average back to $\hat{\tau}^{sdid}_{d^*}$. For pre-treatment event times $r \in \{-1, -2, \ldots, -12\}$, the same construction yields placebo estimates that compare the SDID gap at an individual pre-treatment quarter to its $\hat{\lambda}$-weighted pre-period average. Plotted together (e.g., Figure 6), the placebos provide a visual check on pre-treatment fit: systematically detectable values would constitute evidence that the synthetic counterfactual fails to track the Western district's pre-treatment trajectory.

In addition to the choice of estimator, we must choose a donor pool of non-treated posts. We start with a donor pool that consists of all Baltimore posts outside of the Western district. Because we are interested in estimating the effects of GVRS over an outcome window of 18 months starting in January 2022, and because GVRS launched in the Southwestern district in January 2023 (precluding its posts from being viable donors after that point), we exclude posts in the Southwestern district from the donor pool throughout the analysis. As a robustness check, we further exclude all posts adjacent to the Western and Southwestern districts as well as all posts in the Central district, where pre-existing ties between groups in those areas and in the Western district raise the possibility of spatial crime spillover.

For inference, we compute clustered bootstrap standard errors from 500 iterations, given the robust performance of bootstrap inference for SDID and related methods (Arkhangelsky et al., 2021). We adjust our inference procedure to account for estimating the effects of GVRS across our four reported violent crime outcomes (homicides and non-fatal shootings, aggravated assaults, robberies, and carjackings) by reporting Benjamini-Hochberg-adjusted q-values (Benjamini and Hochberg, 1995) alongside the bootstrap p-values. These q-values control the false discovery rate across this family of tests.

5.3 Results

We present below our estimates of the effects of GVRS on rates of reported violent crime, possible geographic displacement of crime to other areas, arrest patterns, and willingness to

weights is re-estimated for any event time; standard errors come from a clustered bootstrap that re-runs the full SDID estimation, including weight optimization, on each resample.

report crime to police in the Western district, concluding with a discussion of robustness.

Violent crime Table 4 reports our main results: SDID estimates of the average treatment effect of GVRS on rates of homicide and non-fatal shooting victims, carjackings, aggravated assaults, and robberies in the Western district, with both bootstrap p-values and Benjamini-Hochberg-adjusted q-values across these four outcomes.

The clearest effects are on homicide and non-fatal shooting victims and on carjackings—the two offenses with the strongest theoretical and ethnographic link to the group-involved violence GVRS was designed to deter (Jacobs and Cherbonneau, 2023). Rates of homicide and non-fatal shooting victims fell by 39 per 100,000 over the 18-month window—30% of the implied counterfactual—with a bootstrap p-value of 0.009 and an FDR-adjusted q-value of 0.038, comfortably surviving a 5% false discovery rate control across our five tests.[18] Carjackings fell by 15 per 100,000, or 39% of the counterfactual mean; this estimate is statistically significant before adjusting for multiple testing ($p = 0.054$), but falls just short of conventional significance thresholds after adjustment ($q = 0.11$).

We find no statistically detectable effects on aggravated assaults or robberies. Both point estimates are modest relative to their counterfactual means (+7.9% and −5.5%, respectively), and both are far from conventional significance thresholds ($p = 0.22$ and $p = 0.63$, respectively).

Figure 6 plots quarterly SDID event-time coefficients and 95% confidence intervals for these four outcomes. For both homicide and non-fatal shooting victims and carjackings, all of the post-treatment coefficients are negative throughout the 18-month outcome window, with some of them being statistically significant. The results for aggravated assaults and robberies are less clear. For aggravated assaults, the last post-treatment coefficient estimate is positive, implying an adverse effect. For robberies, the pattern of post-treatment coefficient estimates suggest an initially beneficial effect that attenuates over time. Across all four outcomes,

[18] Effects estimated separately for homicide victims and non-fatal shooting victims are -27% ($p = 0.121$) and -32% ($p = 0.028$) of their implied counterfactuals, respectively.

the pre-treatment coefficients show no systematic trend that would raise concerns about a violation of the parallel trends assumption.

We show in Section 5.4 that the homicide/shooting and carjacking estimates are robust to alternative donor pools, inference procedures, and to replacing SDID with either SCM or DID. The robbery estimate is less stable: DID and SCM produce sizable negative effect estimates, while SDID does not. We argue in Section 5.4 that this discrepancy reflects a pandemic-era structural break in robbery rates in Baltimore's donor districts, which SDID's time weights are designed to absorb but which the equal time weighting of DID and the level-matching of SC do not fully reconcile with the Western district's own pre-treatment trend.

Spatial crime spillovers Next, we consider the possibility that these reductions in homicide and non-fatal shooting victims and carjackings were the result of these crimes being displaced to other areas of Baltimore. While there is now a great deal of evidence—including from randomized experiments of hot spots policing interventions—that offenders tend to be strongly tied to place and do not shift seamlessly to other locations (Weisburd et al., 2006), a small number of papers have found evidence of spatial crime spillovers in response to a shift in law enforcement presence (Blattman et al., 2021).

Figure 5 plots annual counts of homicide and non-fatal shooting victims, carjackings, aggravated assaults and robberies in different areas of Baltimore. The figure considers two candidate areas where spatial crime spillovers due to GVRS might occur: the Central district (which has known social ties to groups operating in the Western district) and posts that border the Western and Southwestern districts. Several patterns are worth noting. First, the levels of most violent crimes in Baltimore did not change dramatically immediately after the launch of GVRS in 2022. Relative to 2021, homicide and non-fatal shooting victims and robberies were slightly lower, while aggravated assaults were slightly higher. The exception is carjackings, which rose substantially in 2022, consistent with a nationwide trend in recent years (Robertson, 2022). Second, with the possible exception of carjackings, we see little evidence that violent crime increased in the Central district or neighboring posts in 2022. To more formally test for spatial crime spillovers, Appendix Table A.2 reports SDID treatment effect estimates on the same set of crimes in these potential spillovers areas, analogous to the main results in Table 4. All of the point estimates are negative, consistent with crime having decreased in these areas relative to the counterfactual. However, none of the estimates are statistically significant.

Arrests Next, we consider the effects of GVRS on enforcement activity, as measured by arrests. We are interested in testing two distinct theories for how enforcement may have contributed to the estimated reductions in violent crime reported above. First, police may have expanded the intensity of broad-based, routine enforcement, which could have a "net widening" effect of sweeping more people into the criminal justice system. Second, police may have concentrated their enforcement activity on violent offending specifically. Table 5 reports SDID treatment effect estimates for six arrest-related outcomes, grouped into two panels that speak to these two theories in turn.

Panel A shows no detectable increase in broad enforcement activity. The point estimate for total arrests in the Western district is a decline of 11 per 100,000 per quarter (2.0% of the counterfactual post-period mean), far from statistical significance ($p = 0.82$). Arrests with charges that do not include a violent index crime—a category of arrests that officers are typically thought to have greater discretion over whether to make—show a similarly imprecise null (-24 per 100,000 per quarter, or -4.9%, $p = 0.61$). For a subset of these arrests where officer discretion may play a particularly large role—drug possession or distribution arrests that are not co-charged with a violent index crime—the estimated effect is negative but not statistically significant ($p = 0.40$). Taken together, the Panel A results are inconsistent with GVRS having widened the net of routine enforcement in the Western district and rule out large increases in total, non-violent index crime, or discretionary drug arrests.[19]

[19] GVRS itself generated approximately 123 arrests in the Western district during the 18-month post-intervention window. The fact that total arrests did not rise therefore implies that these direct GVRS arrests were offset by declines in other categories.

Panel B shows a sharp increase in enforcement targeted at violent index crime. Arrests with charges that include a violent index crime rose by 31 per 100,000 per quarter, an 81% increase relative to the counterfactual post-period mean ($p = 0.001$). This pattern appears to be driven by an increase in aggravated assault arrests (+19 per 100,000 per quarter, or +47%, $p = 0.019$).[20]

The effect on violent crime arrests, on its face, appears to be very large but this estimate must be interpreted in light of the fact that the few of these crimes are cleared and so the base rate is very low. Referring to the final column of Panel B, the 81% increase in violent index crime arrests increased the violent index crime quasi-clearance rate, defined as the ratio of arrests to reported crimes, by just 2.8 percentage points—from a counterfactual post-period mean of 0.076 (or roughly one arrest per 13 reported violent index crimes absent GVRS) to an observed post-period rate of 0.104 (roughly one per 10), an effect that is marginally statistically significant ($p = 0.079$).[21] As such, the effect of GVRS on violent crime clearances is best thought of as large in percentage terms but nevertheless as only a modest change in the probability that a violent crime is solved through an arrest. Taken together, the Panel A and Panel B results are consistent with GVRS having redirected enforcement—away from the high-volume, routine arrest categories that generate most exposure to the criminal justice system and toward serious violence—without scaling up arrests overall.

[20] Appendix Figure A.2 plots quarterly SDID event-time coefficients and 95% confidence intervals for violent index crime and aggravated assault arrests. Neither outcome exhibits a systematic pre-treatment trend: the 12 quarterly placebo coefficients are individually close to zero and none are statistically distinguishable from zero. The two post-treatment profiles differ in ways that are informative. For violent index crime arrests (top panel), all six post-treatment point estimates are positive and of similar magnitude through the first year of the intervention, with modest attenuation over the final two quarters. For aggravated assault arrests (bottom panel), the point estimate is largest in the launch quarter (the first quarter of 2022) and declines gradually thereafter. In both cases the effect appears immediately upon the launch of GVRS rather than emerging gradually, consistent with a rapid reorientation of enforcement toward violent offending.

[21] We refer to this as a "quasi-clearance rate" because we cannot link individual arrests to the specific reported crimes they resolve in our data; we instead compute the ratio of violent index crime arrests in a post-quarter to reported violent index crimes in the same post-quarter. This ratio can exceed one in finite periods, but it is informative about the intensive margin of enforcement relative to reported offending.

Willingness to report gunfire Finally, we leverage data on ShotSpotter incidents and 911 calls to estimate whether GVRS affected willingness to report crime to the police. This outcome is important for two reasons. First, if GVRS eroded residents' willingness to engage with the police—for example, by generating friction in the course of enforcement actions or by signaling a heightened risk of unwanted police contact for bystanders—then this may be an additional cost of the strategy. Second, if GVRS reduced willingness to report, then the decline in reported crime that we observe could have occurred absent any real reduction in gun violence.

To assess this, we draw on data from the city's ShotSpotter acoustic sensor network, which is reporting-independent, and on 911 call records, which are reporting-dependent. Comparing the two lets us separate movement in actual gunfire activity from movement in reporting behavior, in the spirit of Ang et al. (2025). Table 6 reports SDID treatment effect estimates over the same 18-month post-treatment window for four outcomes, computed on the donor pool restricted to posts that were continuously inside the ShotSpotter sensor footprint from January 2019 through June 2023 (see Section 5.1). The resulting estimates should be interpreted as the effect of GVRS in the Western district relative to covered areas of Baltimore, not relative to the city as a whole.

The first three columns all point in the same direction. ShotSpotter incidents in the Western district fell by 8.6% relative to the counterfactual (59 fewer incidents per 100,000 per quarter), directionally consistent with the 30% reduction in homicide and non-fatal shooting victimizations documented above, though substantially smaller in percentage terms and not statistically significant at conventional levels ($p = 0.18$).[22] The estimate for rounds detected per 100,000 is much noisier and not statistically significant (-3.8%, $p = 0.69$), but is not inconsistent with the incident-level decline. 911 calls to the Baltimore Police Department fell by 9.1% (2,572 fewer calls per 100,000 per quarter), an effect that is marginally statistically

[22] The gap in percentage terms partly reflects the much higher base rate of ShotSpotter incidents, which include gunfire that does not result in physical victimization, than of homicide and non-fatal shooting victims. As a result, a reduction in the number of homicide and non-fatal shooting victims of a given magnitude will yield a smaller percentage change in ShotSpotter incidents.

significant ($p = 0.056$) and could in principle reflect either fewer underlying incidents to report or reduced willingness to call.

The critical diagnostic is the ratio of 911 calls to ShotSpotter incidents, following the measure used by Ang et al. (2025). If willingness to report fell—that is, if Western district residents responded to acoustically-detected gunfire by calling 911 less often—then this ratio would *fall*, with each detected incident generating fewer public calls. In fact, the 911 call decline is matched by a similar-magnitude percentage decline in ShotSpotter incidents, so the calls-to-incidents ratio is virtually unchanged: the point estimate is +0.28, or +0.5% of the counterfactual mean, and is essentially indistinguishable from zero ($p = 0.97$). We read this near-exact null as strong evidence against the interpretation that the observed declines in reported violent crime and 911 calls in the violent crime analysis above are an artifact of reduced public engagement with the police; acoustically-detected gunfire and public calls moved in lockstep, as would be expected if the underlying rate of gun violence fell without any change in reporting behavior.

5.4 Robustness

We test the robustness of our findings to two key estimation decisions: the choice of estimator (SDID, SCM, DID) and the choice of donor pool of non-treated police posts. Appendix Figures A.4 and A.5 report the full grid of point estimates and 95% confidence intervals for each outcome in Tables 4 and 5, respectively. For most outcomes, the three estimators and two donor pool choices deliver similar magnitudes and the same sign. The two outcomes where the estimators diverge meaningfully are robberies and drug arrests not co-charged with a violent index crime. The remainder of this subsection explains why this happens and why SDID is our preferred estimator in this setting.

Consider robberies first. The top panel of Figure A.1 plots the Western district's actual quarterly robbery rate alongside the counterfactual trajectories implied by DID and SDID, with SDID's estimated time weights shaded in gray on a secondary axis.[23] Both Baltimore's

[23] The implied counterfactual for the Western district at date t under estimator $s \in \{\text{DID}, \text{SDID}\}$ is $\widehat{Y}^{s}_{d^*,t} =$

and the Western district's robbery series see a structural break in 2020: both series have a higher mean in 2015–2019, then stabilize at a lower level from 2020 onward through the post-treatment period. SDID assigns greater weight to pre-treatment time periods in which the donor units' outcomes resemble those in the post-treatment period; in practice, this ends up assigning approximately 70% of the weight in 2020–2021. As a result, SDID's implied counterfactual is anchored to the later pre-treatment periods and tracks the actual Western district trajectory closely just before GVRS launched. In contrast, because DID weights all 84 pre-treatment periods uniformly, the larger resulting difference between the pre- and post-treatment periods drives down the implied counterfactual and inflates the DID point estimate.

To isolate the role of SDID's time weights, we re-estimate DID on a pre-period restricted to 2020–2021. The restricted DID ATT for robberies is -5.4% ($p = 0.64$), essentially identical in magnitude to the main-specification SDID estimate (-5.5%). With a common pre-treatment period that avoids the structural break, DID and SDID agree; the full DID estimate is an artifact of the 84-month uniform pre-period averaging across two distinct regimes.

The story for drug arrests not co-charged with a violent index crime is qualitatively similar, and the structural break more extreme. Arrests in Baltimore fell by nearly half from 2015–2019, driven by particularly steep drops in more discretionary arrests for offenses like drug possession and disorderly conduct that bottomed out in 2020 and have remained at historically low levels. This structural break is visible in the bottom panel of Figure A.1. As with robberies, DID's uniform weighting implicitly assigns more weight to earlier pre-treatment periods when drug arrest rates were much higher, while SDID explicitly assigns greater weight to later pre-treatment periods when drug arrest rates were much lower. As a result, relative to the SDID point estimate (-16%), the magnitude of the DID point estimate

$\overline{Y}_{d^*}^{\text{pre},\hat{\lambda}^s} + \sum_i \hat{\omega}_i^s \big[Y_{idt} - \overline{Y}_i^{\text{pre},\hat{\lambda}^s}\big]$, where the sum is over donor posts i; $\hat{\omega}_i^s$ and $\hat{\lambda}_t^s$ are estimator s's unit and time weights (uniform for DID, estimated for SDID); and $\overline{Y}_j^{\text{pre},\hat{\lambda}^s} = \sum_{t<T_{post}^{d^*}} \hat{\lambda}_t^s Y_{jdt}$ is unit j's pre-period average under s's time weights. The post-treatment average gap between the Western district's actual outcome and $\widehat{Y}_{d^*,t}^s$ recovers estimator s's ATT.

is artificially inflated (-56%). Restricting DID's pre-period to 2020–2021 nearly closes the gap fully (-13%).

6 Effects on GVRS subjects

To better understand the mechanisms through which GVRS may affect violent crime rates in the Western district, we estimate the intervention's effects on arrests and serious victimization among GVRS subjects. We first identify people who are observationally similar to GVRS subjects among a pool of potential controls—people who are not GVRS subjects themselves and who had not been previously arrested in the Western district—and then leverage the panel nature of our data to estimate event study specifications that compare changes in outcome rates among GVRS subjects and their matched controls relative to a reference period just before the subjects were identified for the intervention.

6.1 Data

The person-level analysis uses three types of administrative data: GVRS contact records, person-level arrest records from BPD, and homicide and non-fatal shooting victimizations.

GVRS contacts Records maintained by GVRS contain information on each subject and on the contact events the program initiates with them; we describe these records and summarize the activities they capture in Section 4. For our person-level analysis, the most important fields in these data are those capturing the date on which each subject was first identified for the intervention—either first referred for communication or first targeted for arrest—as well as the dates of any GVRS-recorded arrests. Because subjects are identified on an ongoing basis during BPD's weekly shooting reviews, each subject has their own first-referral date. Since our arrest and victimization data run through June 30, 2023, we focus the analysis on GVRS subjects identified in 2022 so that we can observe at least two quarters of outcome data for each of them. Of the 276 GVRS subjects whose first referral was in the Western district before July 2023, 173 were identified in 2022. This excludes 15 who were identified in 2021 during the preparation for the intervention's official launch the following year, and 88 who were identified in the first six months of 2023.

Arrests and victimizations We use non-public, linked, person-level data from BPD on the universe of arrests (from 2011 through June 30, 2023) and homicide and non-fatal shooting victimizations (from 2017 to June 30, 2023) in Baltimore. These data also contain basic demographic information about people such as race, age, and sex. We supplement BPD's arrest data with arrest events recorded in the GVRS contact file, harmonized into a single "any arrest" indicator. This is necessary because indictment-warrant arrests served against GVRS targets—typically the apprehensions made during takedowns—are not always entered into BPD's internal arrest viewer and are therefore absent from the linked person-level arrest data, even when the GVRS contact records show that an arrest occurred.[24] Our person-level arrest outcomes use this harmonized measure throughout.

6.2 Estimation strategy

We estimate the effects of GVRS on its subjects by matching each subject to one or more observationally similar controls and then estimating an event study specification on the resulting panel dataset, following the approach of Smith et al. (2019).

We begin by building a stacked panel dataset. Let t_0 represent the date in 2022 on which one or more GVRS subjects were identified for the intervention. For each unique t_0, we construct a panel of GVRS subjects identified on t_0 together with all candidate controls—people never identified for GVRS, never homicide victims, and not previously arrested in the Western district before t_0—and stack these t_0-specific panels together.

We identify matched controls separately within three strata of GVRS subjects, given differences between their underlying populations and in the treatments directed at them: (1) communication referrals who had a non-fatal shooting (NFS) victimization in the 90

[24] To construct the harmonized indicator, we match each GVRS-recorded arrest event to the nearest BPD arrest record for the same subject within ± 1 day using greedy bipartite matching (with a small set of hand-verified exceptions for matches at wider gaps). GVRS-recorded arrest events with no matching BPD arrest record are treated as additional arrest events. Of the 157 raw GVRS-recorded arrest events in our data, six records containing information indicating that they were not actually arrests are excluded; of the remaining 151, 99 collapse into BPD records and 52 enter the harmonized measure as GVRS-only arrests.

days preceding their referral, (2) communication referrals who did not, and (3) arrest target referrals.[25] Within the two communication referral strata, candidate controls must satisfy the same recent-NFS condition as their stratum's treated subjects.

Let $X_i^{t_0} = \{X_{1,i}^{t_0}, ..., X_{k,i}^{t_0}\}$ be a vector of k characteristics for person i defined relative to time t_0. The characteristics include 90-day quarterly counts of prior arrests by charge type and of NFS victimizations.[26] The four quarters immediately preceding t_0 (quarters -1 through -4) are deliberately *excluded* from the matching specification; we hold them out as a placebo window over which to test for conditional parallel trends in our event study below.[27] Limiting our stacked panel to one stratum together with the corresponding candidate controls, we estimate the propensity to be assigned to the GVRS treatment, $T_i = 1$, using ordinary least squares (OLS):

$$T_i = \beta X_i^{t_0} + \alpha Z_i + \gamma_{t_0} + \varepsilon_i.$$

In addition to the matching variables $X_i^{t_0}$, we control for demographic characteristics (Z_i) including race, age (in 10-year bins, plus a 60+ bin and a missing-age indicator), and sex. The identification date fixed effects (γ_{t_0}) ensure estimation occurs within stacks of the panel, i.e., among groups of people for whom $X_i^{t_0}$ is defined relative to a common date. Using the resulting fitted propensity scores $\hat{T}_i$, for each GVRS subject $i = 1, \ldots, N$, we calculate the absolute difference with each candidate control i' ($\Delta\hat{T}_{ii'} \equiv |\hat{T}_i - \hat{T}_{i'}|$), rank candidates in

[25] Stratifying communication referrals on recent NFS victimization is necessary because a sizeable share of communication referrals were recently victimized; without the stratification, the matched-control comparison would be confounded by the mechanical correlation between recent victimization and selection into communication referral. We do not impose an analogous restriction within the arrest target stratum, where the recent victimization rate is much lower.

[26] The charge categories we use are: homicide, sexual assault, robbery, carjacking, aggravated assault, simple assault, Part 1 non-violent crimes, weapons, drug possession, drug distribution, disorderly conduct, and other offenses. The time bins prior to t_0 are 90-day quarters: each of quarters 5 through 16 prior to referral contributes a separate count per charge type and for NFS victimizations, plus a single count covering all events at 17 or more quarters prior to referral. For example, a single element of $X_i^{t_0}$ is the count of person i's arrests that included a drug possession charge in the eighth quarter (i.e., 631–720 days) before t_0.

[27] We have also estimated effects under a matching specification that additionally includes the four held-out quarters, and find similar results; see Appendix Table 7 and accompanying discussion.

ascending order of $\Delta\hat{T}_{ii'}$, and keep up to 100 controls i' within a caliper of 0.01 ($\Delta\hat{T}_{ii'} \leq 0.01$).

After identifying up to 100 matched controls, i', for each GVRS subject, i, we estimate an event study specification using weighted OLS:

$$\Delta Y_{ii't} = \sum_{r \in \{-4,-3,-2,0,+1\}} \beta_r T_{it}^r + \varepsilon_{it},$$

where $\Delta Y_{ii't} \equiv Y_{it} - Y_{i't}$ is the difference in the outcome of interest between a GVRS subject and one of their matched controls in quarter t relative to t_0, T_{it}^r is an indicator for whether the GVRS subject was identified for the intervention r quarters ago, and $r = -1$ (the quarter immediately before referral) is the omitted reference period. The coefficients of interest are β_r, and we focus on β_0 and β_{+1}, the post-referral effects on the GVRS subject's outcome relative to that of their matched controls, analogous to intent-to-treat (ITT) effect estimates. The remaining coefficients ($\beta_{-4}, \beta_{-3}, \beta_{-2}$) cover the four-quarter pre-referral window held out from matching; statistically detectable values would constitute evidence against conditional parallel trends. Because each GVRS subject can have a different number of matched controls, we ensure that each subject contributes equally to the estimation by assigning each pair-quarter observation a weight equal to one over the number of matched controls assigned to that subject. We compute heteroskedasticity-robust standard errors clustered at the GVRS subject level. Our outcomes are an indicator for any arrest in the quarter (using the harmonized BPD/GVRS arrest measure described above) and an indicator for any homicide or NFS victimization in the quarter.

The identifying assumption for our matching difference-in-differences estimator is conditional parallel trends: that, conditional on the matching variables, treated subjects and their matched controls would have followed parallel trajectories in the absence of treatment (Heckman et al., 1998; Smith and Todd, 2005). Because our matching specification holds out the four quarters immediately before referral, the post-treatment coefficients we estimate at quarters 0 and +1 rest on a parallel trends test that the matching procedure does not have the information to mechanically satisfy.

6.3 Results

Our analysis sample comprises GVRS subjects whose first referral was in the Western district and occurred in 2022 ($N = 163$), matched via propensity score to candidate controls drawn from the broader Baltimore population. Appendix Table A.3 traces the funnel from the broad sample of all Western district GVRS subjects with a first referral before July 2023 (the descriptive sample of Table 2) down to the matched analysis sample.

Table 7 reports baseline characteristics of these 163 GVRS subjects alongside their matched controls, separately for the three matching strata: 13 communication referrals with a recent NFS victimization, 90 communication referrals without one, and 60 arrest targets, matched to 79, 6,302, and 4,766 controls, respectively. The propensity score matching produces well-balanced samples on the matching variables—prior arrests by charge type and NFS victimizations measured in 90-day bins beginning five quarters before referral. In the two larger strata, the share of GVRS subjects with any arrest more than one year before referral matches almost exactly the share among their matched controls, and average prior arrest counts by charge type are similarly close. The two groups are broadly balanced on demographics as well, though GVRS subjects average one to two years younger than their matched controls and are slightly more likely to be Black—reflecting the difficulty of perfectly matching a young, almost-uniformly-Black treated group to a more heterogeneous candidate pool, even after conditioning on demographics in the propensity score. The four quarters immediately preceding referral, which the matching procedure deliberately holds out, show somewhat larger differences—most visibly in the small recent-NFS communication referral stratum, where 46% of GVRS subjects had any arrest in the year before referral compared to 20% of their matched controls. We use these held-out quarters as a placebo test for conditional parallel trends in the event study below.

Effects on arrest Figure 7 (top row) plots quarterly event study coefficients on an indicator for any arrest, with corresponding ATTs reported in Panel A of Table 8. The pattern differs sharply between the two referral types.

For communication referrals (top-left panel), the post-treatment coefficients are small and statistically indistinguishable from zero: the ATT at the quarter of referral is +0.01 (s.e. 0.04), and at the subsequent quarter it is essentially zero (s.e. 0.05). The placebo coefficients in the held-out window are similarly small, consistent with the matching successfully balancing pre-trends in this stratum. Read literally, this null might suggest that the deterrence-focused messaging communication referrals received does not translate into a measurable change in their arrest rates. We caution against this interpretation on two grounds. First, the estimate is too imprecise to speak to a *decrease* in arrests of any plausible magnitude: with a counterfactual quarterly arrest rate of roughly 7% and standard errors of 4–5 percentage points, the smallest decrease detectable at conventional power—about 11–13 percentage points—exceeds the counterfactual arrest rate itself, so even a complete cessation of arrests among these subjects would not register as statistically significant. The design can detect only large *increases* in arrests, of the kind that heightened enforcement scrutiny might produce. Second, even a genuinely unchanged arrest rate is consistent with reduced offending if communication referrals changed their behavior in ways that lowered offending while simultaneously increasing law enforcement scrutiny; and our short post-treatment window may in any case miss longer-run shifts. The qualitative evidence in Section 7 is consistent with reduced offending among these subjects.

For arrest targets (top-right panel), the contrast is dramatic. The ATT in the quarter of referral is +0.74 (s.e. 0.06, $p < 0.001$), corresponding to roughly an eleven-fold increase relative to the matched-control mean of 0.07. The pattern of charges underlying this spike is informative, though it can be characterized only for the roughly four-fifths of these arrests that appear in BPD's records with charge detail; the remainder are indictment-warrant apprehensions recorded only in the GVRS contact data, whose underlying charges we do not observe and which plausibly include serious violent offenses. Among arrests with recorded charges, weapons and drug distribution charges together account for the overwhelming majority of the increase, with simple assault, aggravated assault, and homicide charges contributing smaller but meaningful shares. Property and disorderly conduct charges, and other offense categories, do not contribute appreciably. This concentration in serious gun- and drug-trafficking offenses is consistent with GVRS's stated focus on group-involved violence rather than a broad-based enforcement push.

The arrest target ATT at the subsequent quarter is +0.12 (s.e. 0.05), substantially smaller than the quarter of referral effect but positive and statistically distinguishable from zero. Two features of this estimate are informative. First, if arrest target subjects were incapacitated for a prolonged period due to pretrial detention or post-conviction incarceration following their arrest, then their arrest rates in the subsequent quarter would be mechanically *lower* than those of their matched controls and the ATT would be negative. The fact that the subsequent quarter ATT is positive instead suggests that most arrest target subjects are released within the same quarter and remain at risk of subsequent enforcement contact. Second, the persistence of an elevated arrest rate one quarter after referral is consistent with continued GVU enforcement attention on these subjects—for example, follow-on indictment warrants in extended group cases or additional arrests of subjects whose initial takedown is recorded in our data on its arraignment date but whose remaining co-defendants are picked up over the following weeks.

Effects on shooting victimization Figure 7 (bottom row) plots the same event study coefficients for an indicator of homicide or NFS victimization in the quarter, with corresponding ATTs in Panel B of Table 8. We find no statistically detectable effects of GVRS on shooting victimization for either communication referrals or arrest targets over the post-treatment window. The point estimates are uniformly small in magnitude—under two percentage points in absolute value across cells—and confidence intervals comfortably include zero. The placebo coefficients in the held-out pre-treatment window are similarly small and provide no evidence against parallel trends.

Two design features bear on the interpretation of these victimization estimates. First, a subset of communication referrals are subjects who experienced a NFS victimization in the 90 days preceding referral; matching is conducted within this substratum, and the resulting analysis combines those with recent shooting victimizations and those without. Both communication referrals and their matched controls in the recent-victimization substratum mechanically register a NFS in the quarter prior to referral (this drives the peak in this quarter visible in Appendix Figure A.6, but since both groups have this in common it contributes nothing to the post-treatment ATT). Second, pre-period shooting victimizations in our data are non-fatal by construction, since people who were homicide victims before a given referral date are excluded from the matching pool; post-treatment values include both fatal and non-fatal events. With these features in mind, the appropriate interpretation of these estimates is that we cannot detect any change in subsequent homicide or NFS victimization for either referral type over the available post-treatment window.

7 Interviews and focus groups with GVRS subjects

We complement the quantitative analyses above with qualitative evidence drawn from interviews and focus groups with GVRS subjects in the Western district. The aim is exploratory: to characterize the mechanisms through which the intervention may have shaped subjects' behavior—deterrence and incapacitation, social service delivery, and community involvement and procedural justice—and to surface interpretations that the administrative-data analyses cannot.

7.1 Data and method

Qualitative interviews and focus groups were conducted with GVRS subjects in the Western district to develop a better understanding of how the intervention may have changed their violent behavior. In consultation with Roca and YAP, purposive sampling was used to select gang members who were targeted by GVRS, received custom notifications, and were actively participating in social service programs. Two focus groups were held with GVRS service recipients: one with Roca clients ($N = 6$) and one with YAP clients ($N = 6$). Semi-structured individual interviews were completed with an additional twelve clients (six from each organization). Data collection occurred between January and March 2024; interviews and focus groups typically lasted 60–90 minutes and were audio-recorded and professionally transcribed with participant consent.

Qualitative data were managed and analyzed using NVivo 14 software to facilitate systematic coding and retrieval of key insights. Interview and focus group transcripts were analyzed using a two-stage thematic coding process (Saldaña, 2021). "Open coding" was used to identify initial emergent themes in participant narratives. This was followed by "focused coding" explicitly organized around five key GVRS components: deterrence, incapacitation, social service delivery, community involvement, and procedural justice. The goal of this two-stage analytic strategy was to provide some exploratory insights on how these program components may have influenced the behavior of GVRS subjects. The perceptions of a small number of people in a purposively selected sample are not necessarily representative of the broader population of GVRS subjects. As such, ample caution is needed when making inferences about focused deterrence violence reduction mechanisms based on these data.

7.2 Findings

Deterrence and incapacitation Interview participants consistently reported that GVRS implementation immediately increased their perceived risk of apprehension. After learning that they were "on the radar" for GVRS attention, interview subjects reported quickly changing their normal routines to evade possible law enforcement surveillance. As Marcus described:

> I came home one day and my Dad was like some people from the city came by here looking for you... I ain't gonna lie, as soon as I heard that that shit had me

> paranoid for real... So I started cutting myself from everybody, for real. Even stopped going to the store by my house cause I ain't want them to catch me out there.

As suggested by this quote, short-run simple evasion activities evolved into more strategic network avoidance (Fader, 2021), a calculated form of risk management with the potential to reconfigure social connections that facilitate violence. As subjects developed further knowledge of their enhanced punishment risks, many made deliberate decisions to distance themselves from specific high-risk associates and avoid locations vulnerable to police surveillance.

For GVRS subjects already under state supervision (probation, parole, or pretrial monitoring), the changes in perceived risks were particularly pronounced. These individuals, already navigating restrictions on their behavior, experienced GVRS as an additional layer of scrutiny that significantly amplified their punishment risk calculations. Andre described his experience receiving a custom notification while at his parole office:

> When they hit me with that letter at the parole office, that really woke me up. My [parole officer] was sitting there while they were reading it, so I knew it wasn't no joke. I left out of there knowing I had to move different... not that I was moving crazy or nothing, but I had to cut certain people off. I told my girl I can't even go to certain parts of the city no more. I used to be out on Pennsylvania Ave heavy, but after I got that letter, I ain't step foot over there... My PO already told me one dirty urine or one wrong person in the car with me, and I'd [get violated]... now I got these GVRS people watching too?

The direct communications of apprehension risks for continued violent behavior were supported by follow-up enforcement actions as needed. Arrests and subsequent prosecution provided tangible evidence that the GVRS enforcement threats were credible. Beyond deterrence, these actions also removed people who were likely to commit shootings and be shot by rivals from the street. As suggested by Jason, GVRS incapacitation efforts involved a deliberate, methodical approach to building cases against targets that was more robust than previous routine enforcement actions in the Western district:

> I feel like the police are going harder to catch people... I've been seeing a lot more people going to jail. Just because they'll let you go on a three or four month run, but they've been watching you the whole four months, just racking up evidence... then they'll come get you that last fifth month. A couple of my boys went down just like that...

Participants further recognized that their past actions might already be documented, creating greater concern that accumulated evidence could trigger future enforcement. This understanding seemed to further increase perceptions of apprehension risk, as individuals could no longer rely on the absence of immediate consequences as evidence of successful evasion. After a GVRS investigation resulted in the apprehension of his close friend, Deon observed, "He got life... he ain't never coming home... I got lucky. I [was only] locked up for a gun."

The impact of these targeted removals extended beyond individual deterrence and incapacitation to create structural voids in violent networks. When key members were arrested and incarcerated, their absence disrupted established patterns of coordination, leadership, and decision-making within their groups. Interview participants suggested that targeted groups attempted advanced novices within their groups to replace those removed; however, it was difficult to do so in the wake of enforcement actions. These individuals often lacked the experience, connections, and credibility of their predecessors, further weakening group capacity for coordinated violence. Further, as multiple targeted individuals altered their routines—reducing their presence in high-risk locations, limiting their social connections, and disengaging from violent conflicts—changes in their behavior rippled outward, disrupting the very social networks that sustain street violence. In effect, GVRS did not simply deter and incapacitate individuals; it seemed to destabilize the broader ecology of violence in the Western district.

Social service delivery Street outreach workers and social service agencies provided an important complement to GVRS enforcement efforts by connecting subjects to concrete services and opportunities. These tangible offerings provided practical alternatives to persisting in risky illegal street economic activity that made them vulnerable to GVRS sanctions. As Jermaine described:

> They can get you anything you need, if you come in and you don't have a GED, they help you get that. If you want to work, they can help you with certifications... They got a lot of resources. Even if they don't offer the programs, they will definitely put a good word in and connect you somewhere else if they see that you're serious. They're going to put their faith and their hope in you and push you, but they going to also leave enough room for you to feel as though that you did it yourself... when something is handed to you too easy, you feel like it's not yours. But they just set you up so you can set yourself up. That's what I like about them.

Beyond addressing underlying individual needs such as employment, housing, education, and substance abuse counseling, interview participants viewed services as facilitating transitions into new social networks that reinforced disengagement from violence. Tyrell described how this process worked:

> This program gave me more of an appreciation for surrounding myself with people who want the same as you... and not just people who want something from you. They basically teach us to be around people that want to be something better in life, not just being out in the streets... Now, I just try to surround myself with more people that want the same as me or even better than me. I surround myself with good people that got good mindsets...

Tyrell's insight reveals how programming provided an opportunity for mutual growth in peer groups ("people that want the same as me or even better") that collectively reinforced behavioral change across multiple participants and addressed the relational vacuum that often undermines individual attempts to disengage from street networks.

Participants reported that life coaching and cognitive behavioral therapy helped them avoid negative thinking and make better life decisions. Therapists and street outreach workers adopted a practical harm reduction approach that acknowledged the realities of dangerous environments while encouraging incremental behavioral changes by the GVRS subjects. Darryl described this nuanced process:

> Being at Roca made me look at violence different... it made me more conscious of how I think about it. I'm not like how I used to be... I don't feel like I'm going to go kill somebody when there's really not a real purpose or reason... But being in the city we're in, you always still gonna have some of that mindset... you can't really take that from us, because we've been seeing it our whole lives. It sounds a little crazy, but when you grew up seeing so much... it's hard to shake that mindset... but here we learn to minimize it... to tame it.

Darryl's reflection describes developing metacognitive awareness ("more conscious of how I think about it") that enabled greater behavioral regulation within his existing identity framework. The idea of "taming" rather than eliminating violent impulses recognized that exposure to violence creates deeply embedded cognitive patterns that cannot simply be unlearned or denied. By focusing on regulation rather than replacement, these services created space for participants to maintain aspects of street identity necessary for navigating their environments while reducing the likelihood of perpetrating violence to deal with problems.

Community involvement and perceptions of legitimacy GVRS leveraged existing community structures and relationships to reinforce anti-violence messages. Community voices—including clergy, elders, returning citizens, and family members—provided local sources of informal social control and moral engagement that complemented law enforcement interventions. As James described:

> One of the older guys from my neighborhood was there, and he told me, ’These people aren’t here to lock you up. They’re here to keep you safe. Hear them out.’ And because it was him saying it, I actually listened.

This anecdote illustrates how community members not only exerted influence over the willingness of GVRS subjects to participate in anti-violence programming but also bridged the legitimacy gap between formal authorities and subjects with high levels of institutional distrust. For people socialized to view law enforcement with suspicion, these community voices provided alternative pathways to engagement that didn’t require immediate trust in formal authorities. The phrase “because it was him saying it” highlights how message effectiveness depends not only on content but on messenger credibility—a resource that community members often possess in abundance within neighborhoods where official authority may be contested.

Community participation at call-ins also seem to activate emotional and moral dimensions of decision-making that formal deterrence messages do not address. Community members described the trauma of persistent shootings on neighborhood well-being and expressed grave concern over losing young people to senseless violence and to the criminal justice system. The words of a mother who lost her son profoundly affected Michael:

> She talked about raising her grandkids alone, how her son was gone, and how she didn’t want to see any of us put our mothers through that same pain. That hit different. I kept thinking about my own mom and what she would go through if something happened to me.

The phrase “that hit different” suggests that these kinds of appeals activate different appraisals of risks and rewards associated with continued violent behavior (e.g., “my mother would suffer as a result of my actions” relative to “I would be punished for my behavior”).

All GVRS subjects expressed strong feelings of mistrust and cynicism towards the Baltimore Police Department during interviews. Nevertheless, these subjects seemed to appreciate that the GVRS represented a fundamentally different approach to violence prevention than past enforcement efforts. Kevin described his evolution from initial rejection to a positive acknowledgment:

> When they first came at me with that letter I ain't even want to hear it. Like, here go the police again... But then the YAP lady came and started talking about opportunities, about how they could help me get right... they let me know the cops [are] watching me, they know what's going on, but they also trying to give me a chance to move different.

This acceptance seemed to stem from a key feature of the strategy that distinguished it from previous law enforcement approaches: a respectful and transparent warning of stiff enforcement in response to very specific behaviors (i.e., continued gun violence) coupled with genuine offers of help and concern. The inclusion of non-law enforcement messengers provided alternative sources of legitimacy that circumvented established distrust with the criminal justice system.

The participants' narratives described a bounded legitimacy granted to the GVRS approach: many GVRS subjects complied with the anti-violence message promoted by the police despite participants' continued reports of abusive and overly aggressive police behavior in other contexts. This specific, intervention-focused legitimacy created sufficient acceptance of the procedural approach to violence prevention that facilitated GVRS subject engagement with services, even without resolving broader concerns about policing practices. The legitimacy of the GVRS approach in the eyes of the treated group members seemed to bolster the credibility of the deterrence message in a way that made service uptake more attractive. Other studies that have found similar results highlight the importance of procedurally-just communication of carrots and sticks in focused deterrence strategies (Papachristos et al., 2007; Trinkner, 2019).

8 Baltimore's homicide and violent crime rates relative to peer cities

The within-Baltimore empirical evidence presented in Sections 5 and 6 suggests that GVRS substantially reduced gun violence in the Western district during its first 18 months of operation. GVRS implementation began to expand to other parts of Baltimore in the years that followed, including the Southwestern district (January 2023), the Central district (January 2024), the Eastern district (April 2024), and the Southern district (July 2025).[28] Over this same window, Baltimore experienced a 60% decline in its homicide rate between 2022 and 2025, outpacing the homicide rate drop in most other U.S. cities. In this section, we compare changes in homicide and violent crime rates in Baltimore and its peer cities over this period and discuss how much GVRS may have contributed to any of those differences.

8.1 Data

Because the FBI's National Incident Based Report System data are released with a considerable time lag, we draw on monthly crime data from the Real-Time Crime Index, a continuously updated platform that aggregates and standardizes reported crime statistics from participating law enforcement agencies across the U.S.[29] The estimation panel covers the 32 quarters from January 2018 through December 2025 for the 252 municipal police departments serving populations of at least 100,000 with complete homicide or violent crime data over the window.[30]

We estimate effects on two outcomes: homicides and non-fatal violent index crimes (rape,

[28] Since the continued rollout of GVRS prioritized the areas in Baltimore with the highest levels of violence and since a large portion of the city has now been exposed to the intervention, it is not possible to identify a credible counterfactual to evaluate the broader city rollout of GRVS.

[29] The Real-Time Crime Index, developed by AH Datalytics, is publicly accessible at https://realtimecrimeindex.com. Its data are not subject to some of the limitations of FBI Uniform Crime Reports data during the transition to the National Incident-Based Reporting System and they allow us to include 2025 in our analysis sample.

[30] We restrict to municipal police departments to maintain comparability with Baltimore's jurisdictional structure, dropping county sheriffs, state police, and other non-municipal agencies. We further drop agencies with multi-month gaps in their crime series; single-month gaps are filled by linear interpolation.

robbery, and aggravated assault). All outcomes are expressed as quarterly counts per 100,000 population, annualized (quarterly count $\div$ population $\times$ 100,000 $\times$ 4). Separating homicides from the other violent index crime components allows us to test whether any city-level effect is concentrated in lethal and near-lethal violence, as the within-Baltimore evidence in Section 5.3 would suggest.

8.2 Estimation strategy

We apply the same synthetic difference-in-differences (SDID) estimator used in our within-Baltimore analysis (Section 5.2), but with U.S. cities (rather than Baltimore police posts) as the units of observation. Baltimore is the single treated unit, with a treatment date of January 2022, the launch of GVRS in the Western district. Donor units are the other agencies in our panel.

Because the choice of donor pool can meaningfully affect SDID estimates with a single treated unit, we estimate effects using two donor pool restrictions designed to retain agencies with broadly comparable pre-treatment violence levels and population scale. First, we restrict the donor pool to agencies with pre-2022 mean rates of the corresponding outcome above the median across all eligible donors; this restriction keeps Baltimore in the upper portion of the rate distribution where the available comparison units are themselves above-average-violence cities.[31] Second, we restrict the donor pool to agencies serving populations of at least 250,000, ensuring that Baltimore (population approximately 565,000) is compared to other large municipal jurisdictions.

Inference is based on placebo p-values computed by re-estimating the model treating each donor unit as the treated unit in turn (500 placebo iterations).

8.3 Results

Table 9 reports the main SDID estimates. The estimated effect on Baltimore's homicide rate is large, negative and highly statistically significant using either of the two donor pools.

[31] Pre-2022 medians (excluding Baltimore) are 5.4 homicides per 100,000 residents and 432 non-fatal violent index crimes per 100,000 residents, both annualized.

Using the above-median pre-2022 rate donor pool (125 donor cities), Baltimore's annualized homicide rate over January 2022 through December 2025 was 13.6 per 100,000 below the synthetic counterfactual—a 24.9% decline relative to the implied counterfactual mean of 54.6 per 100,000 (placebo $p < 0.001$). The estimate from the population-restricted donor pool (83 donor cities) is essentially identical: a decline of 13.5 per 100,000 (-24.7%, $p < 0.001$). The top panel of Figure 8 plots Baltimore's actual quarterly homicide rate alongside the SDID counterfactual using the above-median donor pool. Prior to 2022, the two series demonstrate similar trends and seasonality. However, starting in 2022, a gap opens up between Baltimore and its synthetic comparison and widens through 2025. The pre-treatment fit is reasonable: SDID's time weights load principally on the second half of the pre-period, the years closest to the treatment date, and over those quarters the actual and synthetic series track each other closely.

The estimated effect on the non-fatal violent crime rate, by contrast, is small in percentage terms and statistically indistinguishable from zero in both donor pools. The point estimates are -30 per 100,000 (annualized) in the above-median donor pool (-1.7% of the counterfactual mean, $p = 0.81$) and -59 per 100,000 in the population-restricted donor pool (-3.1%, $p = 0.54$). The bottom panel of Figure 8 shows that Baltimore's non-fatal violent crime trajectory tracks the synthetic counterfactual closely throughout the post-treatment period, with no widening gap analogous to the one observed for homicides. The contrast between the two outcomes mirrors the within-Baltimore findings of Section 5.3: in both the cross-post and cross-city analyses, the clearest treatment effect appears in lethal violence rather than in non-fatal violent index crimes.

Two patterns are worth noting about the magnitude of the city-level homicide effect. First, the cross-city ATT (a $\sim$25% reduction in Baltimore's overall homicide rate) is substantially larger than what a Western district-only effect would mechanically produce at the city level. The Western district accounts for roughly 16% of Baltimore's homicides and non-fatal shootings, despite containing only about 5% of the city's population, so a 30% reduction in homicide and non-fatal shooting victims in the Western district alone (the within-Baltimore ATT from Section 5.3) would translate into only a ∼5% citywide reduction. The fact that the citywide ATT is much larger is consistent with the geographic expansion of GVRS to additional districts after 2022 and the broader change in law enforcement culture that this has produced. Second, the cross-city analysis cannot, on its own, isolate the contribution of GVRS proper from any other contemporaneous changes in Baltimore policing or in the city's broader environment; the SDID estimate captures the full Baltimore-specific deviation from the synthetic counterfactual, of which GVRS is one component. While other explanations have been offered for Baltimore's recent homicide decline, including a change in the elected state's attorney, the within-Baltimore evidence—that the homicide decline has been greatest in the Western district—and the cross-city evidence—which shows that only homicides fell in Baltimore and not violent crimes more broadly—is consistent with GVRS playing an important role.

8.4 Robustness

Appendix Table A.5 reports SCM and two-way fixed effects difference-in-differences (DID) estimates as robustness checks. The headline homicide finding survives the move to DID and is in fact slightly larger in magnitude: the DID estimates are −16.8 per 100,000 (−29.1%, $p < 0.001$) in the above-median donor pool and −17.1 per 100,000 (−29.5%, $p < 0.001$) in the population-restricted donor pool. SCM, on the other hand, produces much smaller homicide estimates that are not statistically distinguishable from zero (−4.6, $p = 0.12$ in the above-median pool; −1.5, $p = 0.66$ in the population-restricted pool). For the non-homicide violent crime rate, DID delivers small, negative, statistically insignificant estimates similar to the SDID baseline (−7% of the counterfactual in both donor pools), while SC produces large positive estimates of approximately +316 per 100,000 (+21%, statistically significant in both donor pools)—i.e., the SC estimator implies that Baltimore's non-homicide violent crime rate *exceeded* its synthetic counterfactual by roughly a fifth over the post-treatment window.

The discrepancy between SCM and the other estimators reflects the same mechanics described in Section 5.4 for the within-Baltimore analysis: SCM weights donor units to match the treated unit's pre-treatment level using only cross-sectional weights and no time weights, which makes its counterfactual sensitive to structural breaks in the pre-period and to whether the donor pool's post-treatment trajectory diverges from the treated unit's pre-treatment trend. The non-homicide violent crime SCM counterfactual (∼1,491 per 100,000) is far below the SDID counterfactual (∼1,838) and the actual Baltimore post-treatment mean (∼1,807); the synthetic match is being constructed from donor cities whose violent crime rates fell substantially over the post-period, generating a large positive treated-versus-synthetic gap that is unlikely to reflect a genuine treatment effect on Baltimore. Because both SDID and DID generate substantively similar conclusions for both outcomes—a large, robust homicide effect and a near-zero non-homicide violent crime effect—we treat the SCM estimates as informative about estimator sensitivity rather than as a competing substantive finding.

A natural concern, given Baltimore's elevated homicide rate in Figure 1, is that its decline reflects mean reversion: homicide rates rose across many U.S. cities during 2020–2021, and those that rose most might be expected to fall most thereafter. Two features of the data argue against this reading. First, Baltimore had almost no pandemic-era homicide rate spike from which to revert. Its annualized homicide rate averaged 57.9 per 100,000 in 2018–2019 and 60.0 in 2020–2021—an increase of about 2 per 100,000, negligible when compared to the swings in many cities and to Baltimore's own subsequent decline. Baltimore's rate was persistently high for years before the pandemic, so its post-2021 fall is a break from a long-standing level rather than the unwinding of a transitory increase. Across donor cities, the post-2021 decline is strongly predicted by the size of the 2020–2021 run-up, and Baltimore's negligible run-up would imply almost no reversion (Appendix Figure A.8). Second, although higher-homicide cities did on average decline somewhat more after 2021—a mean-reversion gradient present in our donor pool—Baltimore's decline far exceeds what that gradient predicts. Regressing each donor city's 2022–2025 change in homicide rate on its pre-2022 level yields a slope of about 0.11; at Baltimore's pre-2022 rate this implies a decline of 6–8 per 100,000 against the roughly 18 observed, leaving Baltimore between 3.7 and 3.9 standard deviations below the fitted line and with a larger raw decline than all but one of the 125 comparison cities (Appendix Figure A.7). This conditional gap, about 10–12 per 100,000, is similar in magnitude to our SDID treatment-effect estimate of 13.6, reflecting that the synthetic counterfactual already nets out the common reversion. Finally, mechanical reversion would be expected across violent crime categories, yet Baltimore shows no comparable deviation in non-fatal violent crime (Section 8.3).

9 Discussion

Investments in law enforcement can be effective in reducing violence but can entail high financial and social costs, including a widening of the net of the criminal justice system by exposing larger numbers of people—including many low-level offenders—to the destructive effects of arrests and prosecution (Chalfin, 2025). The fact that a small number of offenders, often operating in groups, drive an outsize share of gun violence suggests that a more targeted approach—one that concentrates both enforcement and social services on the people at highest risk—has the potential to generate large public safety benefits while minimizing the harms of aggressive and broad-based policing strategies (MacDonald, 2023).

We study the 2022 launch of a targeted strategy to reduce gun violence in Baltimore's Western police district, an area with among the persistently highest rates of gun violence in the U.S. that has, in the past, appeared intransigent to efforts to reduce violence. The approach uses a mixture of "carrots and sticks" to try to change the behaviors of people thought to be at high future risk of carrying out shootings and, accordingly, the strategy was implemented through the collaboration among law enforcement, community-based, and social service organizations.

In the 18 months after the strategy was implemented in the Western district, the number of people shot and killed there declined by approximately one third, and carjackings by approximately 40%, compared to other areas of Baltimore that were, prior to the intervention, equally challenged by endemic violence. We observe no evidence that violence spilled over to other Baltimore communities outside of the Western district. Importantly, the large reductions in gun violence the strategy produced were not accompanied by a broad-based increase in arrests in the Western district: total arrests, including for non-violent and lower-level offenses, did not change relative to the counterfactual, ruling out a "net widening" expansion of enforcement that would have swept additional people into the criminal justice system. However, arrests for serious violent crimes rose and increased the clearance rate from about 8% to 10%. The evidence thus suggests that by focusing attention on the principal drivers of violence, a combination of law enforcement and community partners can meaningfully reduce gun violence, in part by *redirecting* enforcement toward serious violence rather than ramping up enforcement for the community writ large. These effects mirror findings from research on large-scale "gang takedowns" (Chalfin et al., 2021; Choe et al., 2026; Domínguez, 2026) but, importantly, while gang takedowns deliver only the stick, this strategy tries to reduce harm by warning others of the enforcement they will face if they continue offending, offering low-barrier supportive services to those willing to step away from violence, and engaging community members as key actors to prevent future violence.

With respect to the mechanisms behind the observed violence reductions, the evidence is mixed and points to several channels operating together. From the administrative data, we cannot detect a change in the likelihood of arrest among GVRS subjects who received only communication. This null is largely uninformative about deterrence, however: the arrest rate among these subjects is low enough that our estimates could not detect a reduction even if their arrests had ceased entirely (Section 6.3). The administrative data can detect only large *increases* in arrests, of the kind heightened enforcement scrutiny might produce, leaving the deterrence channel to be assessed primarily through the qualitative evidence. By contrast, we observe a sharp, immediate rise in arrest probability among subjects targeted for arrest, indicating an immediate incapacitation effect, though the lack of a sustained decrease in arrests one quarter later suggests this effect may be modest in duration. Qualitative evidence from interviews and focus groups with GVRS subjects points to behavioral change consistent with deterrence—participants reported a heightened perception of apprehension risk and deliberate distancing from high-risk associates and locations—as well as meaningful uptake of services that supported transitions away from violence and a central role for credible community messengers, whose legitimacy with group members exceeded that of law enforcement, in making both the strategy's warnings and its service offers credible. The timing of the two data sources may also bear on how to weigh them against each other. The administrative data analysis follows subjects identified in 2022 through mid-2023, when GVRS's outreach and service infrastructure was still being built out, whereas the interviews and focus groups were conducted in early 2024, after the program had matured. If the strategy's effects on subjects strengthened as that infrastructure developed, then the early administrative data window may capture a less mature, less effective phase of the program than these accounts reflect. Read together, the administrative and qualitative evidence suggest that deterrence, incapacitation, services, and the legitimacy of community-based messengers each plausibly contribute to the observed reductions, with the most direct quantitative evidence bearing on incapacitation and the evidence on the remaining channels resting largely on subjects' own accounts.

The plausibility of a deterrence channel is reinforced by how weak the baseline enforcement environment was. Absent GVRS, the clearance rate for violent index crimes in the Western district is estimated to have been just 7.6%—more than nine in ten reported violent crimes would not result in an arrest. GVRS raised it to 10.4%: an increase of 2.8 percentage points but, against so low a base, a proportional increase of roughly 37%, alongside an 81% rise in violent index crime arrests. Whether a shift of this kind alters behavior turns on whether those at risk perceive it—and here the strategy's direct communication is likely central, conveying the changed enforcement environment immediately rather than leaving it to be inferred slowly from accumulated experience (Section 4.2).

Two of these channels—incapacitation and the take-up of services—can be assigned rough magnitudes. Neither, on its own, appears to account for the bulk of the reduction. Over the 18-month window, we estimate that GVRS averted approximately 71 homicide and shooting victimizations in the Western district.[32]

Take the incapacitation channel first. The 110 people arrested directly through GVRS enforcement during the 18-month study period represent about 0.4% of the Western district's population but roughly 16% of the estimated 615–735 active group members in the district as of 2021.[33] The risk borne by such individuals can be strikingly concentrated: Heller et al. (2024) find that, among the 500 people at highest risk of being shot in Chicago, 13% were shot within 18 months. If the GVRS arrestees faced comparable risk and half of the 110 were incapacitated over our 18-month outcome window, preventing their own victimization alone would avert roughly 7 shootings—about a tenth of the 71—and the true contribution is plausibly larger, since incapacitation would also forestall shootings these individuals might have committed against others. The gang takedown literature points the same way: Chalfin et al. (2021) and Choe et al. (2026) find that arresting roughly 0.1% of a community's population during a takedown—with no accompanying message or services—produced substantial reductions in homicides and shootings, indicating that incapacitation can be powerful even without the deterrence and service components GVRS adds.

Next, consider the services channel. Of the GVRS subjects, 90 took up services through Roca or YAP, 86 of them receiving counseling or life coaching that often incorporated cognitive behavioral therapy. The closest experimental benchmark is READI Chicago, which offered up to 18 months of subsidized employment alongside cognitive behavioral therapy to men at extremely high risk of gun violence involvement (Bhatt et al., 2024). We use the

[32] This applies the cumulative 18-month SDID estimate—a reduction of 39 homicide and shooting victims per 100,000 residents per quarter—to the district's population of roughly 30,600 over six quarters. The corresponding figure for calendar year 2022 alone, used in the benefit-cost calculation below, is approximately 57.

[33] This counts only arrests made directly through GVRS enforcement; it excludes any additional arrests the strategy may have induced indirectly through its broader reshaping of enforcement toward serious violence.

estimates for READI's outreach-referred participants, who were identified much as GVRS subjects are—through human judgment rather than by an algorithm or reentry status—and we draw on the instrumental variables estimates, which recover effects on those who actually took up the program, the analog of the GVRS subjects who engaged with services. Among these participants, the control complier mean for shooting and homicide victimization was 0.132 over READI's 20-month follow-up, which the program is estimated to have reduced by 0.056 per participant, a 43% decline. Scaling this effect to our 18-month window and applying it to the 90 GVRS service recipients would avert on the order of four to five victimizations, about 6% of the roughly 71 victimizations we estimate the strategy prevented. As in the incapacitation calculation, this counts only participants' own averted victimization and so understates the channel: READI also cut these participants' shooting and homicide arrests by 79%, implying averted offending on top of averted victimization. Cutting the other way, READI delivered a more intensive and sustained services package than most GVRS subjects received, so applying its effects likely overstates what GVRS services alone could achieve. We read these considerations as roughly offsetting.

Taken together, the two calculations point the same way: incapacitation and services each plausibly account for a meaningful but minority share of the reduction—incapacitation about a tenth, services somewhat less—and neither approaches the full decline. The sizable residual is consistent with an important role for the deterrence channel that the administrative data cannot capture (Section 6.3) and for the at-scale, network-level effects on which the strategy is premised, reinforcing the reading that no single mechanism explains the reduction on its own.

Beyond the within-Baltimore evidence, our cross-city analysis finds that Baltimore's homicide rate over the post-2022 period was approximately 25% below a synthetic counterfactual constructed from other large U.S. cities, while its non-fatal violent crime rate tracked the counterfactual closely.[34] The citywide effect on homicide is much larger than

[34] The concentration of the cross-city deviation in homicide rather than in non-fatal violent crime is consistent with the design of GVRS. The strategy targets patterned, group-involved violence—the kind that

what would mechanically arise from a Western district-only effect of the magnitude we estimate: the Western district accounted for roughly 16% of Baltimore's homicides in 2021, so a 30% Western-only reduction would translate into only a ∼5% citywide reduction.[35] The discrepancy is consistent with the geographic expansion of GVRS to additional districts after 2022 and with broader changes in how Baltimore polices group-involved violence that the original strategy may have inspired. While the cross-city analysis cannot, on its own, isolate the contribution of GVRS proper from other contemporaneous changes in Baltimore policing or in the city's broader environment, taken together with the within-Baltimore evidence it is consistent with GVRS—and the broader citywide shift toward a focused, partnership-based response to group-involved violence that it may have catalyzed—accounting for a meaningful share of Baltimore's recent historic homicide decline.

The strategy also appears highly cost-effective. Its dedicated hires and contracts were budgeted at roughly $4.8 million for the 2022 pilot year in the Western district. But several service lines were under-expended or never programmed during 2022—including transitional employment, additional incentive funds, and parts of the relocation, microgrant, and stipend budgets—so that actual first-year outlays came in lower, on the order of $3.5 million. The program is small in absolute terms: the actual outlay works out to roughly $6 per Baltimore resident and around one-tenth of one percent of the city's $3.87 billion fiscal 2021 budget. These amounts cover only the program's dedicated spending; they exclude the BPD and State's Attorney personnel reallocated to GVRS from other duties, whose costs were largely absorbed into existing budgets, and they reflect the 2022 pilot year rather than the full 18-month evaluation window. Matching the benefit calculation to this same first year, we

disproportionately results in lethal and near-lethal shootings. The broader non-fatal violent crime category aggregates a much wider set of incidents, most of which fall outside the intended scope of GVRS, and the non-fatal shooting subset is too small on its own to move the aggregate. We would not expect a strategy focused on group-involved gun violence to produce a proportional reduction across the broader non-fatal violent crime category.

[35] For this calculation we scale our combined homicide and non-fatal shooting estimate (−30%), which is far more precisely estimated than the homicide-only estimate. The two are similar in magnitude—the homicide-only ATT is −27%—but the homicide-only estimate is less precise ($p = 0.12$). The choice of outcome for the population share is immaterial: the Western district accounted for 15.6% of citywide homicides and 15.5% of citywide homicides and non-fatal shootings in 2021.

estimate that GVRS averted approximately 57 homicide and non-fatal shooting victimizations in the Western district during 2022. Valuing each at \$2.15 million puts the value of averted gun violence at roughly \$122 million: about 35 times actual first year spending, and at least 25 times even the full budgeted amount.[36] These benefits accrued disproportionately to some of Baltimore's most disadvantaged and socially isolated communities, and are if anything conservative: they count only averted shootings, setting aside the strategy's effects on other crime and the broader social costs of violence.

With respect to the approach's political and operational sustainability, several matters are worth discussing. Focused deterrence strategies can be difficult to implement and, once launched, challenging to sustain (Braga and Kennedy, 2021). Baltimore twice previously discontinued focused deterrence programs that were generating some promising preliminary impacts on serious violence. What distinguishes Baltimore's recent experience from earlier focused deterrence efforts—including Baltimore's own—is not the underlying playbook, which is well established, but the fidelity with which it was implemented. GVRS was not simply a collection of interventions (custom notifications, services, and focused enforcement) but an operating model for organizing violence reduction around a clearly defined population and a disciplined management process: weekly violence reviews grounded in shared analysis, cross-agency coordination, direct communication with high-risk people, services delivered through trusted community organizations, and focused enforcement aligned around a common understanding of who was driving violence and why. Sustained leadership from the Mayor and Police Commissioner helped hold these routines in place, and performance indicators tracking activities and outcomes helped limit the implementation drift that ended Baltimore's earlier attempts and has undermined similar efforts elsewhere.

This bears directly on replicability. Because "focused deterrence" refers to a broad family of programs, what jurisdictions do under that label can differ substantially from day to day, and much of that variation—and, plausibly, much of the variation in measured effects across

[36] Based on the contingent valuation estimate of Ludwig and Cook (2001), \$1.2 million in 1998 dollars, inflated to 2022 dollars to match the budget year.

the prior literature—reflects management and process fidelity rather than the design of the intervention itself. The encouraging implication is that the practices that produced high fidelity implementation in Baltimore—shared problem analysis, recurring violence reviews, cross-agency coordination, performance management, and a disciplined operating rhythm—are ordinary and legible. They do not depend on conditions unique to Baltimore and can, in principle, be adopted elsewhere; what they require is not exceptional local circumstances but the sustained commitment to follow them, which Baltimore's own earlier lapses suggest is achievable but not automatic.

10 References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, 105 (490), 493–505.

——— (2015) "Comparative politics and the synthetic control method," *American Journal of Political Science*, 59 (2), 495–510.

Abadie, Alberto and Javier Gardeazabal (2003) "The economic costs of conflict: A case study of the Basque Country," *American Economic Review*, 93 (1), 113–132.

Aboaba, Oludamilare, Aaron Chalfin, Michael LaForest-Tucker, Lucie Parker, and Patrick Sharkey (2026) "Can Crime Be Deterred at Low Cost? Evidence From a Randomized Experiment in New York," *Journal of Policy Analysis and Management*, 45 (2), e70092.

Agan, Amanda, Jennifer L Doleac, and Anna Harvey (2023) "Misdemeanor prosecution," *The Quarterly Journal of Economics*, 138 (3), 1453–1505.

Ang, Desmond, Panka Bencsik, Jesse Bruhn, and Ellora Derenoncourt (2025) "Community engagement with law enforcement after high-profile acts of police violence," *American Economic Review: Insights*, 7 (1), 124–140.

Ariel, Barak, Ashley Englefield, and John Denley (2019) "I heard it through the grapevine: A randomized controlled trial on the direct and vicarious effects of preventative specific deterrence initiatives in criminal networks," *Journal of Criminal Law & Criminology*, 109, 819.

Arkhangelsky, Dmitry, Susan Athey, David A Hirshberg, Guido W Imbens, and Stefan Wager (2021) "Synthetic difference-in-differences," *The American Economic Review*, 111 (12), 4088–4118.

Bacher-Hicks, Andrew and Elijah de la Campa (2021) "The Impact of New York City's Stop and Frisk Program on Crime: The Case of Police Commanders,"Technical report, Working Paper.

Benjamini, Yoav and Yosef Hochberg (1995) "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal statistical society: series B (Methodological)*, 57 (1), 289–300.

Bhatt, Monica P, Sara B Heller, Max Kapustin, Marianne Bertrand, and Christopher Blattman (2024) "Predicting and preventing gun violence: An experimental evaluation of READI Chicago," *The Quarterly Journal of Economics*, 139 (1), 1–56.

Blattman, Christopher, Sebastian Chaskel, Julian C. Jamison, and Margaret Sheridan (2023) "Cognitive Behavioral Therapy Reduces Crime and Violence over Ten Years: Experimental Evidence," *American Economic Review: Insights*, 5 (4), 527–45, 10.1257/aeri.20220427.

Blattman, Christopher, Donald P Green, Daniel Ortega, and Santiago Tobón (2021) "Place-based interventions at scale: The direct and spillover effects of policing and city services on crime," *Journal of the European Economic Association*, 19 (4), 2022–2051.

Blattman, Christopher, Julian C Jamison, and Margaret Sheridan (2017) "Reducing crime and violence: Experimental evidence from cognitive behavioral therapy in Liberia," *The American Economic Review*, 107 (4), 1165–1206.

Braga, Anthony A and Brenda J Bond (2008) "Policing crime and disorder hot spots: A randomized controlled trial," *Criminology*, 46 (3), 577–607.

Braga, Anthony A and David M Kennedy (2021) *A framework for addressing violence and serious crime: Focused deterrence, legitimacy, and prevention*: Cambridge University Press.

Braga, Anthony A, David M Kennedy, Elin J Waring, and Anne Morrison Piehl (2001) "Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire," *Journal of Research in Crime and Delinquency*, 38 (3), 195–225.

Braga, Anthony A, Brandon S Turchan, Andrew V Papachristos, and David M Hureau (2019a) "Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis," *Journal of Experimental Criminology*, 15 (3), 289–311.

Braga, Anthony A, Brandon S Turchan, and David L Weisburd (2026) "Focused deterrence can reduce crime: A systematic review of randomized controlled trials and quasi-experiments," *Criminology & Public Policy*.

Braga, Anthony A, Greg Zimmerman, Lisa Barao, Chelsea Farrell, Rod K Brunson, and Andrew V Papachristos (2019b) "Street gangs, gun violence, and focused deterrence: Comparing place-based and group-based evaluation methods to estimate direct and spillover deterrent effects," *Journal of Research in Crime and Delinquency*, 56 (4), 524–562.

Bruhn, Jesse (2021) "Competition in the Black Market: Estimating the Causal Effect of Gangs in Chicago."

Chalfin, Aaron (2025) "Investments in policing and community safety," *Annual Review of Criminology*, 8 (1), 403–429.

Chalfin, Aaron, Benjamin Hansen, Emily K Weisburst, and Morgan C Williams (2022) "Police Force Size and Civilian Race," *American Economic Review: Insights*.

Chalfin, Aaron, Michael LaForest, and Jacob Kaplan (2021) "Can Precision Policing Reduce Gun Violence? Evidence from "Gang Takedowns" in New York City," *Journal of Policy Analysis and Management*.

Chalfin, Aaron and Justin McCrary (2017) "Criminal deterrence: A review of the literature," *Journal of Economic Literature*, 55 (1), 5–48.

——— (2018) "Are US cities underpoliced? Theory and evidence," *The Review of Economics and Statistics*, 100 (1), 167–186.

Cho, Sungwoo, Felipe Gonçalves, and Emily Weisburst (2024) "The impact of fear on police behavior and public safety," *Review of Economics and Statistics*, 1–45.

Choe, Hyeonggeun, Alyssa Mendlein, and Aaron Chalfin (2026) "Gang Takedowns and Neighborhood Violence: Evidence from a Case Study in Philadelphia," *Working Paper*.

Ciccia, Diego (2024) "A short note on event-study synthetic difference-in-differences estimators," *arXiv preprint arXiv:2407.09565*.

Cook, Philip J. (1980) "Research in criminal deterrence: Laying the groundwork for the second decade," *Crime and Justice*, 2, 211–268.

Cook, Philip J. and Jens Ludwig (2000) *Gun violence: The real costs*: Oxford University Press, USA.

Cook, Philip J., Ariadne E. Rivera-Aguirre, Magdalena Cerdá, and Garen Wintemute (2017) "Constant lethality of gunshot injuries from firearm assault: United States, 2003-2012," *American Journal of Public Health*, 107 (8), 1324–1328, 10.2105/AJPH.2017.303837.

Corsaro, Nicholas and Robin S Engel (2015) "Most challenging of contexts: Assessing the impact of focused deterrence on serious violence in New Orleans," *Criminology & Public Policy*, 14 (3), 471–505.

Cullen, Julie and Steven Levitt (1999) “Crime, Urban Flight, And The Consequences For Cities,” *The Review of Economics and Statistics*, 81 (2), 159–169.

Davis, Jonathan M V, Tracey Meares, and Emily Arnesen (2025) “Improving programming in juvenile detention: The impact of project safe neighborhoods youth outreach forums,” *Journal of Quantitative Criminology*, 41 (1), 23–50.

Del Pozo, Brandon, Alex Knorre, Michael J Mello, and Aaron Chalfin (2022) “Comparing risks of firearm-related death and injury among young adult males in selected US cities with wartime service in Iraq and Afghanistan,” *JAMA Network Open*, 5 (12), e2248132–e2248132.

Denley, John (2023) *Prevention not cure: Targeting influenced, influencers and convergence spots to stop organised crime from happening* Ph.D. dissertation, University of Cambridge, Institute of Criminology.

Dobbie, Will, Jacob Goldin, and Crystal S Yang (2018) “The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges,” *The American Economic Review*, 108 (2), 201–40.

Domínguez, Magdalena (2026) “Sweeping up gangs: The effects of tough-on-crime policies from a network approach,” *American Economic Journal: Economic Policy*, 18 (1), 127–158.

Durlauf, Steven N and Daniel S Nagin (2011) “Imprisonment and crime: Can both be reduced?” *Criminology & Public Policy*, 10 (1), 13–54.

Evans, William N and Emily G Owens (2007) “COPS and Crime,” *Journal of Public Economics*, 91 (1-2), 181–201.

Fader, Jamie J (2021) ““I don’t have time for drama”: Managing risk and uncertainty through network avoidance,” *Criminology*, 59 (2), 291–317.

Fagan, Jeffrey and Garth Davies (2000) “Street stops and broken windows: Terry, race, and disorder in New York City,” *Fordham Urban Law Journal*, 28, 457.

Gramlich, John (2023) *What the data says about gun deaths in the US*: Pew Research Center.

Green, Ben, Thibaut Horel, and Andrew V Papachristos (2017) “Modeling contagion through social networks to explain and predict gunshot violence in Chicago, 2006 to 2014,” *JAMA Internal Medicine*, 177 (3), 326–333.

Harcourt, Bernard E and Jens Ludwig (2006) “Broken windows: New evidence from New York City and a five-city social experiment,” *The University of Chicago Law Review*, 271–320.

Heckman, James J, Hidehiko Ichimura, and Petra Todd (1998) “Matching as an econometric evaluation estimator,” *The review of economic studies*, 65 (2), 261–294.

Heller, Sara B, Benjamin Jakubowski, Zubin Jelveh, and Max Kapustin (2024) “Machine learning can predict shooting victimization well enough to help prevent it,” *Review of Economics and Statistics*, 1–45.

Heller, Sara B, Anuj K Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A Pollack (2017) “Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago,” *The Quarterly Journal of Economics*, 132 (1), 1–54.

Jabri, Ranae (2021) “Algorithmic Policing.”

Jacobs, Bruce A and Michael Cherbonneau (2023) “Carjacking: Scope, Structure, Process, and Prevention,” *Annual Review of Criminology*, 6, 155–179.

Kapustin, Max, Terrence Neumann, and Jens Ludwig (2022) "Policing and management,"Technical report, National Bureau of Economic Research.

Kennedy, David M (1997) *Juvenile Gun Violence and Gun Markets in Boston: A Summary of a Research Presesentation*: US Department of Justice, Office of Justice Programs, National Institute of . . . .

——— (2011) "Whither streetwork: The place of outreach workers in community violence prevention," *Criminology & Public Policy*, 10, 1045.

Kennedy, David M, Anthony A Braga, Anne M Piehl, and Elin J Waring (2001) *Reducing gun violence: the boston gun project's operation ceasefire*: US Department of Justice Office of Justice Programs.

Kennedy, David M, Anne M Piehl, and Anthony A Braga (1996) "Youth violence in Boston: Gun markets, serious youth offenders, and a use-reduction strategy," *Law & Contemporary Problems*, 59, 147.

Klößner, Stefan, Ashok Kaul, Gregor Pfeifer, and Manuel Schieler (2018) "Comparative politics and the synthetic control method revisited: A note on Abadie et al.(2015)," *Swiss Journal of Economics and Statistics*, 154, 1–11.

Leovy, Jill (2015) *Ghettoside: A true story of murder in America*: Spiegel & Grau.

Leslie, Emily and Nolan G Pope (2017) "The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments," *The Journal of Law and Economics*, 60 (3), 529–557.

Ludwig, Jens and Philip J Cook (2001) "The benefits of reducing gun violence: Evidence from contingent-valuation survey data," *Journal of Risk and Uncertainty*, 22 (3), 207–226.

MacDonald, John (2023) "Criminal justice reform guided by evidence: social control works—The Academy of Experimental Criminology 2022 Joan McCord Lecture," *Journal of Experimental Criminology*, 1–18.

Mas, Alexandre (2006) "Pay, reference points, and police performance," *The Quarterly Journal of Economics*, 121 (3), 783–821.

McCrary, Justin and Deepak Premkumar (2019) "Why we need police," in *The Cambridge Handbook of Policing in the United States*, 65–84.

Mello, Steven (2019) "More COPS, less crime," *Journal of Public Economics*, 172, 174–200.

Nagin, Daniel S (2013) "Deterrence in the twenty-first century," *Crime and Justice*, 42 (1), 199–263.

Pailañir, Daniel and Damian Clarke (2023) "SDID: Stata module to perform synthetic difference-in-differences estimation, inference, and visualization."

Papachristos, Andrew V, Anthony A Braga, and David M Hureau (2012) "Social networks and the risk of gunshot injury," *Journal of Urban Health*, 89 (6), 992–1003.

Papachristos, Andrew V, David M Hureau, and Anthony A Braga (2013) "The corner and the crew: The influence of geography and social networks on gang violence," *American Sociological Review*, 78 (3), 417–447.

Papachristos, Andrew V, Tracey L Meares, and Jeffrey Fagan (2007) "Attention felons: Evaluating project safe neighborhoods in Chicago," *Journal of Empirical Legal Studies*, 4 (2), 223–272.

Ratcliffe, Jerry H (2004) "The hotspot matrix: A framework for the spatio-temporal targeting of crime reduction," *Police Practice and Research*, 5 (1), 5–23.

Ready, Justin, Devon Cowan, and Barak Ariel (2023) "Reducing serious public violence: Testing the effects of focused deterrence to reduce serious public violence,"Technical report, Victoria Police Southern Metro Region.

Robertson, Campbell (2022) "'I Honestly Believe It's a Game': Why Carjacking Is on the Rise Among Teens.," *International New York Times*, NA–NA.

Roman, Caterina G, Nathan W Link, Jordan M Hyatt, Avinash Bhati, and Megan Forney (2019) "Assessing the gang-level and community-level effects of the Philadelphia focused deterrence strategy," *Journal of Experimental Criminology*, 15, 499–527.

Saldaña, Johnny (2021) "The coding manual for qualitative researchers."

Saunders, Jessica, Russell Lundberg, Anthony A Braga, Greg Ridgeway, and Jeremy Miles (2015) "A synthetic control approach to evaluating place-based crime interventions," *Journal of Quantitative Criminology*, 31, 413–434.

Sharkey, Patrick and Robert J Sampson (2010) "Destination effects: Residential mobility and trajectories of adolescent violence in a stratified metropolis," *Criminology*, 48 (3), 639–681.

Smith, Jeffrey A and Petra E Todd (2005) "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of econometrics*, 125 (1-2), 305–353.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick (2019) "Capitalists in the Twenty-First Century," *The Quarterly Journal of Economics*, 134 (4), 1675–1745, 10.1093/qje/qjz020.

Tita, George and Allan Abrahamse (2004) "Gang homicide in LA, 1981-2001," *At the local level: Perspectives on violence prevention*, 3, 1–20.

Trinkner, Rick (2019) "Addressing the "black box" of focused deterrence: An examination of the mechanisms of change in Chicago's Project Safe Neighborhoods," *Journal of Experimental Criminology*, 15, 673–683.

Tyler, Tom R, Jonathan Jackson, and Avital Mentovich (2015) "The consequences of being an object of suspicion: Potential pitfalls of proactive police contact," *Journal of Empirical Legal Studies*, 12 (4), 602–636.

Vitale, Alex S (2021) *The end of policing*: Verso Books.

Weisburd, David, Laura A Wyckoff, Justin Ready, John E Eck, Joshua C Hinkle, and Frank Gajewski (2006) "Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits," *Criminology*, 44 (3), 549–592.

Weisburst, Emily K (2019a) "Police use of force as an extension of arrests: Examining disparities across civilian and officer race," in *The American Economic Review*, 109, 152–56.

——— (2019b) "Safety in police numbers: Evidence of police effectiveness from federal COPS grant applications," *American Law and Economics Review*, 21 (1), 81–109.

Zimring, Franklin E, Gordon Hawkins, and James Vorenberg (1973) "Deterrence: The legal threat in crime control."

11 Tables and Figures

Table 1: Sociodemographic characteristics of Baltimore and the Western district

	Baltimore	Western district
Population	592,211	31,017
Median age	37.0	40.6
% Black	60.9%	93.2%
Median household income	60, 891	30, 938
% below poverty	20.3%	35.9%
% Bachelor's degree or higher	34.2%	9.2%
% vacant units	16.6%	36.0%

Note: American Community Survey 5-year estimates, 2017–2021. Western district figures are aggregated from Census tracts using area-weighted apportionment.

Table 2: GVRS subjects: baseline characteristics

	All referrals	Communication referrals	Arrest targets
N	276	179	97
Demographics			
Age at referral	31.1	30.9	31.4
Black	1.00	0.99	1.00
Male	0.94	0.94	0.94
Share with a prior non-fatal shooting victimization			
Ever before referral	0.24	0.30	0.12
In year before referral	0.14	0.20	0.04
Share with a prior arrest			
Ever before referral	0.81	0.85	0.73
In year before referral	0.26	0.28	0.21
Ever before referral, Western district	0.57	0.64	0.44
In year before referral, Western district	0.13	0.16	0.07
Among those with a prior arrest, average number of:			
Total arrests	6.3	6.5	5.9
Arrests with the following charge:			
Homicide	0.1	0.1	0.1
Carjacking	0.0	0.0	0.1
Aggravated assault	0.2	0.2	0.2
Robbery	0.2	0.2	0.1
Simple assault	0.8	0.9	0.6
Drug possession	2.2	2.3	2.0
Drug distribution	1.7	1.7	1.6
Disorderly conduct	0.5	0.4	0.5
Weapons-related	0.8	0.9	0.6
Other	1.5	1.4	1.6

Note: The sample comprises GVRS subjects whose first referral was in the Western district and occurred before July 2023, classified by the type of their *first* referral (communication referral or arrest target); the two categories are mutually exclusive at the person level. Race is set to missing for subjects with no recorded race in the arrest data; sex is set to missing for the small number of subjects whose recorded sex is unknown. Prior non-fatal shooting victimizations and prior arrests are measured over the period preceding each subject's first referral; "in year before referral" refers to the 365 days before the referral date. The bottom panel reports average arrest counts conditional on the subject having any prior arrest record in the panel.

Table 3: GVRS subjects: activities

	All referrals	Communication referrals	Arrest targets
N	276	179	97
GVRS subjects receiving a direct communication			
Any	192	152	40
Call-in	8	1	7
Custom notification	184	151	33
Full	96	93	3
Law enforcement only	90	60	30
GVRS subjects receiving services			
Any	90	81	9
Life coaching, CBT, and other counseling	86	78	8
Employment and licensing help	30	27	3
Education and vocational training	24	22	2
Stipends and emergency relocation help	13	13	0
GVRS subjects arrested since first referral			
Any arrest	150	55	95
Any GVRS-related arrest	110	15	95
Takedown	76	9	67
Other GVRS-related arrest	35	6	29

Note: The sample comprises GVRS subjects whose first referral was in the Western district and occurred before July 2023. Subjects are classified by the type of their *first* referral (communication referral or arrest target). Each cell reports the number of subjects who received the indicated activity at any point through June 2023; subjects may receive multiple activities and so appear in multiple rows. Direct communications include both call-ins and custom (in-person) notifications; custom notifications are further split between full notifications (including service providers) and law-enforcement-only notifications. CBT refers to cognitive behavioral therapy. "Takedown" arrests refer to coordinated multi-subject enforcement actions (see Appendix Table A.1); "other GVRS-related arrests" refer to single-subject targeted arrests through the GVRS pipeline.

Table 4: Estimated effects of GVRS on reported violent crime rates in the Western district

	Homicide/shooting victims	Carjackings	Aggravated assaults	Robberies
Counterfactual post-period mean	130.8	39.7	366.3	160.4
ATT estimate	-39.0***	-15.3*	28.9	-8.7
	(15.0)	(7.9)	(23.6)	(18.1)
Percent change (vs. counterfactual)	-29.8%	-38.5%	7.9%	-5.5%
Bootstrap p-value	0.009	0.054	0.222	0.629
FDR q-value	0.038	0.109	0.296	0.629

Note: Each column reports synthetic difference-in-differences (SDID) estimates of the average treatment effect on the treated (ATT) of GVRS in the Western district, over the 18-month post-intervention window from January 2022 through June 2023, for the indicated violent crime outcome. Outcomes are reported crime rates per 100,000 residents at the post-quarter level. The counterfactual post-period mean is the implied mean of the outcome in the Western district absent GVRS (observed post-period mean less the ATT); the percent change is the ATT divided by this counterfactual mean. Standard errors (in parentheses) and bootstrap p-values are from 500 clustered bootstrap iterations. FDR q-values are Benjamini-Hochberg-adjusted across the four outcomes. The pre-treatment period is January 2015 through December 2021 (28 quarters). The donor pool is all posts in Baltimore excluding the Western and Southwestern districts; the latter became GVRS-treated in January 2023 (N=97 donor posts). $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 5: Estimated effects of GVRS on arrests and clearance rates in the Western district

Panel A: Estimated effects on broad enforcement	Total arrests	Non-violent index crime arrests	Drug arrests (excl. violent index)
Counterfactual post-period mean	548.7	493.1	182.5
ATT estimate	-11.0	-24.1	-28.5
	(48.0)	(46.9)	(34.2)
Percent change (vs. counterfactual)	-2.0%	-4.9%	-15.6%
Bootstrap p-value	0.818	0.608	0.404
Panel B: Estimated effects on targeted enforcement	Violent index crime arrests	Aggravated assault arrests	Violent index crime clearance rate
Counterfactual post-period mean	37.9	40.6	0.076
ATT estimate	30.7***	19.1**	0.028*
	(9.5)	(8.2)	(0.016)
Percent change (vs. counterfactual)	80.8%	47.1%	36.8%
Bootstrap p-value	0.001	0.019	0.079

Note: Each column reports SDID estimates of the average treatment effect on the treated (ATT) of GVRS in the Western district, over the 18-month post-intervention window from January 2022 through June 2023, for the indicated arrest or clearance-rate outcome. Panel A groups three measures of broad enforcement activity: total arrests, non-violent index crime arrests (any arrest whose charges do not include a violent index crime), and drug arrests not co-charged with a violent index crime. Panel B groups three measures of enforcement aimed at firearm-involved violence: violent index crime arrests (any arrest whose charges include a violent index crime), aggravated assault arrests, and the violent index crime quasi-clearance rate (the ratio of violent index crime arrests to reported violent index crimes; "quasi" because individual arrests cannot be matched to the specific reported crimes they resolve). All arrest outcomes are rates per 100,000 residents at the post-quarter level; the quasi-clearance rate is a share. Standard errors (in parentheses) and bootstrap p-values are from 500 clustered bootstrap iterations. The pre-treatment period is January 2015 through December 2021 (28 quarters). The donor pool is the same as in Table 4. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 6: Estimated effects of GVRS on ShotSpotter incidents, 911 calls, and willingness to report in the Western district

	ShotSpotter incidents	ShotSpotter rounds	911 calls	Calls-to-incident ratio
Counterfactual post-period mean	687.8	2064.3	28280.3	50.66
ATT estimate	-59.2	-78.1	-2572.3*	0.28
	(43.9)	(197.8)	(1347.7)	(6.98)
Percent change (vs. counterfactual)	-8.6%	-3.8%	-9.1%	0.5%
Bootstrap p-value	0.177	0.693	0.056	0.969

Note: Each column reports an SDID estimate of the average treatment effect on the treated (ATT) of GVRS in the Western district, over the 18-month post-treatment window from January 2022 through June 2023, for the indicated ShotSpotter or 911-call outcome. The first three outcomes are rates per 100,000 residents at the post-quarter level; the final column is the unitless ratio of 911 calls to ShotSpotter incidents in the post-quarter, following Ang et al. (2025). An increase in this ratio corresponds to more public reporting (more 911 calls) per acoustically-detected gunfire incident. Standard errors (in parentheses) and bootstrap p-values are from 500 clustered bootstrap iterations. The pre-treatment period is January 2019 through December 2021 (12 quarters), set by the availability of ShotSpotter coverage in Baltimore. The donor pool is restricted to posts with at least one ShotSpotter incident in every quarter from January 2019 through June 2023, a proxy for being continuously inside the ShotSpotter sensor footprint, and additionally excludes the Southwestern district which became GVRS-treated in January 2023 (N=25 donor posts). $^{*}p$<0.10, $^{**}p$<0.05, $^{***}p$<0.01.

Table 7: Baseline characteristics of GVRS subjects and their matched controls

	Communication referrals: recent shooting victims		Communication referrals: not recent shooting victims		Arrest targets	
	GVRS	Controls	GVRS	Controls	GVRS	Controls
N	13	79	90	6,302	60	4,766
Demographics						
Age at referral	29.1	29.9	30.3	32.3	30.5	31.9
Black	1.00	0.98	0.99	0.91	0.98	0.90
Male	0.85	0.88	0.92	0.86	0.93	0.89
In year before referral						
Any non-fatal shooting victimization	1.00	1.00	0.00	0.01	0.05	0.02
Any arrest	0.46	0.20	0.26	0.19	0.25	0.20
Part 1 violent crime arrests	0.00	0.05	0.04	0.05	0.02	0.07
Part 1 non-violent-crime arrests	0.00	0.02	0.03	0.03	0.03	0.04
Simple assault arrests	0.15	0.02	0.08	0.06	0.03	0.06
Weapons-related arrests	0.00	0.11	0.18	0.06	0.12	0.07
Drug possession arrests	0.00	0.01	0.00	0.00	0.00	0.00
Drug distribution arrests	0.15	0.04	0.04	0.04	0.13	0.04
Disorderly arrests	0.00	0.00	0.02	0.00	0.00	0.00
Other arrests	0.38	0.01	0.18	0.07	0.12	0.07
More than 1 year before referral						
Any non-fatal shooting victimization	0.00	0.05	0.13	0.10	0.10	0.10
Any arrest	0.62	0.72	0.81	0.81	0.72	0.71
Part 1 violent crime arrests	0.54	0.11	0.47	0.46	0.42	0.39
Part 1 non-violent-crime arrests	0.46	0.23	0.47	0.39	0.35	0.34
Simple assault arrests	0.23	0.19	0.67	0.61	0.47	0.47
Weapons-related arrests	0.38	0.31	0.63	0.56	0.35	0.40
Drug possession arrests	0.85	1.00	1.62	1.31	1.43	1.20
Drug distribution arrests	0.85	0.66	1.17	1.01	1.02	0.86
Disorderly arrests	0.08	0.12	0.43	0.25	0.45	0.28
Other arrests	0.92	0.58	1.11	0.86	1.03	0.82

Note: The sample is GVRS subjects whose first referral was in the Western district and occurred in 2022 and who were matched to at least one control, plus their propensity score-matched controls. Subjects are stratified by first-referral type and recent non-fatal shooting (NFS) victimization: communication referrals with at least one NFS victimization in the 90 days before referral; communication referrals without any NFS victimization in that window; and arrest targets. Matching procedure described in Section 6.2. The Controls columns report means over (treated, matched-control) pairs, so a control matched to multiple treated subjects contributes once per match. The N row reports the count of unique subjects in each cell. "In year before referral" aggregates the 1–90 and 91–365 day pre-referral windows used in matching; "more than 1 year before referral" aggregates the 1–2 year, 2–3 year, 3–5 year, and 5+ year pre-referral windows. Race is set to missing for subjects with no recorded race in the arrest data; sex is set to missing for the small number of subjects whose recorded sex is unknown.

Table 8: Estimated effects of GVRS referrals on person-level arrest and victimization

	Communication referrals	Arrest targets
Panel A: Any arrest		
Quarter of referral		
Counterfactual mean	0.066	0.072
ATT estimate	0.013	0.738***
	(0.041)	(0.062)
Quarter after referral		
Counterfactual mean	0.070	0.075
ATT estimate	-0.010	0.118**
	(0.048)	(0.050)
N (matched treated)	103	60
Panel B: Homicide or non-fatal shooting victimization		
Quarter of referral		
Counterfactual mean	0.003	0.005
ATT estimate	0.016	-0.018
	(0.014)	(0.017)
Quarter after referral		
Counterfactual mean	0.004	0.006
ATT estimate	0.026	-0.002
	(0.017)	(0.024)
N (matched treated)	103	60

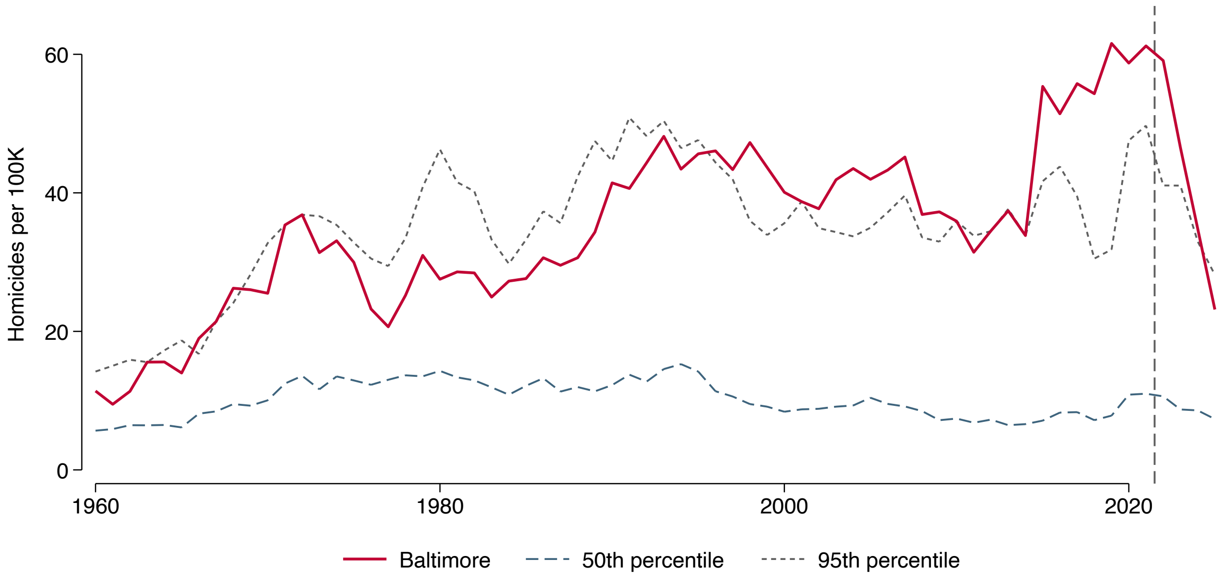
Note: Each cell reports the average treatment effect on the treated (ATT) on the indicated outcome for the indicated GVRS referral type, in the quarter of referral and the subsequent quarter. Estimated using propensity score-matched difference-in-differences with up to 100 nearest-neighbor matches per treated subject (with replacement) and a caliper of 0.01 on the linear-index propensity score. The propensity score is fit on the matching sample using pre-referral covariates from five or more quarters before referral, holding out the four quarters immediately preceding referral as a placebo window for testing of conditional parallel trends. The counterfactual mean is the matched-control mean of the outcome at the indicated post-treatment quarter, treated-id-weighted. Heteroskedasticity-robust standard errors (in parentheses) are clustered at the GVRS subject. Results are robust to a non-holdout matching specification that uses pre-referral covariates from all available quarters. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 9: Synthetic difference-in-differences estimates of Baltimore homicide and violent crime rates relative to donor pool counterfactuals

	Homicide rate	Violent crime rate (excl. homicides)
Panel 1: Donor pool: above-median pre-2022 rate		
Counterfactual post-period mean	54.6	1837.5
ATT estimate	-13.6***	-30.3
	(3.5)	(126.5)
Percent change (vs. counterfactual)	-24.9%	-1.7%
Placebo p-value	0.000	0.810
N donor units	125	119
Panel 2: Donor pool: population $\geq$ 250K		
Counterfactual post-period mean	54.5	1865.7
ATT estimate	-13.5***	-58.5
	(2.4)	(96.5)
Percent change (vs. counterfactual)	-24.7%	-3.1%
Placebo p-value	0.000	0.544
N donor units	83	77

Note: Each column reports SDID estimates of the average treatment effect on the treated (ATT) for the indicated quarterly outcome, expressed per 100,000 population and annualized (quarterly count $\div$ population $\times$ 100,000 $\times$ 4). Baltimore is treated beginning January 2022. The unit of observation is an agency-quarter from January 2018 through December 2025. Outcomes are constructed from monthly counts in the Real-Time Crime Index. The estimation sample is restricted to municipal police departments serving populations of at least 100,000, with single-month gaps in any component series filled by linear interpolation; agencies with longer gaps in the relevant components are dropped from the donor pool. Each panel uses a different donor pool. The counterfactual post-period mean equals Baltimore's observed post-treatment mean less the ATT; the percent change is the ATT divided by this counterfactual mean. Placebo p-values are computed by re-estimating the model treating each donor unit as the treated unit in turn (500 placebo iterations). $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Figure 1: Baltimore's homicide rate vs. peer-city distribution, 1960–2025

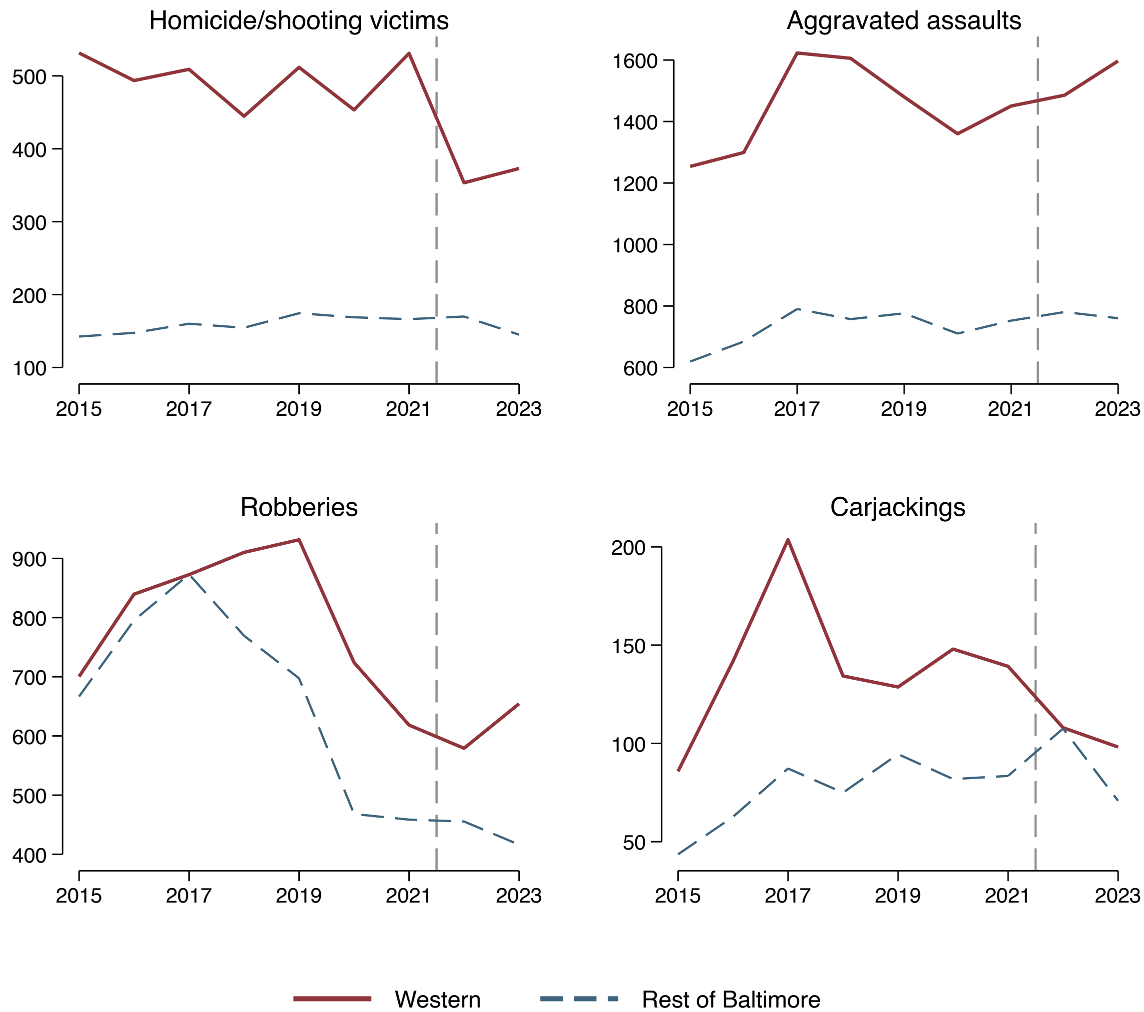


Note: Figure plots Baltimore's annual homicide rate per 100,000 residents (solid line) against the 50th and 95th percentiles of the same outcome across U.S. cities with populations of at least 250,000 (dashed lines), from 1960 through 2025. The vertical dashed line marks the 2021/2022 boundary, immediately preceding the launch of GVRS in the Western district in January 2022. Sources: FBI Uniform Crime Reports (1960–2017) and the Real-Time Crime Index (2018–2025).

Figure 2: Homicides rates across police districts in major U.S. cities

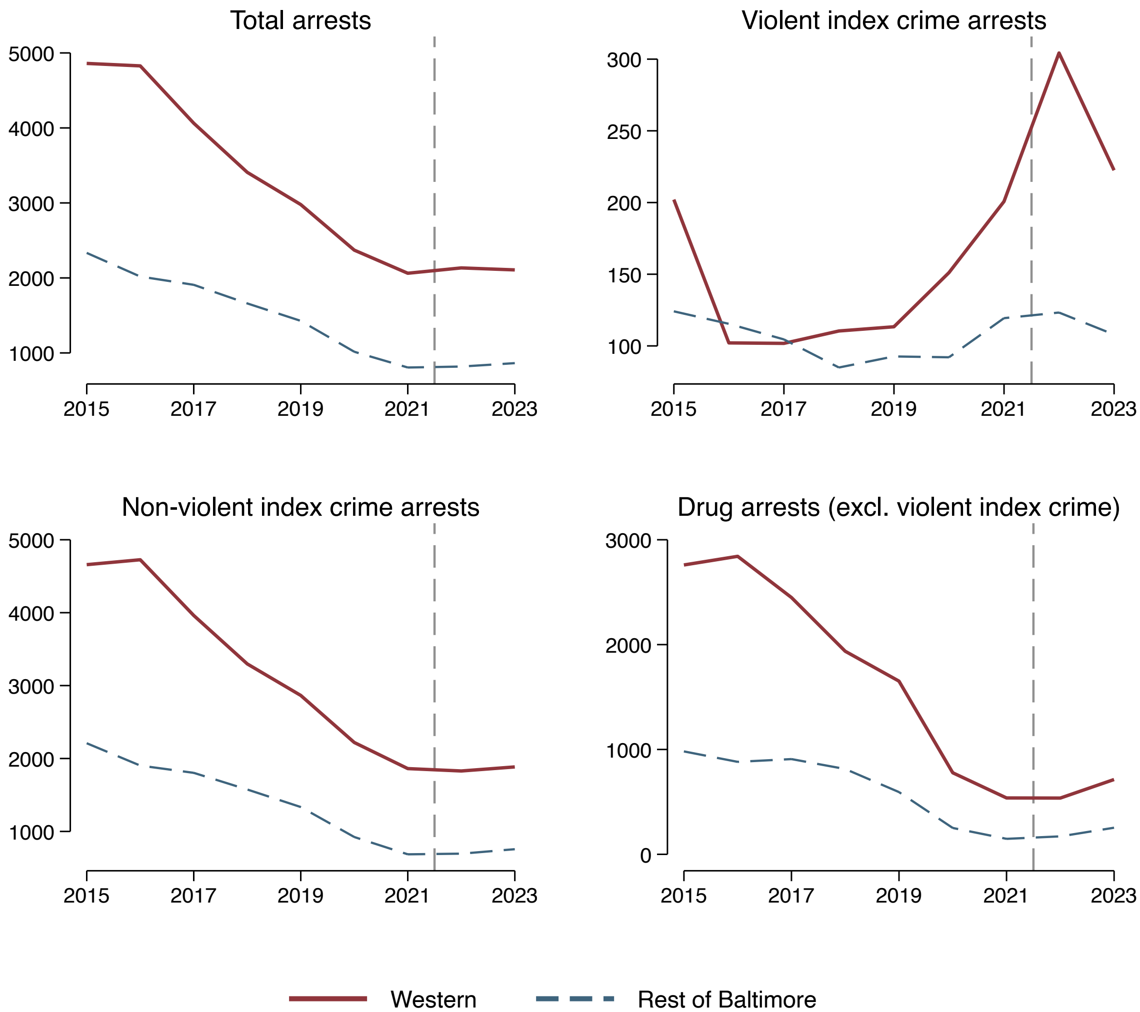
Note: Figure shows annual homicide rates per capita for all 323 police districts in a set of 27 major U.S. cities, ranked from lowest to highest homicide rate. Homicide rates are calculated based on the median number of annual homicides in each district over a 3- to 5-year period ranging from 2017 to 2022, depending on the publicly available data for each city. The Western district in Baltimore ranks 5 out of 323, with only two districts in St. Louis, one in Louisville, and one in New Orleans having a higher homicide rate. The Southwestern and Eastern districts in Baltimore rank 7 and 10, respectively. The set of cities includes Atlanta, Baltimore, Boston, Charleston, Charlotte, Chicago, Cleveland, Dallas, Detroit, Houston, Indianapolis, Kansas City (Missouri), Los Angeles, Louisville, Memphis, Miami-Dade County, Milwaukee, Minneapolis, Nashville, New Orleans, New York City, Philadelphia, Phoenix, San Francisco, Seattle, St. Louis, and Washington, D.C.

Figure 3: Reported crimes per 100K in the Western district and the rest of Baltimore



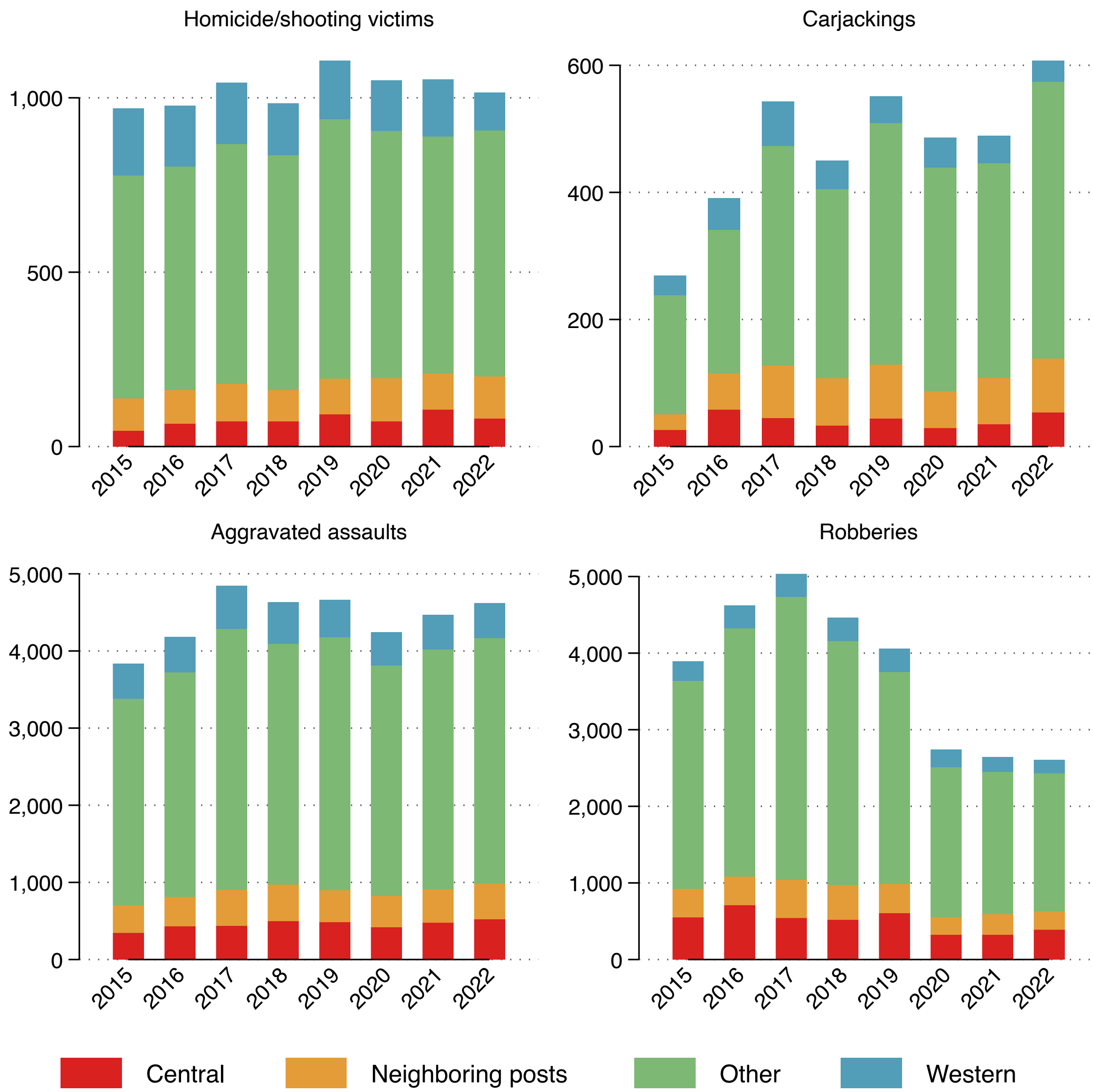
Note: Figure plots annual rates per 100,000 residents for the Western district and the rest of Baltimore, for four reported crimes that commonly involve the use or threat of firearms against a person: homicide and non-fatal shooting victims, aggravated assaults, robberies, and carjackings. Source: Baltimore Police Department records. Rates shown for 2023 are calculated by doubling counts from January 1 through June 30. The vertical line marks the launch of GVRS in the Western district in January 2022.

Figure 4: Arrests per 100K in the Western district and the rest of Baltimore



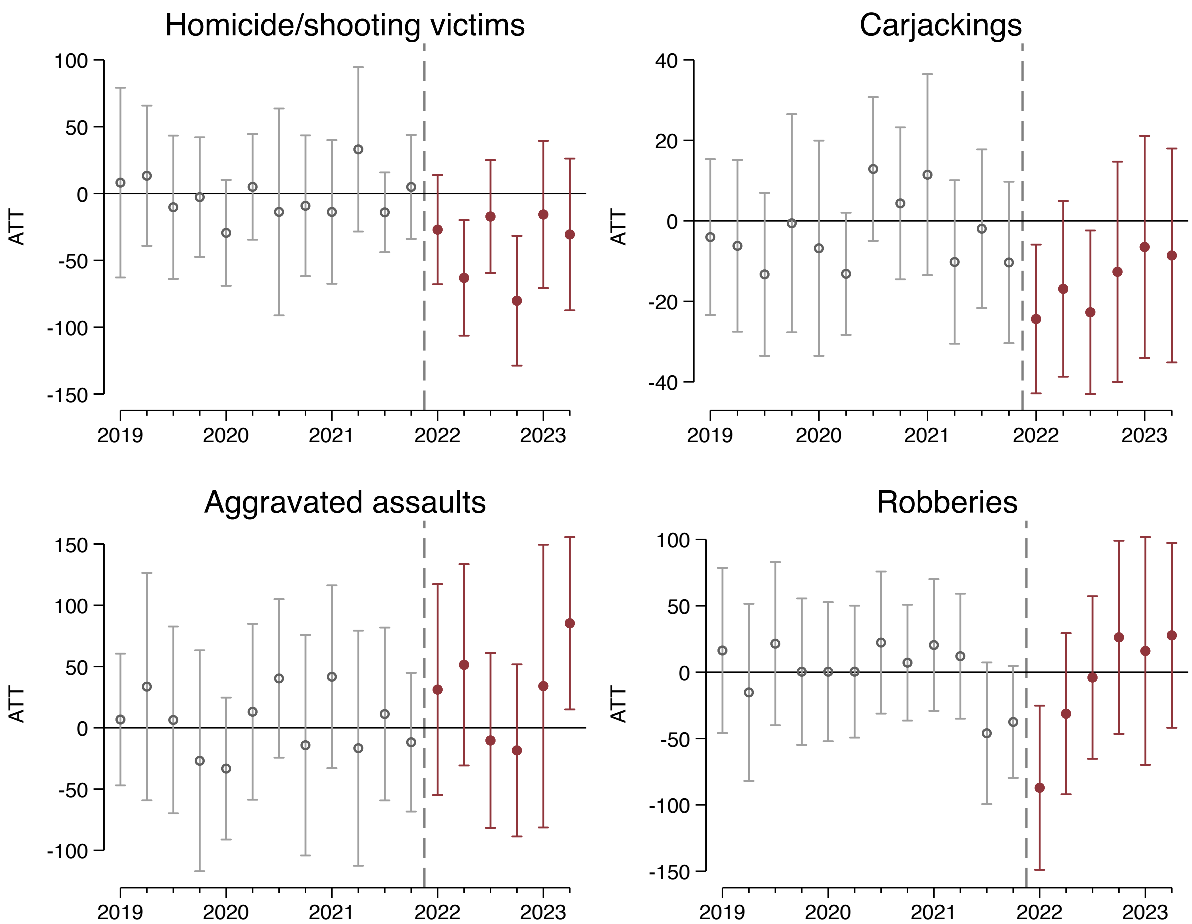
Note: Figure plots annual arrest rates per 100,000 residents for the Western district and the rest of Baltimore, for four arrest measures: total arrests, violent index crime arrests, non-violent index crime arrests, and drug arrests not co-charged with a violent index crime (see Section 5.1 for definitions). Source: Baltimore Police Department records. Rates shown for 2023 are calculated by doubling counts from January 1 through June 30. The vertical line marks the launch of GVRS in the Western district in January 2022.

Figure 5: Reported crimes by area of Baltimore, 2015–2022



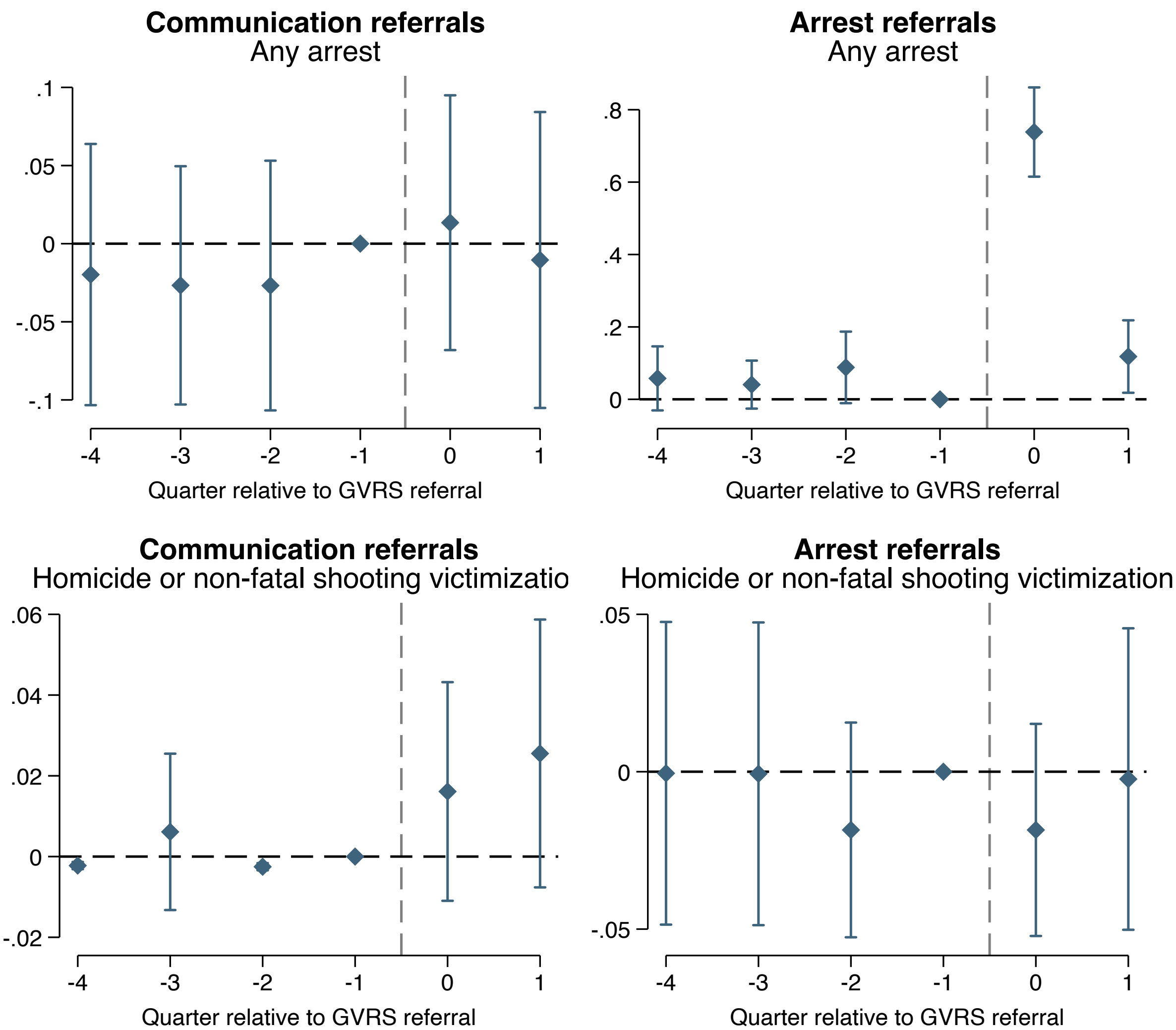
Note: Figure plots annual counts of homicide and non-fatal shooting victims, carjackings, aggravated assaults, and robberies in different areas of Baltimore. The Central district and posts neighboring the Western and Southwestern districts are the areas most likely to experience spatial spillovers from the launch of GVRS in January 2022.

Figure 6: Estimated effects of GVRS on reported crime rates in the Western district



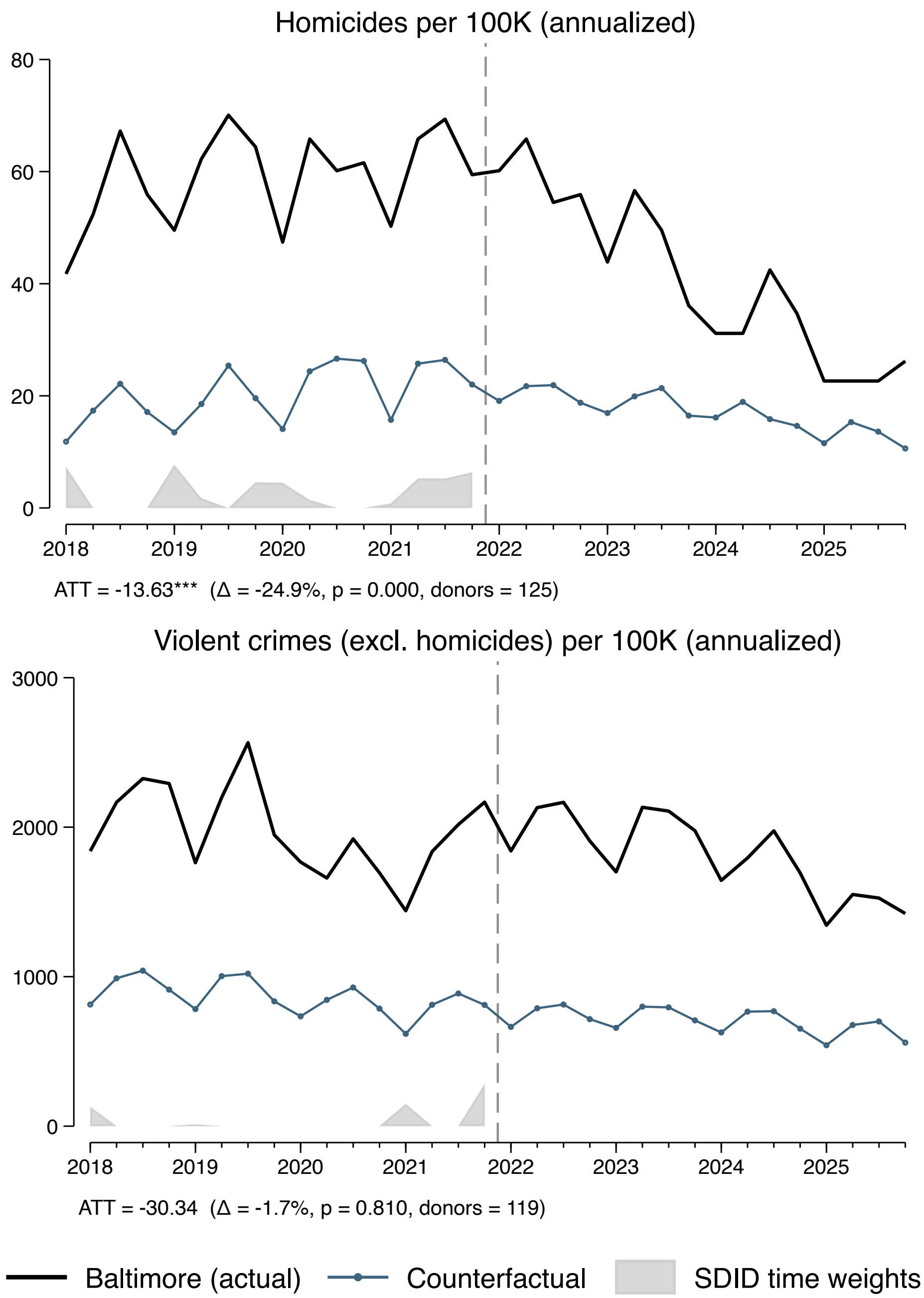
Note: Figure plots quarterly SDID coefficients and 95% confidence intervals for rates of homicide and non-fatal shooting victims, carjackings, aggravated assaults, and robberies in the Western district. Outcomes are reported crime rates per 100,000 residents at the post-quarter level. Estimated using Ciccia's `sdid_event` package; see Section 5.2 for the construction of the post-treatment estimates and pre-treatment in-time placebos. The dashed vertical line marks the pre/post boundary between 2021Q4 and 2022Q1 (the launch of GVRS). Maroon markers (2022Q1 through 2023Q2) are quarterly ATT estimates over the 18-month post window; gray markers (2019Q1 through 2021Q4) are in-time placebo estimates over the twelve quarters preceding GVRS launch. Major x-axis ticks mark each year's Q1; minor ticks mark Q2–Q4. The aggregate 18-month ATTs reported in Table 4 are not plotted here. The donor pool is all Baltimore posts outside the Western and Southwestern districts. Confidence intervals are constructed from 500 bootstrap iterations.

Figure 7: Estimated effects of GVRS referrals on person-level arrest and victimization



Note: Each panel reports event study coefficient estimates and 95% confidence intervals from a propensity score-matched difference-in-differences for GVRS subjects whose first referral was in the Western district and occurred in 2022. Subjects are pooled across communication referrals (top row) and arrest targets (bottom row); within communication referrals, subjects are further stratified by whether they had a non-fatal shooting victimization in the 90 days before referral, with matching done separately within strata. Outcomes are an indicator for any arrest (using a harmonized measure that combines BPD arrest records with GVRS-recorded arrest events not appearing in BPD's data; see Section 6.1) and an indicator for any homicide or non-fatal shooting victimization in the quarter. Quarter -1 is the omitted reference. The dashed vertical line at -0.5 marks the pre/post boundary. Heteroskedasticity-robust standard errors are clustered at the GVRS subject. Pre-period victimizations are mechanically non-fatal, since subjects who experienced a homicide before the matched pair's reference date are excluded from the matching pool; post-period values include both fatal and non-fatal events.

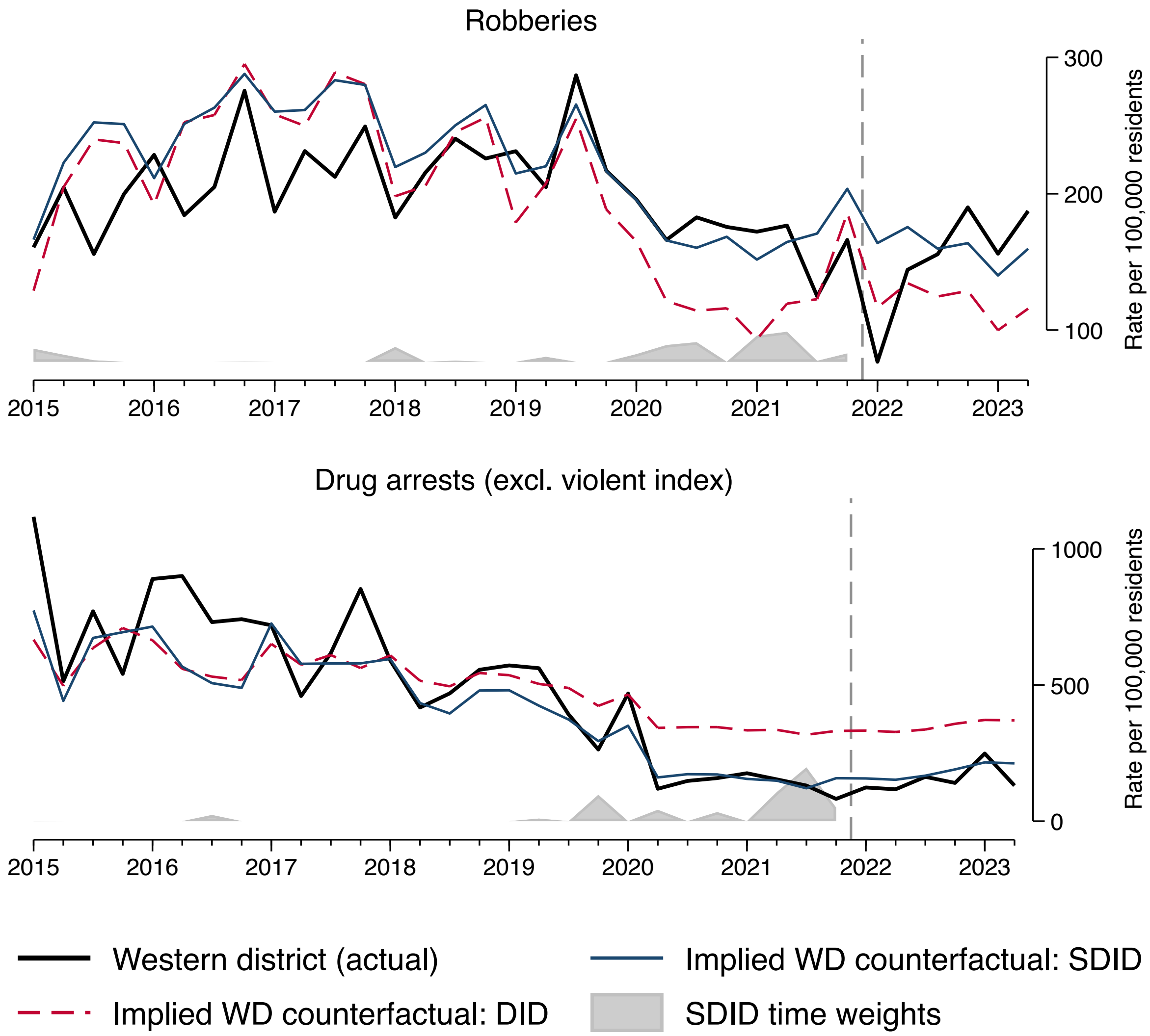
Figure 8: Synthetic difference-in-differences estimates of Baltimore homicide and violent crime rates, above-median donor pool



Note: Top panel: homicide rate. Bottom panel: violent crime rate excluding homicides. Each panel plots Baltimore's observed quarterly outcome (per 100,000 population, annualized) against the synthetic difference-in-differences (SDID) counterfactual, constructed as a weighted average of donor agency outcomes. The grey band along the bottom of each panel shows the SDID time weights, with taller bars indicating quarters that receive greater weight in the counterfactual. The vertical dashed line marks the start of the treatment period in January 2022. The donor pool consists of municipal police departments serving populations of at least 100,000 with a pre-2022 mean rate of the corresponding outcome above the median across all eligible donor agencies. The note below each panel reports the cumulative ATT, the implied percent change relative to the counterfactual post-period mean, the placebo p-value (500 iterations), and the number of donor units. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

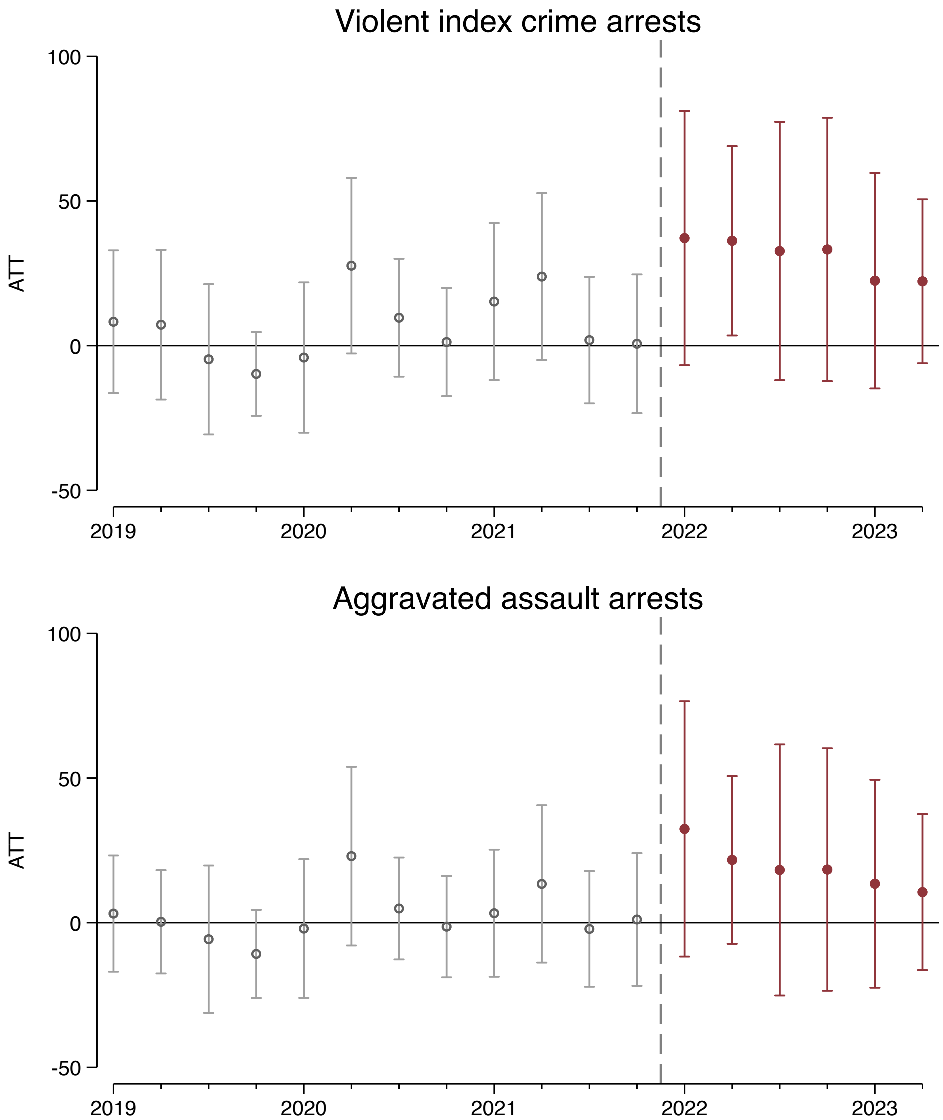
Appendix

Figure A.1: Pre-period structural break in donor outcomes: SDID vs. DID implied counterfactuals for the Western district



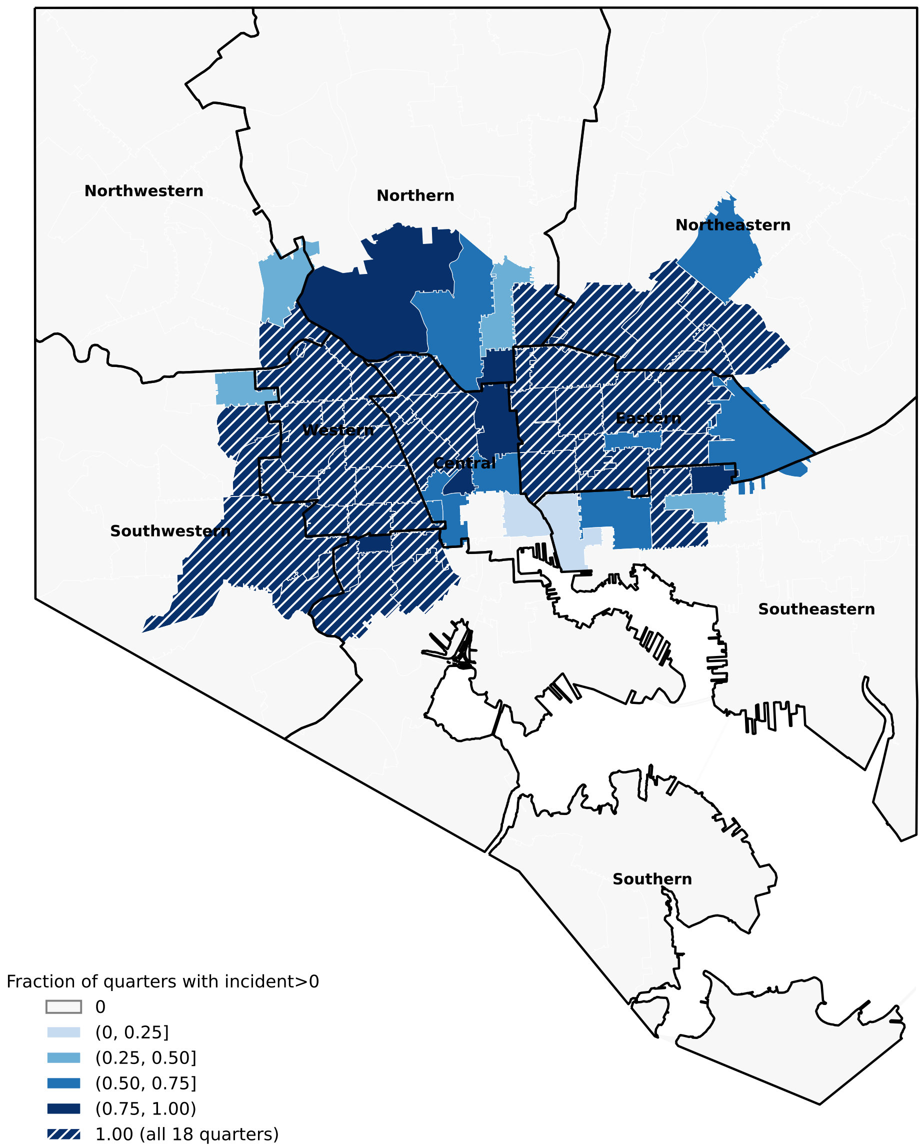
Note: Each panel plots the Western district's actual quarterly outcome (solid black, per 100,000 residents) from January 2015 through June 2023 alongside the counterfactual trajectories implied by SDID (dashed navy) and DID (dashed cranberry). The grey band along the bottom of each panel shows SDID's estimated time weights, with taller bars indicating quarters that receive greater weight in constructing the counterfactual. The vertical dashed line marks the pre/post boundary at year-end 2021. See Section 5.4 for the construction of the implied counterfactual under each estimator; the post-period mean gap between the Western district's actual outcome and the implied counterfactual recovers the estimator's ATT. The donor pool consists of all Baltimore police posts outside the Western and Southwestern districts, matching the main specification. Top panel: robberies. Bottom panel: drug arrests not co-charged with a violent index crime.

Figure A.2: Estimated effects of GVRS on violent index crime and aggravated assault arrest rates in the Western district



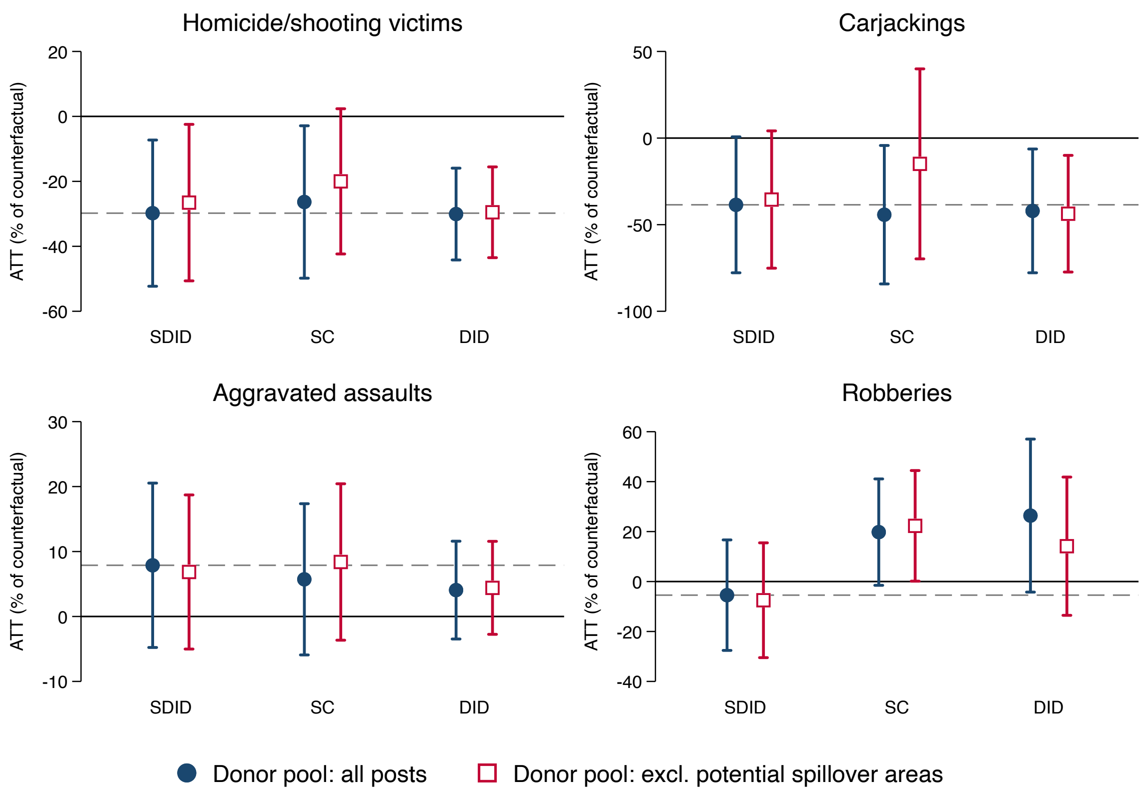
Note: Figure plots event-time SDID coefficients and 95% confidence intervals for rates of violent index crime and aggravated assault arrests in the Western district. Outcomes are reported arrest rates per 100,000 residents at the post-quarter level. Estimated using Ciccia's `sdid_event` package; see Section 5.2 for the construction of the post-treatment estimates and pre-treatment in-time placebos. Event time 0 denotes the first post-treatment quarter (the launch of GVRS in January 2022); the dashed vertical line at -0.5 marks the pre/post boundary. Maroon markers at event times $0, 1, \ldots, 5$ are quarterly ATT estimates over the 18-month post window from January 2022 through June 2023. Gray markers at event times $-1, -2, \ldots, -12$ are in-time placebo estimates over the twelve quarters preceding GVRS launch. The aggregate 18-month ATTs reported in Table 5 are not plotted here. The donor pool is all Baltimore posts outside the Western and Southwestern districts. Confidence intervals are constructed from 500 bootstrap iterations.

Figure A.3: ShotSpotter coverage of Baltimore police posts, 2019–2023



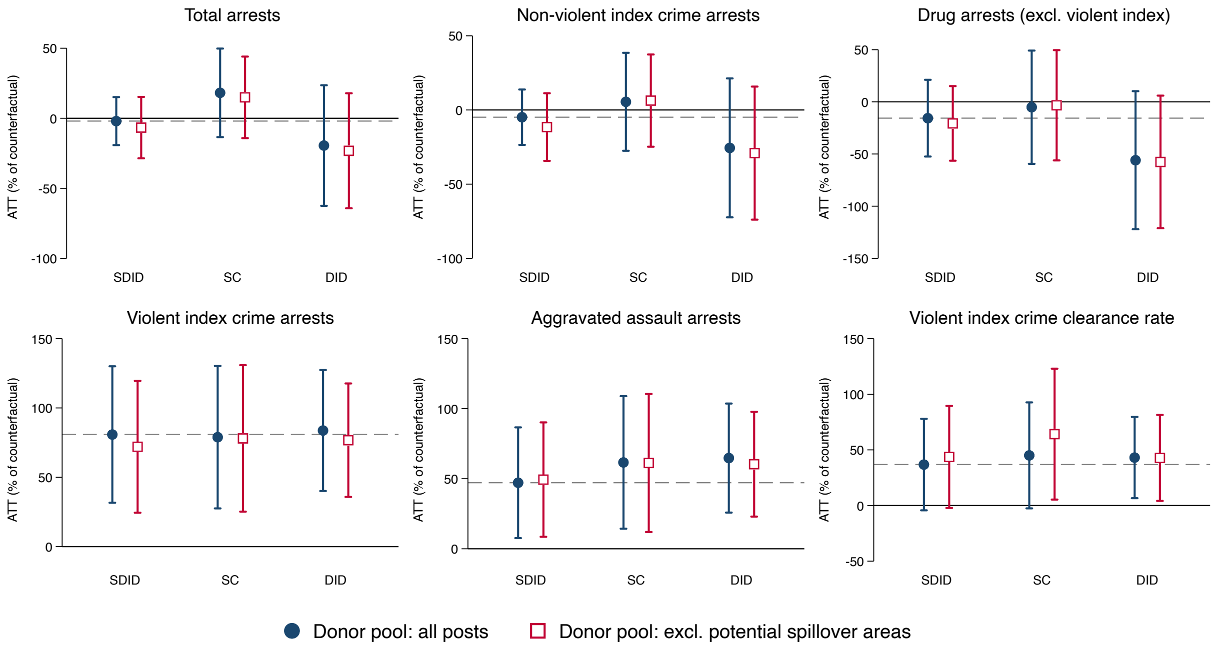
Note: Figure shades each Baltimore police post by the fraction of quarters from January 2019 through June 2023 (18 quarters) in which that post recorded at least one ShotSpotter incident. Posts shaded black are those with at least one ShotSpotter incident in *every* quarter of the 18-quarter window. Posts rendered in the blue gradient recorded a positive ShotSpotter incident in at least one but not all quarters of the window, with darker shading corresponding to higher coverage. Posts left unshaded (near-white) recorded zero ShotSpotter incidents over the entire window. Heavy black lines are Baltimore police district boundaries; thin white lines separate posts. The covered posts form a contiguous zone centered on the Central, Eastern, and Western districts; the Northwestern and Southern districts are essentially uncovered, and the covered posts in Northern, Northeastern, Southwestern, and Southeastern appear only where those districts border the core.

Figure A.4: Robustness of 18-month ATT estimates on violent crime rates to the choice of estimator and donor pool



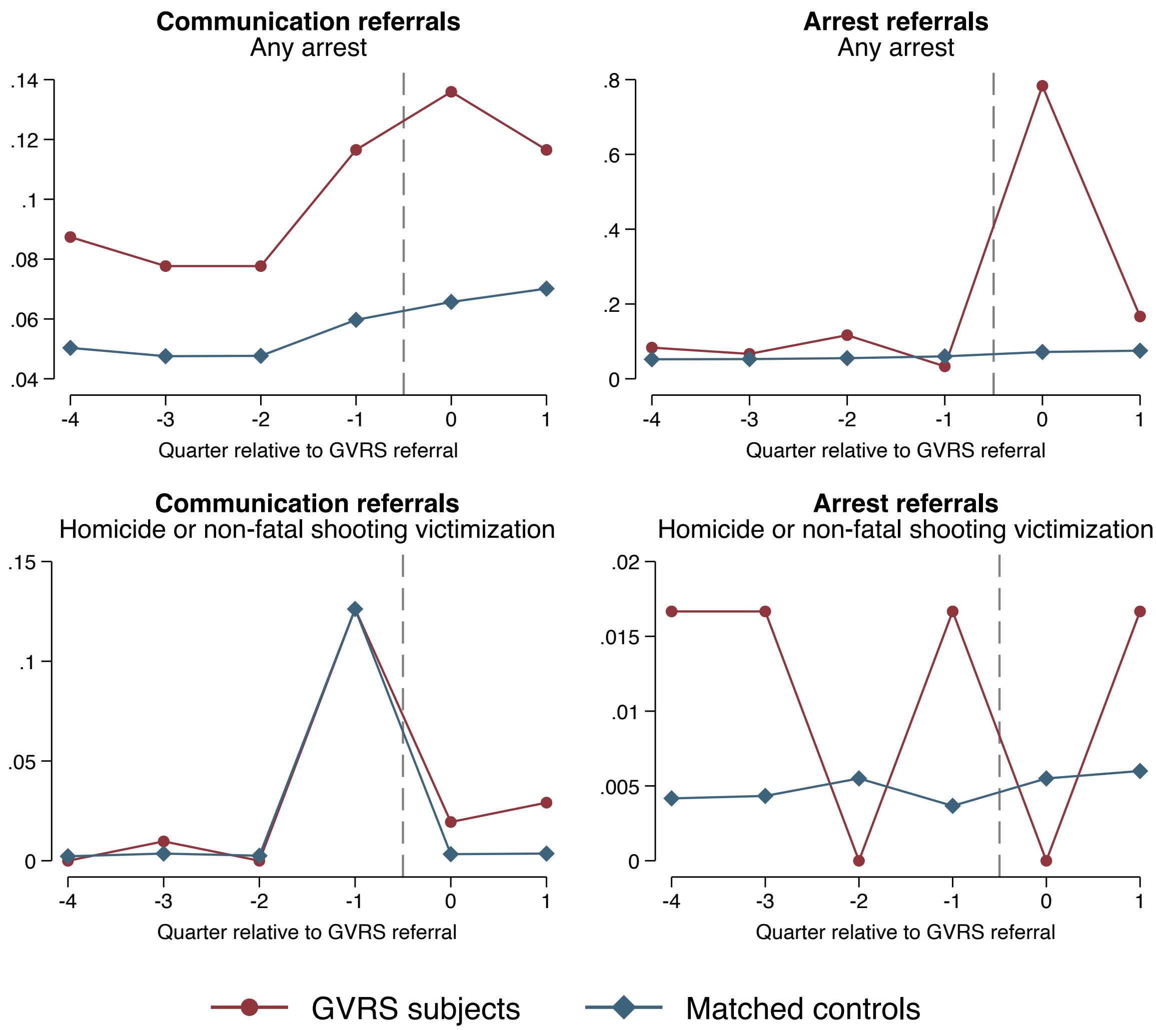
Note: Each panel plots 18-month ATT estimates for the indicated outcome from Table 4, expressed as a percentage of the counterfactual post-period mean implied by that specification. Each panel contains six point-and-whisker markers: three estimators—synthetic difference-in-differences (SDID), synthetic controls (SC), and two-way fixed effects difference-in-differences (DID)—crossed with two donor pools. Filled navy circles denote the "all posts" donor pool, which consists of every Baltimore police post outside the Western and Southwestern districts. Open cranberry squares denote a restricted donor pool that additionally excludes the Central district and posts in other districts that border the Western or Southwestern district boundary; this restriction implicitly accounts for potential spatial spillovers to adjacent areas. The solid black horizontal line in each panel marks zero; the dashed grey horizontal line marks the point estimate from the main specification (SDID, all donors), corresponding to the ATT reported in Table 4. Whiskers are 95% confidence intervals constructed from 500 clustered bootstrap iterations and rescaled to percent units using the same counterfactual post-period mean. The pre-treatment window is January 2015 through December 2021 (84 months); the post-treatment window is January 2022 through June 2023 (18 months). See Section 5.4 for discussion of the robberies panel, where SDID and SC/DID diverge.

Figure A.5: Robustness of 18-month ATT estimates on arrest outcomes to the choice of estimator and donor pool



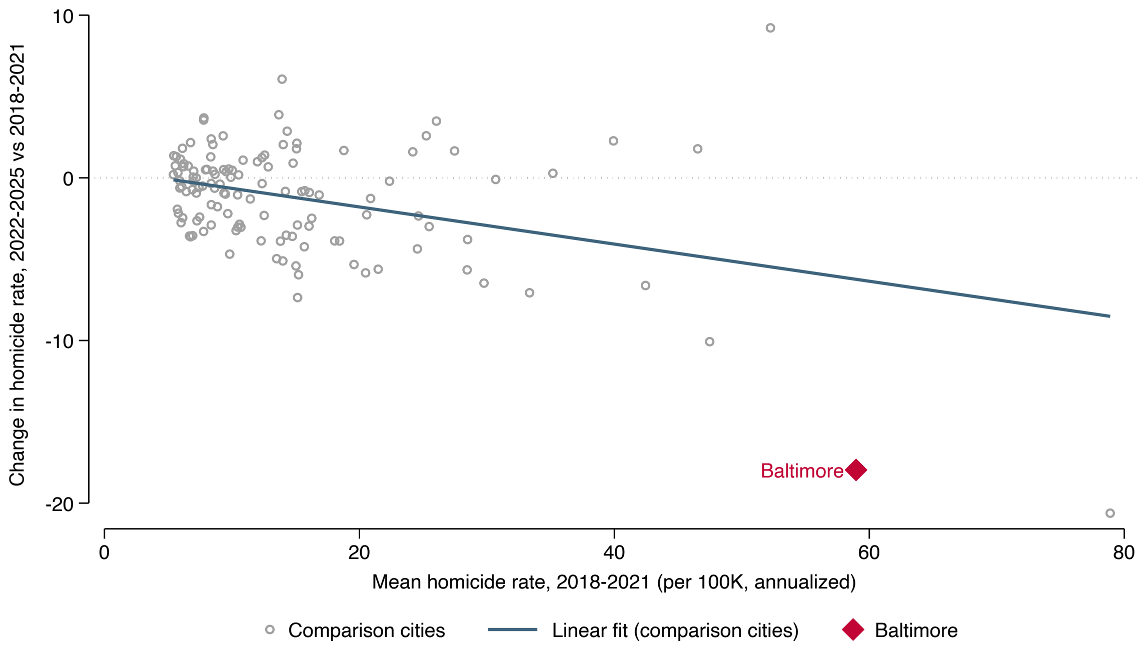
Note: Each panel plots 18-month ATT estimates for the indicated outcome from Table 5, expressed as a percentage of the counterfactual post-period mean implied by that specification. Panels in the top row correspond to Panel A of Table 5 (broad enforcement: total arrests, non-violent index crime arrests, drug arrests not co-charged with a violent index crime); panels in the bottom row correspond to Panel B (targeted enforcement: violent index crime arrests, aggravated assault arrests, violent index crime quasi-clearance rate). Marker conventions, donor pool definitions, reference lines, confidence interval construction, and pre/post windows are as in Appendix Figure A.4.

Figure A.6: Outcome time series for GVRS subjects and matched controls



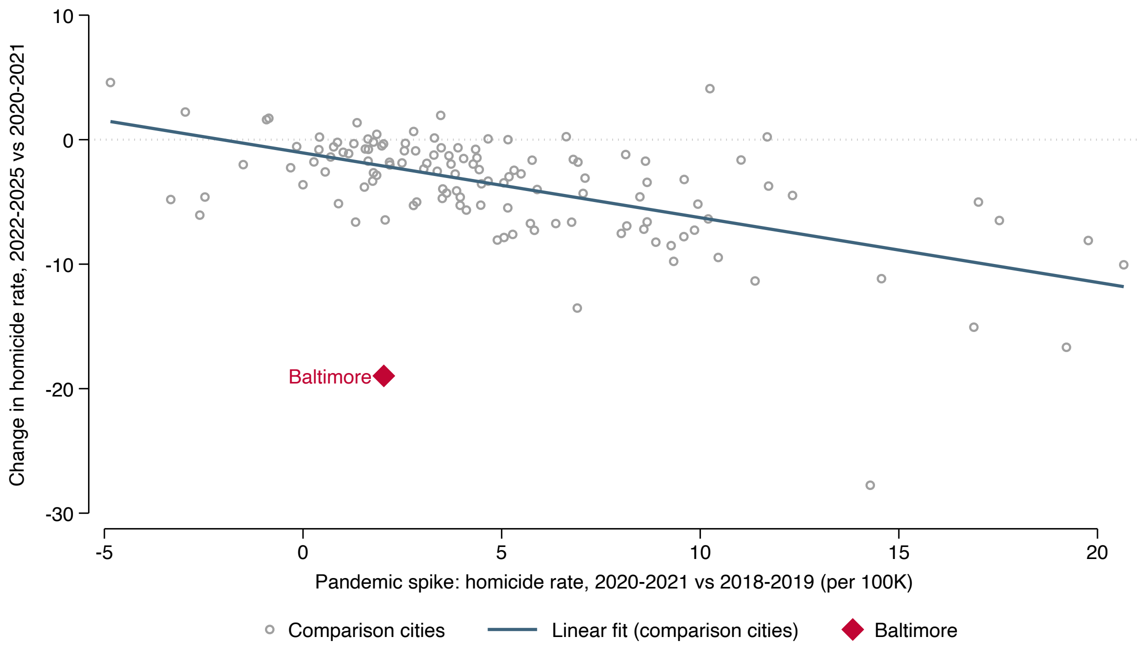
Note: Each panel plots the unweighted average of the indicated outcome at each quarter relative to GVRS referral, separately for GVRS subjects (red) and their propensity score-matched controls (blue). Sample, matching procedure, and outcome definitions are as in Figure 7. The dashed vertical line at -0.5 marks the pre/post boundary. Pre-period victimizations are mechanically non-fatal; post-period values include both fatal and non-fatal events.

Figure A.7: Baltimore's homicide decline against the cross-city mean reversion gradient, by pre-2022 level



Note: Each open circle is a comparison city in the above-median pre-2022 homicide rate donor pool (125 cities); Baltimore is the filled cranberry diamond. The horizontal axis is each city's mean annualized homicide rate per 100,000 over 2018–2021; the vertical axis is the change in that rate between the 2018–2021 and 2022–2025 periods. The navy line is the OLS fit through the comparison cities only (slope ≈ -0.11), capturing the tendency for higher-homicide cities to decline more thereafter—a mean reversion gradient. At Baltimore's pre-2022 rate the fit predicts a decline of about 6 per 100,000, against an actual decline of 18, placing Baltimore roughly 3.7 standard deviations below the fitted line and with a larger decline than all but one of the 125 comparison cities. The dotted horizontal line marks zero change. Data, donor pool, and estimation sample are as in Sections 8.1 and 8.2. See Section 8.4.

Figure A.8: Baltimore's homicide decline against the pandemic spike



Note: Each open circle is a comparison city in the above-median pre-2022 homicide rate donor pool (125 cities); Baltimore is the filled cranberry diamond. The horizontal axis is each city's pandemic-era spike—the change in annualized homicide rate per 100,000 between 2018–2019 and 2020–2021; the vertical axis is the subsequent change between 2020–2021 and 2022–2025. The navy line is the OLS fit through the comparison cities only (slope ≈ -0.52): cities with larger pandemic spikes saw larger subsequent declines, the pattern expected under mean reversion. Baltimore's spike was negligible (about 2 per 100,000), under which the fit predicts a decline of roughly 2, against an actual decline of 19—placing Baltimore about 5 standard deviations below the fitted line. Because Baltimore's homicide rate was persistently high before the pandemic rather than transitorily elevated, its post-2022 decline cannot be attributed to the unwinding of a pandemic spike. The dotted horizontal line marks zero change. See Section 8.4.

Table A.1: List of group/gang takedowns

Takedown	District	First arrest	Last arrest	Press conference	Indictments
Poe	Western	March 2022	March 2022	N/A	11
Princess Plaza	Western	August 2022	August 2022	August 2022	12
Wick Squad	Western	October 2022	January 2023	January 2023	9
Baltimore + Bentalou	Western	March 2023	June 2023	March 2023	33
Carey Boyz	Western	June 2023	June 2023	June 2023	8

Note: Coordinated multi-subject group enforcement actions in the Western district during the 18-month post-intervention window from January 2022 through June 2023. "First arrest" and "last arrest" are the dates of the first and last arrests recorded in the GVRS contact data for each takedown. "Press conference" is the date of the public announcement of the takedown, where one occurred. "Indictments" is the number of subjects indicted as part of the takedown.

Table A.2: Effects of GVRS on reported violent crime rates in potential spillover areas

	Homicide/shooting victims	Carjackings	Aggravated assaults	Robberies
Counterfactual post-period mean	76.3	48.5	332.2	234.4
ATT estimate	-1.6	-10.3	-2.3	-32.7
	(13.1)	(9.7)	(23.6)	(22.6)
Percent change (vs. counterfactual)	-2.1%	-21.2%	-0.7%	-14.0%
Bootstrap p-value	0.902	0.291	0.921	0.147
FDR q-value	0.921	0.581	0.921	0.581

Note: Each column reports SDID estimates of the average treatment effect on the treated (ATT) of GVRS in the Central district and posts neighboring the Western and Southwestern districts, over the 18-month post-intervention window from January 2022 through June 2023, for the indicated violent crime outcome. Outcomes are reported crime rates per 100,000 residents at the post-quarter level. The counterfactual post-period mean is the implied mean of the outcome in the potential spillover areas absent GVRS (observed post-period mean less the ATT); the percent change is the ATT divided by this counterfactual mean. Standard errors (in parentheses) and bootstrap p-values are from 500 clustered bootstrap iterations. FDR q-values are Benjamini-Hochberg-adjusted across the four outcomes. The pre-treatment period is January 2015 through December 2021 (28 quarters). The donor pool is all posts in Baltimore excluding the Western, Southwestern, and Central districts and posts neighboring the Western and Southwestern districts (N=75 donor posts). $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Table A.3: Sample selection for the GVRS person-level analysis

	Communication referrals				
	Recent shooting victims	Not recent shooting victims	Total	Arrest referrals	All
All Western district GVRS referrals through 2023Q2	32	147	179	97	276
Referred in 2022 (matching candidate set)	23	90	113	60	173
Matched to at least one control under holdout spec	13	90	103	60	163

Note: Each cell reports the number of GVRS subjects in the indicated subgroup at the indicated stage of sample selection. The first row is the broad sample used in the baseline descriptives table (Table 2): GVRS subjects whose first referral was in the Western district and occurred before July 2023. The second row restricts to subjects referred in 2022, defining the matching candidate population (we require subjects to be referred in 2022 to ensure at least 6 months of post-referral observation given that arrest data ends June 30, 2023). The third row further restricts to subjects who were matched to at least one control under the holdout propensity score procedure described in Section 6.2; this is the analysis sample used in the event study (Figure 7) and the ATT table (Table 8). Communication referrals are split by whether the subject experienced a non-fatal shooting victimization in the 90 days before referral, the criterion used to stratify communication referrals in the matching procedure. Subjects flagged as both communication and arrest first-referral types contribute to both columns; the All column reports unique subjects.

Table A.4: Baseline characteristics of GVRS subjects and the full pool of candidate controls

	Communication referrals: recent shooting victims		Communication referrals: not recent shooting victims		Arrest targets	
	GVRS	Controls	GVRS	Controls	GVRS	Controls
N	23	602	90	114,287	60	114,307
Demographics						
Age at referral	28.0	33.5	30.3	43.2	30.5	43.1
Black	1.00	0.51	0.99	0.73	0.98	0.73
Male	0.91	0.48	0.92	0.72	0.93	0.72
In year before referral						
Any non-fatal shooting victimization	1.00	1.00	0.00	0.00	0.05	0.01
Any arrest	0.30	0.13	0.26	0.06	0.25	0.06
Part 1 violent crime arrests	0.00	0.05	0.04	0.02	0.02	0.02
Part 1 non-violent-crime arrests	0.04	0.02	0.03	0.01	0.03	0.01
Simple assault arrests	0.13	0.04	0.08	0.03	0.03	0.03
Weapons-related arrests	0.04	0.04	0.18	0.01	0.12	0.01
Drug possession arrests	0.00	0.00	0.00	0.00	0.00	0.00
Drug distribution arrests	0.09	0.02	0.04	0.01	0.13	0.01
Disorderly arrests	0.00	0.00	0.02	0.00	0.00	0.00
Other arrests	0.35	0.04	0.18	0.02	0.12	0.02
More than 1 year before referral						
Any non-fatal shooting victimization	0.04	0.05	0.13	0.02	0.10	0.02
Any arrest	0.78	0.45	0.81	0.92	0.72	0.92
Part 1 violent crime arrests	0.74	0.18	0.47	0.16	0.42	0.16
Part 1 non-violent-crime arrests	0.57	0.17	0.47	0.22	0.35	0.22
Simple assault arrests	0.61	0.30	0.67	0.44	0.47	0.43
Weapons-related arrests	0.48	0.23	0.63	0.16	0.35	0.16
Drug possession arrests	1.70	0.71	1.62	0.60	1.43	0.60
Drug distribution arrests	1.17	0.43	1.17	0.20	1.02	0.20
Disorderly arrests	0.13	0.14	0.43	0.16	0.45	0.16
Other arrests	1.13	0.43	1.11	0.60	1.03	0.60

Note: The sample is GVRS subjects whose first referral was in the Western district and occurred in 2022, plus the full pool of candidate controls available to the matching procedure (see Section 6.2). Subjects are stratified by first-referral type and recent non-fatal shooting (NFS) victimization: communication referrals with at least one NFS victimization in the 90 days before referral; communication referrals without any NFS victimization in that window; and arrest targets. The candidate control pool excludes anyone with a Western-district arrest record before the treated subject's referral date and applies the same NFS condition as the corresponding treated stratum (for the communication-referral strata). Because pre-referral covariates depend on the referral date, the pool is constructed in a stacked panel across referral dates: a control who satisfies the inclusion criteria under multiple referral dates contributes one observation per qualifying date. The Controls columns report means over these stacked pool observations, and the N row reports the count of unique subjects in each cell. "In year before referral" aggregates the 1–90 and 91–365 day pre-referral windows; "more than 1 year before referral" aggregates the 1–2 year, 2–3 year, 3–5 year, and 5+ year pre-referral windows. Race is set to missing for subjects with no recorded race in the arrest data; sex is set to missing for the small number of subjects whose recorded sex is unknown.

Table A.5: Alternative estimators: synthetic control and difference-in-differences estimates of Baltimore homicide and violent crime rates relative to donor pool counterfactuals

	Homicide rate	Violent crime rate (excl. homicides)
Panel 1: SC, donor pool: above-median pre-2022 rate		
Counterfactual post-period mean	45.6	1491.2
ATT estimate	-4.6	316.0**
	(3.0)	(141.3)
Percent change (vs. counterfactual)	-10.2%	21.2%
Placebo p-value	0.119	0.025
N donor units	125	119
Panel 2: SC, donor pool: population $\geq$ 250K		
Counterfactual post-period mean	42.5	1491.1
ATT estimate	-1.5	316.1***
	(3.3)	(120.3)
Percent change (vs. counterfactual)	-3.4%	21.2%
Placebo p-value	0.661	0.009
N donor units	83	77
Panel 3: DID, donor pool: above-median pre-2022 rate		
Counterfactual post-period mean	57.8	1946.4
ATT estimate	-16.8***	-139.2
	(3.4)	(160.6)
Percent change (vs. counterfactual)	-29.1%	-7.2%
Placebo p-value	0.000	0.386
N donor units	125	119
Panel 4: DID, donor pool: population $\geq$ 250K		
Counterfactual post-period mean	58.1	1953.4
ATT estimate	-17.1***	-146.3
	(3.1)	(133.7)
Percent change (vs. counterfactual)	-29.5%	-7.5%
Placebo p-value	0.000	0.274
N donor units	83	77

Note: Companion to Table 9, reporting estimates from the synthetic control (SC) and standard two-way fixed effects difference-in-differences (DID) estimators in place of synthetic difference-in-differences. Each column reports the average treatment effect on the treated (ATT) for the indicated quarterly outcome, expressed per 100,000 population and annualized. Sample, treatment date, and outcome construction are as in Table 9. Placebo p-values are computed by re-estimating the model treating each donor unit as the treated unit in turn (500 placebo iterations). $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.