

THE QUARTERLY JOURNAL OF ECONOMICS

Vol. 139

2024

Issue 1

PREDICTING AND PREVENTING GUN VIOLENCE: AN EXPERIMENTAL EVALUATION OF READI CHICAGO*

MONICA P. BHATT
SARA B. HELLER
MAX KAPUSTIN
MARIANNE BERTRAND
CHRISTOPHER BLATTMAN

Gun violence is the most pressing public safety problem in U.S. cities. We report results from a randomized controlled trial ($N = 2,456$) of a community-researcher partnership called the Rapid Employment and Development Initiative

*The research had support from a wide philanthropic community, including Arnold Ventures, the Partnership for Safe and Peaceful Communities, JPMorgan Chase, the Chicago Sports Alliance, and both the Institute for Firearm Injury Prevention and the Population Dynamics and Health Program at the University of Michigan. A huge team of research staff made the study possible, with enormous thanks to Damilare Aboaba, Xander Beberman, Brenda Benitez, Ryan Carlino, Ran Cheng, Binta Diop, Brandon Domash, Mara Heneghan, Miguel Hernandez-Pacheco, Megan Kang, Leah Luben, Connor McCormick, Melissa McNeill, Evelyn Morris, Danielle Nemschoff, Michelle Ochoa, Priyal Patil, Mark Saint, Michael Tatone, Diamond Thompson, and Nathan Weil. We are indebted to Heartland Alliance and its extraordinary leadership team, to Zubin Jelveh and Ben Jakubowski for their work developing the prediction model, to Roseanna Ander, and to the remarkable local community organizations that made READI happen: Centers for New Horizons, Cure Violence, Heartland Englewood Outreach, Heartland Human Care Services, the Institute for Nonviolence Chicago, Lawndale Christian Legal Center, North Lawndale Employment Network, and UCAN. We are also grateful to our data providers, the Chicago Police Department, Cook County Sheriff's Office, and the Illinois Department of Corrections; to Betsy Levy Paluck and Andrew V. Papachristos for their insights in READI's early stages; and to Jennifer Doleac, Peter Hull, Lawrence Katz, Doug Miller, and Jens Ludwig for helpful comments. All opinions and any errors are our own and do not necessarily reflect those of our funders, implementing partners, data providers, or other government agencies.

© The Author(s) 2023. Published by Oxford University Press on behalf of the President and Fellows of Harvard College. All rights reserved. For Permissions, please email: journals.permissions@oup.com

The Quarterly Journal of Economics (2024), 1–56. <https://doi.org/10.1093/qje/qjad031>.
Advance Access publication on July 6, 2023.

(READI) Chicago. The program offered an 18-month job alongside cognitive behavioral therapy and other social support. Both algorithmic and human referral methods identified men with strikingly high scope for gun violence reduction: for every 100 people in the control group, there were 11 shooting and homicide victimizations during the 20-month outcome period. Fifty-five percent of the treatment group started programming, comparable to take-up rates in programs for people facing far lower mortality risk. After 20 months, there is no statistically significant change in an index combining three measures of serious violence, the study's primary outcome. Yet there are signs that this program model has promise. One of the three measures, shooting and homicide arrests, declined 65% ($p = .13$ after multiple-testing adjustment). Because shootings are so costly, READI generated estimated social savings between \$182,000 and \$916,000 per participant ($p = .03$), implying a benefit-cost ratio between 4:1 and 18:1. Moreover, participants referred by outreach workers—a prespecified subgroup—saw enormous declines in arrests and victimizations for shootings and homicides (79% and 43%, respectively) which remain statistically significant even after multiple-testing adjustments. These declines are concentrated among outreach referrals with higher predicted risk, suggesting that human and algorithmic targeting may work better together. *JEL codes:* C53, C93, I38, J08, K42.

I. INTRODUCTION

Over 170 people are shot each day in the U.S., with young Black men dying of gun homicide—by far their leading cause of death—at almost 20 times the rate of their white peers (CDC 2020).¹ There are strong arguments that more effective policing and punishment can reduce community gun violence (see Braga and Cook 2023). But there is also concern that common law enforcement strategies—aggressive policing that prioritizes street stops and low-level arrests, for example, or much greater use of prisons (Harcourt 2005; Raphael and Stoll 2013)—can impose high social costs, especially on the same communities that already bear the burden of gun violence.² This concern has fueled demand for ways to reduce shootings without the harms of overly aggressive or poorly targeted law enforcement.

One feature of community gun violence—its concentration—may be key to reducing it. In Chicago, for instance, five neighborhoods accounted for a quarter of the shootings in 2022, despite

1. Rates are for non-Hispanic Black and white men ages 18–34. These numbers exclude suicides, accidents, and shootings by police.

2. See Pattillo, Western, and Weiman (2004); Geller et al. (2014); Jones (2014); Ang (2021); Chalfin et al. (2022).

containing a tenth of the city's population. Within such neighborhoods the number of people involved in gun violence is small (e.g., [Braga 2003](#)). If the ex ante risk of being a shooting victim or offender is concentrated enough, then intervening with a small set of people could meaningfully and cost-effectively reduce shootings ([Green, Horel, and Papachristos 2017](#); [Abt 2019](#); [Heller et al. 2022](#)).³

Any individually targeted intervention must overcome two challenges. The first is identifying and engaging people at high risk of gun violence. This a difficult prediction problem given the complex determinants of shootings ([Chandler, Levitt, and List 2011](#); [Wheeler, Worden, and Silver 2019](#); [Heller et al. 2022](#)), and a difficult practical problem given the extreme levels of trauma and disconnection in this population ([Fagan and Wilkinson 1998](#); [Anderson 1999](#)).

The second challenge is finding effective ways to reduce the risk of being involved in a shooting. Unfortunately, we know relatively little about how to find, let alone engage, those at very high risk of gun violence without law enforcement involvement. Nor is there much rigorous evidence on what social services actually reduce this risk (see [Section II.A](#) and [Online Appendix A.1](#)). Nonetheless, many cities are now funding social services that try to overcome both challenges.⁴

This article evaluates a community-researcher partnership designed to tackle both challenges: the Rapid Employment and Development Initiative (READI) Chicago. READI operated in five of Chicago's highest-violence neighborhoods. The program sought to identify men at the very highest risk of shooting involvement using three referral pathways: (i) a machine learning algorithm based on administrative arrest and victimization records; (ii) referrals from local outreach workers; and (iii) screening among

3. A complement to this targeted, individual approach is to address the "root causes" of gun violence, such as concentrated disadvantage and access to guns. A targeted approach can be implemented and help quickly, while making structural changes will take more time (see [Haveman et al. 2015](#); *New York State Rifle & Pistol Association, Inc. v. Bruen*, 597 U.S. (2022)).

4. See <https://www.chicago.gov/content/dam/city/sites/public-safety-and-violence-reduction/pdfs/OurCityOurSafety.pdf>, <https://www.phila.gov/2021-04-14-how-the-city-is-addressing-gun-violence-2021-update-to-the-roadmap-to-safer-communities/>, <https://www.oaklandca.gov/topics/oaklands-ceasefire-strategy>, and https://monse.baltimorecity.gov/sites/default/files/MayorBMS_Draft_ViolenceReductionFrameworkPlan.pdf.

those leaving prison and jail. Over three years, 2,456 men were randomly assigned either to a READI offer or to a control group free to pursue other services.

Men assigned to READI were offered 18 months of subsidized, supported work combined with group cognitive behavioral therapy (CBT). The job was designed to provide several elements: a stable source of income to deter illegal work, an incentive to participate in the therapy, a place to build and reinforce new skills and norms, and a reason to spend less time in dangerous settings. Meanwhile, the CBT-informed programming was designed to foster several complementary behavior changes: to help participants reflect on their own thinking, practice less harmful responses in dangerous situations, and promote more positive behaviors and identities. Due to the significant barriers to participation that this population faces, READI also provided participants with referrals to housing, substance abuse, mental health, and legal services when needed.

With respect to the first challenge—identifying and engaging people at very high risk of shooting involvement—READI was a clear success. Prior to program referral, 35% of men in the study had been shot and 98% had been arrested, with an average of more than 17 prior arrests. Staggeringly, in the 20 months after randomization, there were 11 shooting and homicide victimizations for every 100 men in the control group—52 times more than among average Chicagoans, and 2.7 times more than among other similarly aged men in the five READI neighborhoods.

Despite many barriers to participating, 55% of men assigned to treatment attended at least one day of programming. Overall, participants worked an average of 30% of the hours available, prior to the suspension of in-person programming in March 2020.⁵ The subset of participants who continued to work remained engaged, working 75% of the weeks available to them while in-person work was occurring. This rate is comparable to interventions for much lower-risk populations (such as high school boys) and for much shorter transitional job programs ([Redcross et al. 2016](#); [Heller et al. 2017](#)).

5. The READI study was in the field from August 2017 through October 2021. At the start of the COVID-19 pandemic, CBT sessions shifted online and in-person work was temporarily suspended, though some payments continued; see [Section II.C](#). Because our outcome window is 20 months, about 76% of postrandomization person-day observations occurred before the pandemic.

On the second challenge—reducing serious violence—the results provide reason for both caution and optimism. We track three measures of serious violence involvement over a 20-month outcome period using matched administrative data: (i) shooting and homicide victimizations; (ii) shooting and homicide arrests; and (iii) other serious violent-crime arrests, such as robbery and aggravated battery.⁶ Our primary prespecified outcome is a standardized index that averages all three measures of serious violence with equal weights. We also specified several secondary analyses, including how READI affects the index components, and an index of all crime and violence incidents weighted by their social costs.

There is no detectable impact of READI on the primary outcome—the simple average of the three serious violence measures. The estimated effect of treatment on the treated (TOT) is -0.049 standard deviations ($p = .26$). When we break the index into its three components, however, we find suggestive evidence that READI reduced arrests for shootings and homicides. Relative to men in the control group who would have started programming if offered (control compliers), READI participants had 65% fewer shooting and homicide arrests (2.2 fewer per 100 participants). This result is statistically significant on its own, but not after adjusting inference for the three hypothesis tests involved in breaking the index into its components (unadjusted $p = .05$, adjusted $p = .13$). Results for the other two components are less precise. Point estimates show that participants had 12% fewer shooting and homicide victimizations but 11% more arrests for other types of serious violent crime. Confidence intervals are too wide to draw clear conclusions (adjusted $p = .8$ and $p = .7$, respectively).

When we weight incidents of crime and violence by the costs they impose on society, however, we estimate that READI reduced these social harms by at least \$182,000, and perhaps by as much as \$916,000, per participant—about a 50% decline relative to control compliers ($p = .03$). The increased precision comes from the fact that the large decreases in violence are concentrated in the most socially costly outcomes. Using a range of assumptions, these estimates imply READI's benefit-cost ratio is at least 4 to 1, and could be as high as 18 to 1.

READI also generated heterogeneous treatment effects. We prespecified that we would analyze results by referral pathway,

6. Our preanalysis plan is available at <https://osf.io/ap8fj/>.

and we find that these effects differ significantly from each other ($p = .03$). Participants referred by outreach workers saw serious violence involvement fall by 0.13 standard deviations (adjusted $p = .03$), driven by large and statistically significant reductions in shooting and homicide arrests (79%, adjusted $p = .03$) and victimizations (43%, adjusted $p = .08$) relative to control compliers in the same pathway. Hence, READI was more effective at reducing serious violence among outreach referrals.

It is harder to say why effects were larger for this pathway. The results are consistent with those men receiving a higher dose of programming, but also with them being more responsive to it. Outreach workers were asked to refer men at highest *ex ante* risk of gun violence involvement, which we refer to as selection on $\hat{Y}(0)$, or the predicted level of Y in the absence of treatment. Interviews suggest that, as anticipated, outreach workers also considered referrals' expected gains from treatment, which we refer to as selection on $\beta = Y(1) - Y(0)$, or the treatment effect of READI. Program staff frequently reported filtering out men who they felt were not "ready"—not open to a change of lifestyle or facing too many barriers to participation, and thus unlikely to engage in the program.

In exploratory analysis, we unpack treatment heterogeneity between outreach and algorithm referrals. Men referred by the algorithm had, on average, a higher algorithmically predicted risk of future gun violence involvement. Yet conditional on predicted risk, outreach referrals were subsequently involved in gun violence at higher rates. This suggests that outreach workers' screening methods successfully incorporated risk factors for $Y(0)$ that the algorithm could not observe. However, outreach workers were only partially successful in identifying men with a high β ; violence declined only among the subset of outreach referrals who also had the highest algorithmically predicted risk. Together, these patterns suggest that outreach workers were not simply selecting on expected gains. Rather, the combination of higher observable risk as captured by the algorithm and outreach-identified unobservables appears to predict treatment responsiveness. In other words, human and algorithmic referral mechanisms worked better together than alone.

From the perspective of scientific hypothesis testing, the mixed program effects we document make it difficult to give a definitive answer about how READI changed behavior. It likely reduced shooting and homicide arrests overall, as well as drastically lowered both arrests and victimizations for shootings and

homicides among men referred by outreach. But we fail to reject the null for all forms of serious violence across all subgroups. Future research replicating and refining READI's approach would be valuable to learn whether a combination of work and CBT can reduce gun violence among the men at highest risk of it.

From a policy perspective, binary hypothesis tests may not be the most useful basis for decision making. As others have argued, policy makers should weigh the importance of the outcome, the uncertainty of the estimates, and the availability of other interventions and evidence (Ziliak and McCloskey 2008; Manski 2019; Imbens 2021). From this perspective, a few aspects of the READI results are worth highlighting. For the primary index of serious violence, 74% of the treatment effect's confidence interval is below zero. Policy makers can weigh this against their level of uncertainty about the effectiveness of other social service approaches to reduce gun violence, as well as the increased social costs that aggressive law enforcement responses can generate. They can also use our estimate of READI's benefit-cost ratio—in essence, an importance-weighted sufficient statistic, as in Viviano, Wuthrich, and Niehaus (2021)—which suggests that society values READI's impact on violence 4 to 18 times more than the cost of running the program.

One clear lesson from these results is the potential for a targeted intervention to affect the total amount of gun violence in a city. Despite being less than 0.01% of Chicago's population, the 2,456 men in the study sample would have contained about 6% of Chicago's shooting and homicide victims during an average 20-month period in the absence of READI, costing society between \$711 million and \$3.6 billion.⁷ Despite being disconnected from and distrustful of many social institutions, these men proved willing to engage in READI. The fact that it is possible to identify and engage a relatively small group at such elevated risk of socially costly outcomes emphasizes the potential of continuing to experiment with approaches to help this extremely disconnected and underserved population.

7. We calculate the 6% figure by using the rate of shooting and homicide victimizations during the 20-month outcome period in the control group as the counterfactual rate in the treatment group without READI, yielding 271 victimizations. This is approximately 6 percent of the number of shooting victimizations in Chicago during an average 20-month period between August 2017 and August 2021 (about 4,600).

II. EXPERIMENTAL SAMPLE AND INTERVENTION

II.A. Context and Research Questions

READI was designed in response to an unprecedented 60% spike in Chicago's homicide rate from 2015 to 2016 (Kapustin et al. 2017).⁸ As in many cities, Chicago's shootings are extremely concentrated in neighborhoods with many low-income residents and historically high rates of violence. Within such neighborhoods, gun violence appears to be further concentrated among a small group of people; a wide range of evidence suggests that only a tiny percentage of the population engage in serious violent crime (Wolfgang and Tracy 1982; Braga 2003; Farrington et al. 2006; Green, Horel, and Papachristos 2017; Abt 2019).

Using targeted interventions to tackle concentrated gun violence is a long-standing idea, both in policing and among community violence interventions (CVIs) (Sherman and Rogan 1995; Braga et al. 2001; Skogan et al. 2008; Braga, Papachristos, and Hureau 2014; Butts et al. 2015; Braga, Weisburd, and Turchan 2018). Online Appendix A.1 discusses the research about targeted policing and CVI approaches. One of the most common, community-wide violence interruption, involves intervening in and mediating active disputes in a community. Such mediation is qualitatively and theoretically important to reducing violence, but it is also challenging to evaluate due to the difficulty of finding counterfactual communities (Farrell et al. 2016; Roman, Klein, and Wolff 2018). A review of the evidence characterizes it as mixed (Butts et al. 2015). A newer and complementary set of CVIs proactively offer preventative services to specific people or groups at high risk of gun violence involvement. This kind of targeted, tertiary prevention also has qualitative and theoretical promise, but there is no causal evidence so far of its effectiveness.

Separately, causal evidence exists on READI's two main program components—supported work and CBT—albeit for purposes and populations quite different than CVIs meant to reduce shootings. Studies of transitional jobs programs suggest that they are unlikely to reduce crime and violence on their own, but that strategies combining jobs and enhanced services such as CBT are

8. From 2016 to 2021, Chicago experienced an average of 23 homicides per 100,000 people annually, the vast majority due to guns. This rate is more than eight times higher than those in Los Angeles or New York, comparable to those in Philadelphia and Milwaukee, and around half of those in cities such as St. Louis or Baltimore.

more effective (MDRC 2013; Redcross et al. 2016; Cummings and Bloom 2020). Although CBT-informed programming alone can reduce violence involvement, there is some evidence it may be more effective when paired with an economic intervention (Wilson, Bouffard, and MacKenzie 2005; Lipsey, Landenberger, and Wilson 2007; Blattman, Jamison, and Sheridan 2017; Heller et al. 2017; Dinarte and Egaña del Sol 2019; Arbour 2022; Blattman et al. *forthcoming*). While both of READI's core program elements have shown promise, neither has been evaluated on a population as disconnected and at as high risk of gun violence as the one that READI aimed to serve.

Any CVI that seeks to prevent shootings by intervening with specific people must meet two criteria to be successful: (i) identify and engage a group of people at high enough *ex ante* risk of gun violence for it to be feasible to reduce shootings among them, and (ii) reduce their risk of being involved in a shooting. Both criteria pose significant challenges.

Identifying this population is a difficult prediction problem (Berk et al. 2009; Chandler, Levitt, and List 2011; Wheeler, Worden, and Silver 2019; Heller et al. 2022). Existing research provides little guidance about whether and which observable and unobservable characteristics can identify specific people at very high *ex ante* risk, particularly given the idiosyncrasies inherent in human behavior and in the likelihood of being involved in gun violence (i.e., whether a person is hit or missed). To the extent that gun violence involvement risk is transitory, this population may also change over time.

Once identified, finding these individuals can be extraordinarily challenging, especially if they keep a low profile to avoid encounters with police or their opposition. Once found, they may be disconnected from—and skeptical of—societal institutions and offers of help, in addition to facing many logistical barriers to participation such as housing instability, substance use disorders, and safety concerns about exposing themselves to certain people or places (Fagan and Wilkinson 1998; Anderson 1999).

Most CVIs use outreach workers' local relationships and expertise to find and engage clients—usually young men with high rates of recent violence exposure. But without a random comparison group, evaluations cannot tell us clients' gun violence risk in the absence of services. It is also unclear how many people at high gun violence risk are missed by relying solely on expert referrals, or whether those referrals have the highest gains from participation. Beyond the challenges of identifying, finding, and engaging

the relevant population, we lack evidence on what kinds of interventions can reduce shooting involvement.

As a result, the READI study was not designed with a sole focus on impact evaluation. Rather, we set out to answer three questions: (i) Can we identify men at high enough risk of future gun violence that there is scope to reduce shootings? (ii) Will they participate in a prosocial intervention? (iii) Will a combination of supported work and CBT reduce their involvement in serious violence?

II.B. Sample Selection, Referral Pathways, and Randomization

READI was a partnership between several organizations: Heartland Alliance, an antipoverty and human rights nonprofit based in Chicago that designed, developed, and managed READI; four organizations specialized in outreach; three organizations specialized in employment and CBT-based programming; the principal investigators; and the University of Chicago Urban Labs.⁹ All parts of the program were developed collaboratively with input from community workers and a Participant Advisory Council.

1. *Eligibility.* READI aimed to recruit men 18 and over at the highest risk of gun violence involvement. It focused on 5 of 77 Chicago neighborhoods with the highest levels and rates of gun violence (Figure 1). Providers grouped these neighborhoods into three sites: Austin/West Garfield Park, North Lawndale, and Englewood/West Englewood.¹⁰

Given READI's explicit goal of serving men at the highest risk of gun violence, it is worth emphasizing that being at high ex ante risk for an outcome does not necessarily mean being responsive to an intervention designed to reduce it—that is, having a high $\hat{Y}(0)$, the predicted level of gun violence involvement absent treatment, is not necessarily the same as having a high β , the treatment effect of READI. Understanding the relationship between $\hat{Y}(0)$ and β is nonetheless important for

9. Program implementation started in partnership with Centers for New Horizons, Cure Violence, Heartland Human Care Services, the Institute for Non-violence Chicago, Lawndale Christian Legal Center (LCLC), North Lawndale Employment Network, and UCAN. In September 2018, Heartland Alliance took over outreach services from Cure Violence in Englewood. In April 2021, UCAN took over outreach services from LCLC in North Lawndale.

10. READI continues in these sites using a modified program model for non-study participants.

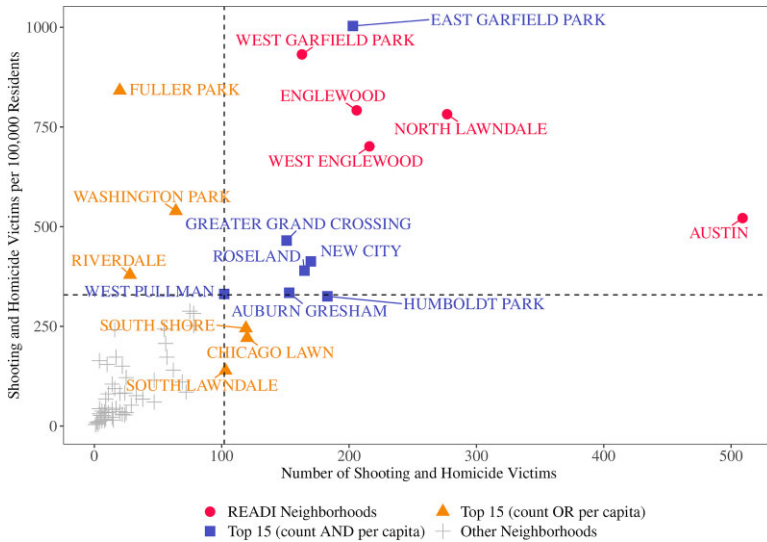


FIGURE I

Shooting Victims per 100,000 Residents (2016), by Neighborhood

Plot shows counts and rates of shooting and homicide victims in each of Chicago's 77 neighborhoods in 2016. Dashed lines represent the top 15 neighborhoods for each dimension.

future interventions. To encourage variation in participants that would allow us to study this, we designed three pathways to recruit men on a rolling basis, described below, with additional details in [Online Appendix A.3](#).

2. Referral Pathways. The algorithm pathway used administrative police data to predict the risk of being involved in gun violence as a victim or an arrestee over the next 18 months (the risk score).¹¹ Each time program slots became available for algorithm referrals, men with the highest risk scores who met READI's eligibility criteria were referred for randomization. This algorithmic approach is useful insofar as observables can successfully predict a person's involvement in gun violence ($\hat{Y}(0)$), but it will miss risk driven by unobservables or fast-moving situations that are not reflected in the police data used in the algorithm.

11. We predict risk scores for those with sufficient recent police contact; see [Online Appendix A.3.1](#). [Heller et al. \(2022\)](#) also provides a full description and analysis of a related prediction model, incorporating lessons from READI's prediction model.

To capture some of these unobservables and allow for possible selection on treatment responsiveness (β), the outreach pathway sourced referrals from outreach workers with extensive on-the-ground experience in the READI neighborhoods. These workers are privy to local information that may be absent from police records, but may be limited by the scope of their social networks or their incentives to offer particular people services and fill caseloads. They were asked to refer men at the highest risk of gun violence.

Finally, the re-entry pathway identified men leaving jail or on parole who may be missed by both outreach workers and the algorithm, and who may be at a particularly sensitive transition point. The re-entry pathway took the longest to become operational due to the complicated logistics involved with operating within carceral facilities. Because COVID-19 ended study recruitment early, this pathway is considerably smaller than the other two and than we initially intended. As such, we focus on differences between the first two pathways but report re-entry results separately for completeness.

Having three referral pathways allows us to assess how the different ways of identifying men at the highest risk of shooting and being shot performed: whether observables are enough to predict future gun violence with a machine learning algorithm, whether the algorithm identified a different set of people than the outreach workers did, and why those groups might differ (i.e., whether on-the-ground knowledge could capture unobservables in a way that improved program targeting, and whether human decision makers chose to select not only on gun violence risk but also expected responsiveness to the program).

3. *Randomization.* READI solicited referrals on a rolling basis from August 2017 to March 2020. New referrals ended early at the start of the pandemic, shifting the sample size to 2,456 from the original target of 3,000. Rolling referrals were made from all three pathways as program slots became available, to accommodate READI's growing capacity to absorb new participants over time and to focus on the people at highest risk, who may change over time. Outreach referrals began in August 2017, algorithm referrals in December 2017, and re-entry referrals in August 2018.

The randomization process varied slightly by pathway but followed the same general structure (see [Online Appendix A.3](#)). In all cases, we randomized at the individual level with a treatment probability of 0.5, within strata defined by site, referral pathway,

and randomization date.¹² After receiving referrals via the outreach or re-entry pathways and matching them to administrative data, or after identifying men with the highest risk scores via the algorithm pathway, we screened out anyone who had previously been randomized, was incarcerated, or had died since their referral.

II.C. The READI Program

READI was a bundled intervention designed to disrupt four proximate causes of gun violence. First is the instrumental use of violence in illegal markets, where illicit organizations lack legal ways to enforce contracts or compete for market share. Second is rational reputation building, where people use violence to signal strength as a way to deter future attacks in a dangerous environment. Third is reciprocity, where violence is a means to punish real or perceived slights, especially in settings where the legal system is not viewed as legitimate or just. Fourth is “irrational” behavior arising from mistakes and misperceptions—instances where fast decision making combined with fear, anger, or persistently biased beliefs about others’ intentions can result in violent escalation.¹³

READI’s key components were designed to address these proximate causes of gun violence. Note, however, that because it was designed and implemented extremely quickly to respond to crisis-level violence in Chicago, READI’s development was a learning process with the model changing slightly over time.

1. *Initial Outreach.* Engagement in all pathways began with outreach workers trying to locate men assigned a READI offer and persuading them to participate. These men were usually mistrustful of organized programming. Once located, outreach workers tried to convince these men to join READI, including by helping them obtain documentation to work legally or negotiate a truce with members of opposition groups in the program. This process of building trust and readiness could take days or months, depending on the person and their existing relationships with the outreach organization. Once a person was willing and ready to

12. In practice there is slight variation in treatment probability within strata; see [Online Appendix A.3.4](#).

13. For an extended theoretical discussion, see [Blattman \(2022\)](#) and [Abt \(2019\)](#). For a historical analysis of how these dynamics have played out in Chicago, see [Aspholm \(2020\)](#).

begin, the outreach worker connected them to their local READI employment organization. Attending an orientation seminar and signing job paperwork denotes the beginning of formal participation in READI—what we define as “taking up” the program.

2. *Supported, Subsidized Work.* To incentivize participation and provide an alternative to work in illegal, violence-prone markets, READI’s first component was supported, subsidized work. Participants could earn money at worksites five days per week (29.5 hours total), for up to 18 months.¹⁴ READI was explicitly not intended as a transitional job program to rapidly place men into full-time work. It was focused on violence reduction. Nonetheless, READI was informed by best practices, including a “career pathway” approach with four stages based on a participant’s progress (Online Appendix Figure A.I).

During the first stage, participants were typically assigned to crews performing outdoor work (such as park cleanup) or other basic services (such as packing meals for food pantries). They received transportation to and from the worksite, as safety was a core challenge. Later stages offered participants a greater variety of jobs and more independence, potentially including subsidized placements with local employers (e.g., some participants eventually worked in a local vehicle seat factory).

This tiered structure allowed participants to expand their skill set and earning potential over time, which prior research suggests is important for keeping them engaged. Participants in the first stage received a minimum hourly wage, initially \$11 but rising during the course of the study to match changes in the local minimum wage. Advancing to each stage was accompanied by a wage increase, among other benefits (Online Appendix Figure A.II).

Given the mixed results of prior jobs programs, even with populations at much lower risk of violence (MDRC 2013; Redcross et al. 2016; Cummings and Bloom 2020), the job component was not expected to reduce violence on its own. Nonetheless, over dozens of interviews and focus groups, participants and program staff said the job was a crucial incentive to participate (see Section III.C and Online Appendix A.8.1 for qualitative data and methods). The job was also designed to complement therapy

14. READI recognized that its target population faced many barriers to employment, so men who stopped attending work were allowed to later resume participating; see Online Appendix A.4.1.

by providing a place to practice and reinforce new thinking and behaviors, guided by staff at the worksites. Finally, given the level of violence involvement the population was expected to have, simply keeping participants busy during the week could potentially have a substantial incapacitation effect.

3. *CBT-Informed Curriculum.* To reduce the kind of “hot” decision making that leads altercations to escalate, and to help build new social identities and norms around violence, social interactions, and illegal employment, the second element of READI involved CBT. CBT is an approach for reducing maladaptive beliefs and behaviors and for promoting positive ones. Its methods can be applied to a range of behaviors, and CBT-informed therapies have been successful at reducing symptoms of depression, anxiety, phobias, traumatic stress, and hostility (Beck 1979; Beck and Dozois 2011).

CBT-informed therapies like READI teach participants that thoughts can influence actions, and help them practice actions designed to shape those thoughts. In 90-minute sessions for three mornings each week, program staff facilitated conversations and conducted exercises with small groups of men designed to help them become more conscious of their automatic thoughts, particularly inaccurate or negative ones about themselves or others that could lead to violence or other behaviors that are inappropriate in a given situation. Facilitators taught participants techniques to recognize these thoughts and respond to them in ways that are more constructive and less harmful. In group sessions and outside of them as “homework” (including at worksites), men had opportunities to practice these techniques on tasks of increasing difficulty. Through such “learning by doing,” CBT can gradually modify participants’ behavior and thinking.

Participants also received individual and small-group professional development and information sessions the remaining two mornings per week. They received a \$25 gift card for each CBT and personal development session attended. Afterward, they departed to their worksites for the day. The CBT sessions were available for all 18 months of READI and were a requirement for participation in supported work.

4. *Ongoing Outreach Support and Referrals to Other Services.* Initially, READI did not plan for outreach staff to be continually engaged with participants but to focus on recruitment. Within the first weeks, however, outreach workers were finding it necessary to devote a significant share of their time to this engagement, so

it was formally incorporated into the program (see the qualitative analysis in [Online Appendix A.8.5](#) for details on how and why outreach workers engaged with participants). Throughout our qualitative interviews and observations, READI staff and participants emphasized the extreme struggles participants faced in continuing with the program day to day. These include frequent fights, confrontations with program staff, and other serious disturbances on site—including occasional incidents of gun violence from outsiders and participants. Outreach workers provided almost continuous support, applying their training in conflict mediation, de-escalation, and restorative justice, including at worksites.

In addition to these safety issues, participants faced other serious challenges, including episodic homelessness, family quarrels, financial difficulties, arrests and other legal troubles, parole commitments, physical and mental health struggles, and other issues that hindered participation. Outreach and program staff helped participants navigate these problems, including by making referrals to organizations providing substance abuse treatment, housing, or legal services.¹⁵ Often, these situations served as opportunities to use skills acquired in the CBT sessions and entrench the new behaviors as habits.

5. Services Available to the Control Group. Men in the study were free to access alternative programming. It is reasonable to assume that some did, particularly men in the control group who lacked access to READI. To our knowledge, no other provider offered programming of READI's length or intensity, with a similar combination of services, delivered to men facing a similarly high level of risk during the study period and in the READI neighborhoods.¹⁶ Though the number of organizations in Chicago offering individual programming intended to reduce gun violence has grown since READI started in August 2017, few provide jobs, let alone for 18 months or in combination with CBT.

15. Criminal legal services were focused on dealing with a client's prosecution. As such, they were unlikely to affect our measure of violent offending, which occurs prior to prosecution, at the point of arrest.

16. The closest comparable program is Chicago CRED, which provides jobs, life coaching, trauma counseling, and education to a similar population of young men but operates mostly in the Roseland and West Pullman neighborhoods. See <https://www.ipr.northwestern.edu/documents/reports/ipr-n3-rapid-research-reports-cred-outreach-jan-22-2021.pdf>

Unfortunately, due to their level of risk, mobility, and distrust, it is impossible to survey control group men to learn about their program participation directly. Service providers also tend to be protective of client identities. We can, however, indirectly learn about this from the control group's primary conduit to alternative programming: outreach workers. Members of the research team conducted focus groups with nearly all program staff, as well as 220 hours of field visits where explicit attention was paid to engagement of the control group. [Online Appendix A.8](#) summarizes the analysis. Briefly, several staff described an intense effort to assist the highest-risk men, largely because they felt these men were in life-or-death situations. These outreach staff tried to stay in touch with control group men, mentor them, and connect them to temporary work agencies or mentorship programs. Although there are scattered reports of success at these efforts, one outreach worker expressed a common view that "there is nothing like READI around here." Staff efforts to connect control group men with other programs were also heterogeneous, with some staff reporting that they did not have the bandwidth to provide continued support.

Any support provided to control group men, however modest, would likely lead us to underestimate READI's treatment effects and overstate its costs. This is probably most acute for men referred by outreach workers, as their pre-existing relationships could have facilitated ongoing contact and access to alternative services. In contrast, outreach workers typically had no information about control group men from the algorithm pathway.

6. *Effects of the COVID-19 Pandemic on Program Delivery.* In March 2020, all randomization, qualitative data collection, and in-person READI programming was suspended. The CBT sessions transitioned to being held online. Supported work was paused altogether, with participants receiving "standby pay" from March through July 2020. In August 2020, in-person work and CBT sessions resumed. Participants had the option of going to work in person or earning up to \$50 per day attending professional development sessions remotely.

Although it did not undermine the study's internal validity, the pandemic affected the study in two ways. First, our sample size is almost 20% smaller than initially intended due to randomization ending early. Second, services were severely disrupted for those already participating, and the resumption of in-person

work after a pause of many months made it difficult to re-engage some men. Since the meaning of “participation” changed with the pandemic, in the main text we report overall take-up rates, program hours, and earnings, but limit reports of retention to the pre-COVID period. [Online Appendix A.5.3](#) provides more detail.

III. DATA AND EMPIRICAL STRATEGY

III.A. Outcomes and Data

The inherently unobservable or “latent” outcome of interest is a person’s involvement in serious violence, either as a perpetrator or a victim. We proxy for both kinds of violence involvement using administrative arrest and victimization records from the Chicago Police Department (CPD). The main advantage of police data is that we can use them to follow a large number of study members over a long period of time, with minimal sample attrition and relatively low cost. This makes the study feasible to conduct, as it would not be possible to track behavior via surveys over long periods of time with such a disconnected and difficult-to-locate population. Administrative records also avoid social desirability bias that might prevent people from honestly disclosing their serious violence involvement.

Police data also have serious limitations. One is that victimizations only appear when victims are willing to report incidents to the police. That said, because all health care workers in Illinois are legally required to report shooting victimizations to local police (20 ILCS 2630/3.2), and because shooting victims are very likely to seek medical care, underreporting of nonfatal shooting victimizations is likely to be minimal.¹⁷ Similarly, homicide victimization is widely thought to be mostly free of underreporting in police data ([Loftin and McDowall 2010](#); [Carr and Doleac 2016](#)).

A more serious challenge is the use of arrests. Arrests likely understate offending due to low clearance rates, even for serious crimes: only 26% of homicides and 5% of nonfatal shootings in Chicago in 2016 resulted in an arrest ([Kapustin et al. 2017](#)). Arrests are also subject to potentially biased police decisions and

17. Some underreporting may still occur if shooting victims self-treat or seek care from medical providers outside of Chicago who may not report such victimizations to the CPD. Based on our conversations with violence prevention, medical, and law enforcement practitioners in Chicago, we think the magnitude of such underreporting is likely to be small.

may be mistaken or wrongful. The chance of an offense resulting in an arrest (or the chance of a false arrest) may also vary by demographics or neighborhood. Although these issues introduce error into arrest measures that could make it harder to detect treatment effects on offending, they affect both treatment and control groups, and so should not bias impact estimates. The key assumption is that treatment does not change the probability of arrest conditional on actual criminal behavior.¹⁸

Finally, the data in this article do not capture victimizations or arrests outside of Chicago. However, if READI's steady paycheck increased the time participants spent in Chicago, then their violence involvement would be more likely to appear in our data, causing us to understate treatment effects.

III.B. Primary Outcomes

To proxy for the latent variable of interest, we prespecified a single primary outcome, serious violence involvement, which is an index that standardizes and averages arrests and reported victimizations for serious violent crimes over the 20 months postrandomization.¹⁹ The three components of this index are (i) shooting and homicide victimizations, (ii) shooting and homicide arrests, and (iii) other serious violent-crime arrests.²⁰

Though READI's emphasis is on reducing shootings and homicides, we included arrests for serious but not gun-related violent offenses in the primary outcome. We did this because we initially underestimated READI's ability to identify people with future shooting involvement, and we were worried that a sole focus on shootings would leave too few incidents to detect program effects. Similarly, in our preanalysis plan we pooled shooting

18. This assumption is more likely to hold for the serious violence measures that are our focus, rather than for lesser offenses where there is more room for police discretion; see [Online Appendix A.2.5](#).

19. The 20-month follow-up was chosen to measure effects during program-ming: 2 months to locate and recruit men, plus 18 months of possible program-ming. We plan a 40-month analysis as well.

20. Arrests are limited to incidents that occurred during the 20-month outcome period. Shooting arrests include any aggravated assault or aggravated battery that involves shooting a gun. Other serious violent-crime arrests include other nonshooting or homicide violent offenses historically included in "Part I" of the Uniform Crime Reporting program: aggravated assault and aggravated battery (excluding homicide, manslaughter, and nonfatal shootings), robbery, and criminal sexual assault. See [Online Appendix A.2.2](#).

and homicide arrests and other serious violent-crime arrests into a single component, as we did not anticipate having sufficient power to detect effects on shooting and homicide-specific outcomes. However, upon seeing the high level of risk among men in the study sample, and given READI's focus on preventing shootings and homicides, we opted to separate these incidents from arrests for other serious violent crimes when constructing the index. This decision has a negligible effect on our estimates of READI's effect on the index itself (Online Appendix Table A.IV). When estimating effects separately for the three (rather than two) components of the index, the multiple-hypothesis adjustment (described below) penalizes our inference for the additional hypothesis test.

An index is useful for increasing power to the extent that all underlying components move in the same direction—something we initially expected. But because READI can affect index components differently, and to better understand what behavior may be changing, we also prespecified that we would estimate effects separately for each component, correcting for the increased probability of Type I error that comes from testing multiple hypotheses, and calculate an index where all arrests and victimizations are weighted by their social cost. Viviano, Wuthrich, and Niehaus (2021) argue that a social cost-weighted index may be a more helpful way to summarize effects and provides a sufficient statistic for policymaking.

We implement two multiple-hypothesis testing adjustments. First, we control for the family-wise error rate (FWER) among the three index components by using a free step-down resampling method (Westfall and Young 1993; Anderson 2008).²¹ Second, we increase power by allowing for some proportion of null hypothesis rejections to be false, controlling for the false discovery rate (FDR) (Benjamini and Hochberg 1995). We report the q -value, an analog to an adjusted p -value, which reports the smallest proportion of false null rejections we would have to accept to reject the hypothesis under the FDR control procedure.

21. We report the adjusted p -value under strong FWER control. When we apply these adjustments to our heterogeneity tests, we treat the primary outcome index across the three referral pathways as a family, and then treat the index components in each referral pathway as its own family.

III.C. Qualitative Data Collection

Between August 2017 and March 2020, two of the investigators and a qualitative research team conducted across all sites 220 hours of formal field observation, 16 focus groups with 90% of READI staff in spring 2019, and 23 semistructured, recorded interviews with participants. In addition, we piloted a survey instrument in the Austin/West Garfield Park site with 66 participants in winter 2020. These data were gathered to describe participants' experiences and perceptions before and after take-up, inform program design, and aid in interpreting impact estimates. Because these data were collected prior to impact analysis, they were used to explore *ex ante* hypotheses of interest rather than to understand *ex post* patterns or outcomes—a form of exploratory and inductive qualitative research known as “emergent themes” (Williams 2008). [Online Appendix A.8](#) describes these data, methods of analysis, and conclusions in more detail.

III.D. Estimating Treatment Effects

We estimate intent-to-treat (ITT) effects via the OLS regression:

$$(1) \quad Y_i = \beta T_i + \lambda \mathbf{X}_i + \gamma_s + \varepsilon_i,$$

where Y_i is the 20-month outcome for individual i ; T_i indicates assignment to an offer of READI; \mathbf{X}_i is a vector of prerandomization arrest, victimization, incarceration, and demographic characteristics; and γ_s is a vector of randomization strata fixed effects.²² We estimate heteroskedasticity-robust standard errors.

The ITT estimate represents the effect of having an offer to participate in READI. Given incomplete take-up, the ITT will understate the effect of READI participation. We therefore also estimate the TOT by using random assignment as an instrumental variable (IV) for participation. We define participation as a binary indicator equal to 1 for those who attended the initial orientation and signed job paperwork. To provide a sense of the proportional change for compliers, we report control complier means (CCMs).²³

22. For more on prerandomization characteristics, deviations from the pre-analysis plan, and robustness to the inclusion of different covariates and alternatives to strata fixed effects, see [Online Appendix A.5.1](#).

23. Because selection into take-up is likely nonrandom, CCMs, rather than control means, are the appropriate benchmarks for interpreting the magnitude of IV estimates (Abadie 2003; Kling, Liebman, and Katz 2007). The CCM is the

1. Threats to Internal Validity

i. *Differential Mortality and Incarceration.* One consequence of evaluating programs for high-risk samples is that death or imprisonment will censor our dependent variables, possibly in ways that are correlated with treatment assignment. For example, if treatment delayed the risk of a homicide, then treatment group men would have had more opportunities than control group men to be arrested or victimized. In some respects, this change in time availability is part of the treatment effect. Overall counts of incidents still reflect how much violence there was among each group, which is why we made these counts our primary outcomes.

Still, if we are interested in whether READI changed violent behavior and not just the number of events, then such a shift in incapacitation could mask changes in outcomes conditional on being able to engage in one's normal activities. Recognizing this, we prespecified that we would test for differential rates of incarceration and death. Although we do not find significant differences in the overall amount of time treatment and control group men are incapacitated, we still verify in [Online Appendix A.5.2](#) that our main results are robust to adjustment for differential censoring.

ii. *Spillovers.* The ITT and TOT are always interpretable as the average difference between the treatment and control groups (or compliers) after READI was implemented. But interpreting them as estimates of the direct effect of being offered or receiving READI relies on the stable unit treatment value assumption (SUTVA)—that one person's treatment status does not affect another's potential outcomes. Given how social crime and violence can be, and that gun violence in particular is often a response to others' behavior, it is plausible that SUTVA could fail due to such spillovers (see [Online Appendix A.5.4](#)).

In a separate project, [Craig, Heller, and Rao \(in progress\)](#) estimate social spillovers. By combining multiple measures of social networks at baseline with the variation in treatment exposure generated by randomization, they identify how an intervention

estimated outcome mean for control group men who would have taken up READI had it been offered to them (control compliers). If no control group men take up (no always-takers), then the CCM is calculated by subtracting the TOT estimate from the outcome mean among treatment group compliers. Since three control group men took up, we use the CCM calculation in [Heller et al. \(2017\)](#). The handful of control crossovers technically make our IV estimate a local average treatment effect (LATE) rather than a TOT. But since crossover is so rare, we refer to the estimate as the TOT for simplicity.

spreads through social networks while overcoming the typical challenges of endogenous ties and common shocks in the peer effects literature. While that project is broader than READI, [Craig, Heller, and Rao \(2022\)](#) reports early READI-specific results to aid with the interpretation of the estimates reported in this article. They show that in the study population, there is no definitive sign of adverse spillovers (if anything, the results in this study may mask declines in other serious violent-crime arrests). But given the sample size, the analysis is somewhat underpowered.²⁴

IV. DESCRIPTIVE STATISTICS, REALIZED RISK, AND TAKE-UP

IV.A. *Baseline Characteristics and Balance*

[Table I](#) reports baseline summary statistics and tests of balance. The average age of study men at referral was 25, and nearly all (97%) are Black. Summary statistics on past arrest and victimization records confirm that the three referral pathways successfully identified men with very high levels of prior violence involvement. Baseline counts cover since 2010 for shooting victimizations and since 1999 for other arrest and victimization outcomes. For every 100 men in the sample, there were an average of 7.6 prior arrests for shootings or homicides, 46 prior shooting victimizations, and 90 arrests for other serious violent crimes (aggravated assault and battery, sexual assault, and robbery). The risk score shows that based on our algorithm, we expected 11.4% of the sample to be either arrested for or the victim of a violent gun crime in the 18 months after referral.

Prior involvement in other kinds of crime and violence was also high. The average study member experienced over 17 arrests prior to randomization and 3.4 reported victimizations, and had spent roughly 175 days in jail or prison in the prior 30 months, with 4% being incarcerated at the time of randomization due to imperfect screening. Though we lack any direct way to measure it with the data available to us, prior ethnographic work (e.g., [Aspholm 2020](#)) and the observations of READI program staff suggest that many of these men were involved in the neighborhood cliques that make up Chicago's fractured gang landscape.

24. They find clearer evidence of a READI-driven decline in drug crime arrests that is masked in the ITT estimate reported here, because control group men who have co-arrest ties to treatment group men also show reduced drug crime arrests. Combined with declines in these arrests among the peers of treatment group men who were outside the study sample entirely, it is likely the ITT estimate reported in this article misses a net decline in drug crime arrests from the intervention.

TABLE I
BASELINE CHARACTERISTICS

	Control mean	Treatment mean	Pairwise <i>p</i> -value
<i>N</i>	1,232	1,224	
Demographics			
Black	0.969	0.971	.864
Age	25.3	25.1	.431
Primary outcome components, counts			
Shooting victimizations	0.479	0.436	.119
Shooting & homicide arrests	0.079	0.073	.571
Other serious violent-crime arrests	0.915	0.878	.420
Risk prediction			
Predicted involvement in a violent gun crime (risk score)	0.115	0.114	.798
Missing risk score	0.108	0.080	.008
Arrest counts			
All arrests	17.1	17.7	.194
Less serious violent-crime arrests	1.5	1.6	.960
Drug crime arrests	4.9	5.3	.057
Property crime arrests	1.6	1.7	.643
Other crime arrests	8.0	8.3	.367
Victimization counts			
All victimizations	3.5	3.3	.297
Other violent victimizations	2.4	2.3	.474
Nonviolent victimizations	0.581	0.560	.649
Incarceration measures			
Days incarcerated	177.3	174.2	.704
Incarcerated at randomization	0.042	0.039	.671
Joint test			
<i>p</i> -value on <i>F</i> -test			.302

Notes. Pairwise *p*-values are from tests of treatment-control difference using heteroskedasticity-robust standard errors and controlling for randomization strata fixed effects. Arrest and victimization measures include all available CPD data from 1999 (2010 for shooting victimizations) through the time of randomization, with counts winsorized at the top 99th percentile. Other serious violent-crime arrests include criminal sexual assault, robbery, and nonshooting aggravated assault and battery. Less serious violent-crime arrests include non-Part I violent-crime arrests, such as simple assault and battery and domestic violence. Other (nonshooting) violent victimizations include sexual assault, robbery, aggravated assault and battery, and simple assault and battery. Nonviolent victimizations include all other incidents such as burglary, stalking, and threats. Risk score is missing for 231 individuals who did not have at least one arrest or two victimizations within the 50 months prior to randomization (risk score *N* for control = 1,099 and for treatment = 1,126). Race is missing for 38 individuals (race *N* for control = 1,211 and for treatment = 1,207). Joint test includes randomization strata fixed effects and the covariates listed here, excluding all arrests and all victimizations since they are linear combinations of other variables.

Of the 17 baseline variables in Table I, two have treatment–control differences with $p < .1$ —no more than would be expected by chance. A joint test of significance has a p -value of .3. We control for these variables in our analysis to account for any chance differences.

We also describe the limited information we have on our sample’s educational attainment and employment before READI. Importantly, this information is only available for treatment group men who took Heartland Alliance’s intake survey during orientation. Because of the selection into this sample, these data are not part of our balance tests or experimental analysis. Still, they are useful for painting a fuller picture of READI participants.

On average, 60% of the 535 survey respondents reported having less than a high school diploma. Thirty-one percent reported having graduated or earned a GED, and 8% reported attending some college without a degree.

Heartland’s survey question about employment changed over time. For the 207 people with data before February 2019, 52% reported working in the last year, and just under 12% reported having never worked. The 305 people with data after that point were asked only about working “prior to READI,” without a specific look-back period. Twenty-eight percent reported working for wages; 15% reported being self-employed; and 55% reported being out of work or unable to work. The Heartland survey did not explain whether participants should consider work in the illegal market as “working.” In a survey we conducted of 66 participants in Austin/West Garfield Park, 45% reported receiving a formal paycheck in the last typical week before READI, consistent with the results of the Heartland survey and suggestive of respondents reporting mostly legal work. In addition, 38% of respondents to our survey reported legal but informal work, referred to as a “side hustle” that did not require risk. Finally, 59% reported a risky side hustle, a proxy for illegal work, including drug sales and theft.

IV.B. Pathway Differences and Realized Risk

The top panel of Table II reports baseline statistics by pathway (Online Appendix Table A.I shows baseline balance by pathway). Relative to outreach referrals, algorithm referrals had longer arrest and victimization histories on most measures, despite being about a year younger. Their number of prior serious

TABLE II
BASELINE CHARACTERISTICS AND REALIZED RISK BY PATHWAY

	All	Algorithm	Outreach	Re-entry	<i>p</i> -value, test of pathway difference
Baseline					
<i>N</i>	2,456	1,232	878	346	
Age	25.2	24.6	25.6	26.4	<.001
Black	0.970	0.963	0.986	0.953	<.001
Shooting victimizations	0.458	0.631	0.287	0.275	<.001
Shooting & homicide arrests	0.076	0.080	0.064	0.092	.179
Other serious violent-crime arrests	0.897	1.0	0.667	1.0	<.001
Ever shot	0.347	0.463	0.233	0.225	<.001
Ever arrested	0.977	1.0	0.945	0.974	<.001
All arrests	17.4	20.4	13.6	16.1	<.001
Ever victimized	0.835	0.908	0.768	0.743	<.001
All victimizations	3.4	4.5	2.3	2.3	<.001
Days incarcerated	175.7	129.3	155.3	392.8	<.001
Predicted involvement in a violent gun crime	0.114	0.137	0.089	0.080	<.001

TABLE II
CONTINUED

	All	Algorithm	Outreach	Re-entry	<i>p</i> -value, test of pathway difference
Realized risk among controls					
<i>N</i>	1,232	616	438	178	
Shooting & homicide victimizations	0.110	0.106	0.114	0.118	.874
Shooting & homicide arrests	0.027	0.023	0.032	0.028	.683
Other serious violent-crime arrests	0.054	0.057	0.050	0.056	.912
Ever shot or killed	0.103	0.106	0.100	0.101	.961
Ever arrested	0.636	0.700	0.578	0.562	<.001
All arrests	1.7	1.9	1.4	1.3	<.001
Ever victimized	0.315	0.354	0.285	0.253	.009
All victimizations	0.526	0.670	0.397	0.343	<.001
Days incarcerated	79.3	88.4	62.7	88.7	.009
Involved in a violent gun crime	0.159	0.174	0.148	0.135	.335

Notes. The top panel shows baseline characteristics. Arrest and victimization measures include all available CPD data from 1999 (2010 for shooting victimizations) through the time of randomization, with counts winsorized at the top 99th percentile. “Predicted involvement in a violent gun crime” shows the risk score; the predicted probability of being a victim or an arrestee in a violent crime involving a gun during the next 18 months. “Involved in a violent gun crime” shows the actual realized rate of that same outcome over the 18 months after randomization for the control group. With the exception of that 18-month outcome period, the rest of the bottom panel shows control means during the 20-month outcome period. *p*-values are from a joint test of the null that referral pathway means are equal using heteroskedasticity-robust standard errors. Other serious violent-crime arrests include criminal sexual assault, robbery, and nonshooting aggravated assault and battery. *N* for the risk score is 2,225 for all, 1,232 for algorithm referrals, 717 for outreach referrals, and 276 for re-entry referrals.

violence incidents was 25% to 119% larger, depending on the measure, with a predicted level of future gun violence involvement of 14% versus 9%. Algorithm referrals also had about twice the number of prior reported victimizations. The only measure on which algorithm referrals looked less involved in crime is on days incarcerated in the past 30 months (129 versus 155 days), possibly because being incarcerated during the baseline period could have reduced the number of recently observed incidents and thus the potential risk score. In general, re-entry referrals were in between algorithm and outreach referrals, though they were about a year older at baseline than the latter. These observable differences across referrals suggest that the pathways identified different kinds of people.²⁵

Because a core question about the referral mechanisms was whether they could anticipate future risk, the bottom panel of Table II shifts from baseline characteristics to realized risk in the control group over the 20-month outcome period. It shows that READI's referral pathways successfully identified men at high risk of future arrest and victimization. Almost two-thirds of the control group were arrested during the 20 months after randomization, 1.7 times on average. About one-third reported at least one victimization. The best measured and most severe indication of gun violence involvement—being shot or killed—is shockingly high: there were 11 shooting and homicide victimizations per 100 control group members during the 20-month follow-up period. This is 52 times higher than the rate among average Chicagoans (0.21 per 100), and 2.7 times higher than the rate among other men 18–34 living in the same neighborhoods READI serves (4.1 per 100).²⁶ In short, READI's referral pathways identified a group at immensely high risk of being shot or killed.

The algorithm was also successful at predicting a broader measure of gun violence. It was trained to predict involvement in a violent gun crime as an arrestee or a victim in the next 18 months (Table II's last row). About 16% were actually arrested

25. Consistent with this is the fact that referrals from one pathway rarely included people who had previously been referred via another pathway; for example, only 35 initial outreach referrals had already been randomized via the algorithm pathway at the time of referral.

26. Rates are calculated based on shootings and homicides in Chicago from READI's launch (August 2017) through the end of the last study member's 20-month outcome period (November 2021) and population data from the American Community Survey.

for or reported being victims of a violent gun crime in that period, higher than the 11.4% predicted at baseline. This could partly be due to rising violent crime rates citywide during the outcome period relative to the data on which the algorithm was trained, and partly from the role of unobservables discussed below.

On general measures of criminal legal involvement such as overall arrests and victimizations, the realized risk levels of referrals from the three pathways mirror the significant differences in their baseline characteristics (top panel of [Table II](#)). Algorithm referrals were more likely to be arrested and victimized on both the extensive and intensive margins than outreach referrals, who in turn had higher rates and counts than re-entry referrals.

Yet, realized rates of gun violence involvement were similar across the pathways in the control group despite the differing risk levels predicted by observables. This suggests that staff using on-the-ground knowledge were leveraging unobservables that predict gun violence ($\hat{Y}(0)$). A key question, given the similarity of gun violence involvement across pathways, is whether the human decision makers were also selecting on expected gains from program participation (β), and not just levels of the outcome. We address this in [Section V.C](#).

IV.C. Take-up and Participation

Given the extraordinary risk of violence in the study population, their interest in participating in a program with rules and restrictions was not self-evident. [Table III](#) reports rates of participation among men in the treatment group, overall and by pathway. Among all men randomized to READI offers, 55% started the program (defined as attending orientation). This take-up rate is comparable to interventions working with much less disconnected populations, such as teenage boys still attending Chicago public high schools in similar neighborhoods ([Heller et al. 2017](#)).

As expected, take-up was highest (78%) among outreach referrals. These men often already knew some program staff and had been screened on an interest in, and “readiness” for, programming (see [Online Appendix A.8.5](#)). Recruiting re-entry referrals was also relatively successful, with take-up of 60%. Fewer of the algorithm referrals actually participated (37%). Based on staff interviews and field observations, a primary reason is that outreach

TABLE III
PROGRAM TAKE-UP RATES AND PARTICIPATION CONDITIONAL ON TAKE-UP, BY PATHWAY

	Take-up rate (%)	Conditional on take-up						Total earnings (\$) Average
		Work hours			CBT/training hours		Total hours Average	
		25th Percentile	50th Percentile	75th Percentile	Average	Average		
All participants	55	149	437	868	560	159	719	9,651
Algorithm	37	91	388	830	517	142	659	8,838
Outreach	78	196	502	942	615	172	787	10,465
Re-entry	60	128	438	792	506	157	663	9,183

Notes. Take-up is defined as attending the first day of READI orientation. Hours and earnings are averages over the entire 20-month outcome period (including after the start of COVID-19) among those who took up. The maximum number of hours someone could have participated during the 20-month outcome period depends on their time from randomization to take-up; for the pre-COVID period, see [Online Appendix Table A.II](#). Work hours correspond to time spent at a worksite. CBT/training hours correspond to time spent in group CBT sessions, professional development sessions, and online trainings. Hours and earnings are limited to men who took up and appear in the payroll data, which excludes 124 men who either took up prior to consent forms allowing the release of their payroll data being distributed or who attended orientation but did not start work within 20 months of randomization. Participation records for CBT and training before April 2020 are incomplete; hours shown are extrapolated based on available data. For additional details on hours and earnings data, see [Online Appendix A.2.4](#).

workers were unable to locate many algorithm referrals. Unfortunately, the outreach organizations did not keep formal and complete records, so we do not know exactly how many were unreachable. Once found, algorithm referrals were sometimes less trustful of outreach workers, or had other reasons to decline participation (such as already being employed, or now living in a distant neighborhood; see [Online Appendix A.3.1](#)).

The rest of [Table III](#) breaks out hours by activity and total earnings conditional on take-up, reflecting participation in the 20-month outcome period. The first few columns report the distribution of work hours from the payroll data from Heartland Alliance (the employer of record for READI participants), which shows a positive skew in participation (mean hours worked, 560 hours, are greater than median hours worked, 437 hours). Due to incompleteness in the CBT and training attendance data requiring us to use extrapolation and to sum across periods, we can only report average nonwork participation (159 hours). Our estimates suggest that participants earned an average of about \$9,650.

Because people took up the program after different lengths of time, and because in-person work paused in March 2020, it is difficult to read retention rates directly from [Table III](#). [Online Appendix Table A.II](#) reports the average maximum possible hours worked for each group during the 20-month outcome period pre-COVID, based on how long it took them to begin participating. Some of the differences in hours across pathways in [Table III](#) reflect differences in time to take-up; in fact, prior to the pandemic, participants from all three pathways worked between 28% and 31% of their possible post-take-up hours during the outcome period. [Online Appendix Table A.III](#) details participation during and after the pandemic.

[Figure II](#) reports two measures of job retention, weekly from the time participants first attended orientation. The first, shown by the solid line, is the proportion of participants who worked at least one day after the time noted on the x-axis, as measured through payroll data. The top panel shows that in the first few weeks after orientation, roughly 15% of participants stopped returning to work. Afterward, the decline in participation becomes roughly linear and relatively slow. About 75% of participants

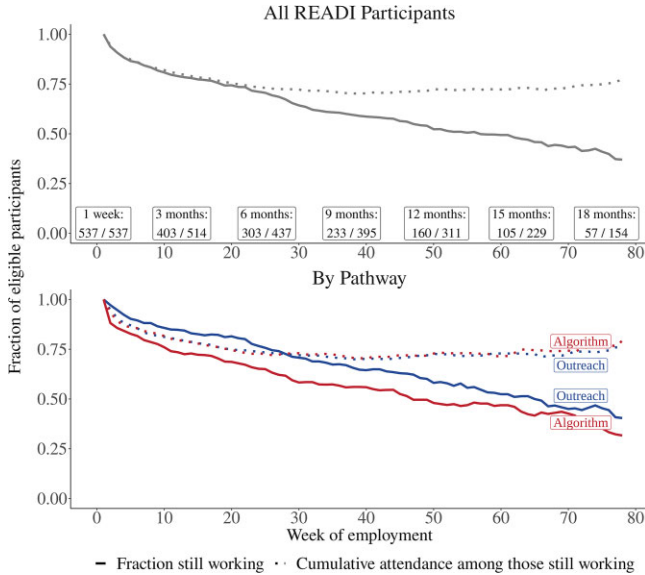


FIGURE II

READI Job Retention, Overall and by Pathway

Figure shows two measures of job retention for men who started READI measured from payroll data. The solid line shows the proportion of participants who work at least once after the time shown on the x-axis conditional on observing them for that long. The boxes show the number of workers contributing to each point. The dotted line shows the average proportion of possible weeks worked among those still working at each point in time. At 18 months after first taking up, $N = 19/60$ algorithm referrals and $N = 38/94$ outreach referrals are still observed working. Since COVID-19 changed what “participation” looked like in practice and payroll policies, we report these measures of retention using only data through the start of the pandemic. Because there were too few re-entry participants with sufficient pre-COVID data to measure retention, they are omitted. For a description of the pandemic’s effect on READI, see [Section II.C](#) and [Online Appendix A.5.3](#). For retention measures inclusive of the pandemic period, see [Online Appendix Figure A.III](#).

continued showing up to work after the 20-week mark, and a little over half continued to work after one year.²⁷

The second measure, shown by the dotted line, is the proportion of weeks worked by participants who are still working.

27. [Figure II](#) excludes the period starting in March 2020 when READI’s in-person employment programming was suspended for several months. A version of the plot including this period shows similar patterns ([Online Appendix Figure A.III](#)), but with slightly lower participation toward the end of the program period.

Unlike the first measure, which captures the extensive margin (whether people still work), this second measure captures the intensive margin (how much people work). After the initial fall-off, active participants consistently worked about 75% of the weeks that they could have.

The bottom panel of [Figure II](#) shows retention by pathway. (This excludes re-entry since the figure is limited to pre-COVID data and too few re-entry participants started this early.) The overall patterns are quite similar for algorithm and outreach participants, with the key difference being a faster fall-off in early participation among algorithm participants. This resulted in the larger dosage for outreach participants shown in [Table III](#).

V. RESULTS

V.A. *Average Treatment Effects*

[Table IV](#) reports estimates of average treatment effects on our prespecified primary outcome index and its components. The point estimate for being randomly assigned a READI offer (the ITT) is a 0.027 standard deviation reduction in the serious violence index (a 0.049 standard deviation reduction for the effect of participating, the TOT), but the result is not statistically significant ($p = .26$).

The second panel shows results on the three components of the index. The two measures of shooting and homicide involvement—arrests and victimizations—have negative and substantively large point estimates. There were 1.3 fewer shooting and homicide victimizations for every 100 participants, a 12% decline relative to the CCM. But the confidence interval is too wide to rule out a similarly sized increase instead. The decline in arrests is a proportionally huge 65% (2.2 fewer per 100 participants). It is statistically significant on its own, but not after adjusting for the three hypothesis tests across components (unadjusted $p = .05$, adjusted $p = .13$).²⁸

The small reduction in the overall index reflects the fact that not all forms of violence move in the same direction. There were 0.6 additional arrests for other serious violent crimes for every 100 participants, an 11% increase, though the standard errors are

28. We focus on reporting adjusted p -values that strongly control the FWER in the text, which is a conservative inference adjustment. The tables also report the less conservative q -values from controlling the FDR.

TABLE IV
READI'S ESTIMATED EFFECTS ON SERIOUS VIOLENCE INVOLVEMENT

	Estimates				p-values		
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR-q
Primary index of serious violence	0	-0.0266 (0.0234)	0.0155	-0.0487 (0.0413)	.257		
Primary outcome components, counts							
Shooting & homicide victimizations	0.1104	-0.0072 (0.0134)	0.1105	-0.0132 (0.0237)	.593	.833	0.733
Shooting & homicide arrests	0.0268	-0.0120 (0.0060)	0.0340	-0.0220 (0.0106)	.045	.126	0.134
Other serious violent-crime arrests	0.0544	0.0034 (0.0101)	0.0551	0.0063 (0.0177)	.733	.728	0.733

Notes. $N = 2,456$. Primary index standardizes the three components shown in the bottom panel based on the control group's distribution and takes their unweighted average. Other serious violent-crime arrests include criminal sexual assault, robbery, and nonshooting aggravated assault and battery. Multiple-hypothesis testing adjustments define the three components of the primary index as a family. FWER p -values control for the family-wise error rate using the Westfall and Young (1993) method for step-down resampling. FDR q -values control for the false discovery rate using the Benjamini and Hochberg (1995) method. CM shows control means; ITT shows intent-to-treat estimates; CCM shows control complier means; and TOT shows treatment-on-the-treated estimates from an IV regression. Regressions include baseline covariates and randomization strata fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses.

large relative to the point estimate.²⁹ As we show in [Section VI](#), however, the disproportionately high costs of shootings and homicides means that weighting arrests and victimizations by their social cost leads to different conclusions, which is relevant for policy makers to consider along with the lack of clear change in our primary outcome.

These results remain similar when using: a count model; different (or no) baseline covariates; randomization inference; adjustments for the possibility that incarceration or death are generating censoring (with some indication of a decline in the rate of shooting and homicide victimization that is masked by the additional days that treatment group men are alive); and the pre-COVID period only (see [Online Appendix A.5](#)). We also find that there are few statistically significant changes in other types of arrests, victimizations, or incarceration outcomes. Last, [Online Appendix A.5.5](#) shows how treatment effects accrue over time.

A natural question is whether the large decline in shooting and homicide arrests is a direct result of incapacitation from the program itself—whether keeping people busy during the workday mechanically reduced violence during that time. [Online Appendix A.5.6](#) reports estimated effects on arrests and victimizations separately by day and time, as measured by the time of incident (not the time of arrest). Although point estimates on the total number of arrests and victimizations are negative and substantively large during work days, the fall in incidents underlying shooting and homicide arrests happens during weekends, suggesting that incapacitation does not seem to be driving the change.³⁰ This is corroborated by interview data, in which participants report changing with whom and where they spend time outside of READI hours (see [Online Appendix A.8.6](#)).

V.B. Heterogeneity Analysis

Here we examine heterogeneity of effects by prespecified subgroups. Because they were built into the experimental design and

29. Because the index components do not all move in the same direction, using the first component of a principal component analysis affords a bit more precision than an unweighted index: ITT = -0.0569 , std. err. = 0.0377 , $p = .13$. Because this approach was not prespecified, we do not emphasize these findings.

30. [Online Appendix A.5.7](#) shows that there may be some role for incapacitation in reducing arrests for drug crimes.

prediction process, we focus on differences across referral pathways and predicted risk levels.³¹ Our preanalysis plan noted that these tests would be exploratory, as we did not anticipate being powered to detect moderate heterogeneity. Nonetheless, we still adjust our inferences for multiple testing.

Table V shows significant heterogeneity across referral pathways: we can reject the null that the estimated effects on the primary outcome index are the same across pathways ($p = .03$). There is a clear, statistically significant decline in serious violence in the outreach pathway. The decline in the index of 0.13 standard deviations for outreach participants remains statistically significant after adjusting inference for the three tests across pathways (adjusted $p = .03$). Breaking the index into components shows a similar but more precise pattern as in the overall results: large declines in both arrests (79%) and victimizations (43%) for shootings and homicides, which remain statistically significant after adjusting for the three tests across outcomes (adjusted $p = .03$ and $.08$, respectively), with a small decline (but large standard error) in other serious violent-crime arrests.

No results in the other referral pathways approach statistical significance. In many cases this is because the tests are underpowered, not because point estimates are substantively small. Standard errors for the TOT in the algorithm pathway are particularly large, despite the larger sample there, because of the weaker first stage resulting from lower take-up.

Table VI shows that pathway differences are not driven by differences in baseline predicted risk. The table reports separate effects for those over and under median predicted risk, as well as the 231 people with missing scores.³² There are some indications that READI's effects may be concentrated among those with higher predicted risk: the point estimate on the primary

31. Online Appendix A.6 summarizes heterogeneity by two other prespecified subgroups: site and age. Briefly, we find no evidence of heterogeneity by age, but there is some variation by site: significantly larger declines in two sites (Austin/West Garfield Park and Greater Englewood) but a positive point estimate in the third (North Lawndale).

32. Missing a risk score (which can only happen for outreach and re-entry referrals) indicates that someone had too little recent police contact to meet the criteria for inclusion in the algorithm: at least one arrest or two victimizations in the past 50 months. Consistent with the idea that less recent police contact indicates lower risk, the control means for the missing group are considerably lower.

TABLE V
READI'S ESTIMATED EFFECTS ON SERIOUS VIOLENCE INVOLVEMENT, BY PATHWAY

	Estimates			<i>p</i> -values			
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR- <i>q</i>
Primary index of serious violence by pathway							
Algorithm (<i>N</i> = 1,232)	−0.0097	0.0360 (0.0346)	−0.0640	0.0990 (0.0917)	.298	.529	0.312
Outreach (<i>N</i> = 878)	0.0085	−0.1007 (0.0379)	0.0563	−0.1293 (0.0469)	.008	.029	0.024
Re-entry (<i>N</i> = 346)	0.0126	−0.0600 (0.0594)	0.0501	−0.1003 (0.0958)	.312	.529	0.312
Primary outcome components by pathway, counts							
Algorithm							
Shooting & homicide victimizations	0.1055	0.0231 (0.0197)	0.0654	0.0635 (0.0524)	.243	.552	0.415
Shooting & homicide arrests	0.0227	−0.0048 (0.0085)	0.0222	−0.0133 (0.0226)	.572	.571	0.572
Other serious violent-crime arrests	0.0568	0.0169 (0.0155)	0.0468	0.0465 (0.0411)	.277	.552	0.415

TABLE V
CONTINUED

	Estimates			p-values			
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR- <i>q</i>
Outreach							
Shooting & homicide victimizations	0.1142	−0.0438 (0.0214)	0.1320	−0.0562 (0.0264)	.040	.080	0.061
Shooting & homicide arrests	0.0320	−0.0255 (0.0098)	0.0415	−0.0328 (0.0121)	.009	.025	0.028
Other serious violent-crime arrests	0.0502	−0.0045 (0.0165)	0.0583	−0.0059 (0.0203)	.783	.792	0.783
Re-entry							
Shooting & homicide victimizations	0.1180	−0.0214 (0.0355)	0.1358	−0.0358 (0.0569)	.546	.823	0.819
Shooting & homicide arrests	0.0281	−0.0028 (0.0171)	0.0346	−0.0046 (0.0274)	.869	.883	0.869
Other serious violent-crime arrests	0.0562	−0.0246 (0.0194)	0.0612	−0.0412 (0.0312)	.204	.562	0.613

Notes. $N = 2,456$. Estimates for each outcome are from a single regression that interacts pathway indicators with treatment. Primary index standardizes the three components shown in the bottom panel based on the control group's distribution and takes their unweighted average. Other serious violent-crime arrests include criminal sexual assault, robbery, and nonshooting aggravated assault and battery. For multiple-hypothesis testing adjustments in the primary index, we define the three different pathways as a family. For the component adjustments, we define the three different outcomes in each pathway as a family. FWER p -values control for the family-wise error rate using the [Westfall and Young \(1993\)](#) method for step-down resampling. FDR q -values control for the false discovery rate using the [Benjamini and Hochberg \(1995\)](#) method. CM shows control means; ITT shows intent-to-treat estimates; CCM shows control complier means; and TOT shows treatment-on-the-treated estimates from an IV regression. Regressions include baseline covariates and randomization strata fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. F -tests of the null hypothesis that treatment effects are equal across the three pathways are as follows: primary index: $p = .028$; shooting and homicide victimizations: $p = .069$; shootings and homicide arrests: $p = .236$; other serious violent-crime arrests: $p = .235$.

TABLE VI
READI'S ESTIMATED EFFECTS ON SERIOUS VIOLENCE INVOLVEMENT, BY ALGORITHMICALLY PREDICTED RISK

	Estimates			<i>p</i> -values			
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR- <i>q</i>
Primary index of serious violence by predicted risk							
Over median (<i>N</i> = 1,112)	0.0850	−0.0669 (0.0405)	0.1717	−0.1482 (0.0866)	.099	.259	0.298
Under median (<i>N</i> = 1,113)	−0.0511	0.0100 (0.0321)	−0.0671	0.0169 (0.0510)	.754	.942	0.852
Missing (<i>N</i> = 231)	−0.1489	−0.0092 (0.0492)	−0.1482	−0.0128 (0.0697)	.852	.942	0.852
Primary outcome components by predicted risk, counts							
Over median							
Shooting & homicide victimizations	0.1434	−0.0193 (0.0223)	0.1573	−0.0427 (0.0475)	.388	.619	0.582
Shooting & homicide arrests	0.0376	−0.0206 (0.0103)	0.0575	−0.0456 (0.0219)	.045	.125	0.135
Other serious violent-crime arrests	0.0771	−0.0048 (0.0175)	0.1017	−0.0108 (0.0373)	.784	.781	0.784

TABLE VI
CONTINUED

	Estimates				p-values		
	CM	ITT	CCM	TOT	Observed ITT	FWER	FDR-q
Under Median							
Shooting & homicide victimizations	0.0906	0.0056 (0.0191)	0.0853	0.0093 (0.0303)	.770	.773	0.770
Shooting & homicide arrests	0.0203	-0.0069 (0.0083)	0.0227	-0.0112 (0.0132)	.410	.691	0.615
Other serious violent-crime arrests	0.0407	0.0135 (0.0137)	0.0292	0.0224 (0.0218)	.324	.691	0.615
Missing							
Shooting & homicide victimizations	0.0526	-0.0109 (0.0348)	0.0614	-0.0159 (0.0493)	.754	.971	0.754
Shooting & homicide arrests	0.0075	0.0048 (0.0133)	0.0076	0.0075 (0.0188)	.715	.971	0.754
Other serious violent-crime arrests	0.0150	-0.0060 (0.0154)	0.0090	-0.0090 (0.0218)	.698	.971	0.754

Notes. $N = 2,456$. Estimates for each outcome are from a single regression that interacts indicators for above/below-median risk score and missing risk score with treatment. The median risk score in the study sample, the predicted probability at baseline of being arrested for or the victim of a violent gun crime in the next 18 months, is 0.11. Primary index standardizes the three components shown in the bottom panel based on the control group's distribution and takes their unweighted average. Other serious violent-crime arrests include criminal sexual assault, robbery, and nonshooting aggravated assault and battery. For multiple-hypothesis testing adjustments in the primary index, we define the three different risk score groups as a family. For the component adjustments, we define the three different outcomes in each risk score group as a family. FWER p -values control for the FWER using the Westfall and Young (1993) method for step-down resampling. FDR- q values control for the false discovery rate using the Benjamini and Hochberg (1995) method. CM shows control means; ITT shows intent-to-treat estimates; CCM shows control complier means; and TOT shows treatment-on-the-treated estimates from an IV regression. Regressions include baseline covariates and randomization strata fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. F -tests of the null hypothesis that treatment effects are equal across the three risk score groups are as follows: primary index: $p = .330$; shooting and homicide victimizations: $p = .700$; shootings and homicide arrests: $p = .299$; other serious violent-crime arrests: $p = .569$.

index is large and negative only for that group, and the only statistically significant estimate among the components is the decline in shooting and homicide arrests for those at above-median predicted risk. Overall, we cannot reject the null that the three predicted risk groups have the same effect ($p = .33$), and the confidence interval is wide enough to include large proportional declines from a lower baseline for the below-median predicted risk group.

V.C. Interpreting Pathway Heterogeneity

Motivated by the patterns in the previous section, we explore outcomes by pathway and predicted risk, focusing on how the outreach workers' selection process or role in participant engagement might explain the observed treatment heterogeneity. [Figure III](#) presents all the results we discuss in this section. It shows outcomes by pathway and baseline risk score quartile. The quartiles are defined across the full distribution of risk scores in the study sample, so average predicted risk in each quartile is very similar across pathways.³³

The top panel of [Figure III](#) shows ITT estimates for the primary outcome index (see [Online Appendix Figure A.VI](#) for components). For outreach referrals with below-median or missing predicted risk, the point estimates are very close to zero. The large estimated declines in serious violence for outreach referrals in [Table V](#) are driven by the smaller group of above-median predicted risk outreach referrals. Yet there do not appear to be parallel declines among the above-median predicted risk algorithm referrals. Thus, there seems to be an interaction between the predictable part of the risk of gun violence involvement and something outreach workers are doing.

The two bottom panels of [Figure III](#) help rule out several possibilities about how outreach workers matter. The bottom left panel shows average rates of realized gun violence involvement (the outcome predicted by the algorithm) by pathway and risk quartile for the control group (i.e., $\bar{Y}(0)$). The bottom right panel shows take-up and hours worked for the treatment group. Marker

33. For distributions of risk scores within referral pathway, see the top left panel of [Online Appendix Figure A.VI](#). While we describe referrals in quartiles 1 and 2 as having “lower” predicted risk, in absolute terms it is still quite high: almost 7% of this group was expected to be arrested for or the victim of a violent gun crime within 18 months, compared to 11% for the full sample.

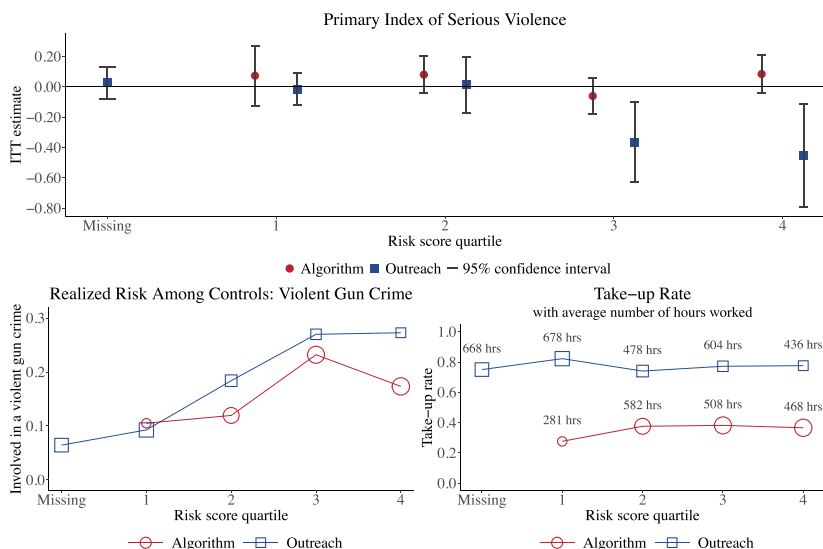


FIGURE III

READI Estimated Effects, Realized Risk, and Take-up by Pathway and Algorithmically Predicted Risk

The top panel shows coefficient estimates and 95% confidence intervals (using heteroskedasticity-robust standard errors) on three-way interactions of pathway indicators, risk score quartile indicators, and an indicator for being randomized to receive a READI offer, from a regression of the primary index on baseline covariates, randomization strata fixed effects, and all two-way interactions of pathway indicators, risk score quartile indicators, and an indicator for being randomized to receive a READI offer. The bottom left panel shows the realized rate of involvement in a violent gun crime as a victim or an arrestee during the 18 months after randomization by quartile of the risk score, which is the predicted probability at baseline of the same outcome, separately for algorithm and outreach referrals. The bottom right panel shows the take-up rate, defined as the share of the treatment group attending the first day of READI orientation, by quartile of the risk score, separately for algorithm and outreach referrals. Above the markers are the average hours worked during the 20-month outcome period conditional on take-up. Markers in both of the bottom panels are weighted to reflect the share of algorithm and outreach referrals in each risk score quartile. $N = 231$ for the missing risk score group (of whom 161 are outreach referrals). For ITT estimates of each individual index component by pathway and risk score quartile, see [Online Appendix Figure A.VI](#).

sizes correspond to the share of referrals in each pathway in each quartile.

The marker sizes in the bottom panels show that outreach workers did not prioritize the highest predicted risk referrals: most outreach referrals (75%) had below-median or missing predicted risk, while most algorithm referrals (67%) had

above-median predicted risk. The fact that most outreach referrals had lower predicted risk—not the subgroup most responsive to READI—suggests that outreach workers were not consistently successful at selecting on gains, even if they were trying to do so.

The bottom panels suggest two other factors with a stronger relationship to outreach referral decisions: the unobservables that contributed to a high risk of gun violence and the proclivity to take up the program. Outreach workers' success at identifying unobservable risk factors is shown in the bottom left panel. Although we know from [Table II](#) that both pathways referred men whose realized risk was greater than their predicted risk, outreach referrals typically had higher realized risk than algorithm referrals at similar predicted risk levels. The gap between the pathways suggests that outreach workers selected men partly based on information unobservable to the algorithm that is predictive of high $Y(0)$. The bottom right panel shows that outreach referrals were about 40 percentage points more likely to start READI, regardless of predicted risk. This pattern is consistent with outreach workers referring men partly based on the determinants of take-up, potentially including the ability to find the referrals in the first place.³⁴

Importantly, neither factor appears to explain the pattern of treatment heterogeneity. The gap in realized risk between algorithm and outreach referrals is similar across quartiles 2 and 3, and if anything, slightly larger in quartile 2. But the estimated treatment effect is much larger for outreach referrals in quartile 3. Although this is not a strong test of the role of unobservable determinants of $Y(0)$, the pattern seems inconsistent with the idea that outreach workers' use of such unobservables in their referral decision is the key driver of the larger effects among that

34. This possibility was also highlighted in the qualitative data, where staff frequently discussed selecting on “readiness for READI” (see [Online Appendix A.8.5](#)). Examples of not being ready offered by outreach workers included: men who were still focused on settling a score, men who preferred illegal work (such as drug dealing), men who feared for their safety if they left their block or associated with opposition members, men who could not overcome their impatience with or skepticism of the therapy and jobs, and men unwilling to attend the program sober. While the concentration of outreach referrals in the less responsive, lower predicted risk quartiles suggests that being “ready for READI” does not necessarily mean ready for change, it could still mean ready to take up the program.

pathway.³⁵ Moreover, take-up rates are fairly flat across the risk quartiles.³⁶ Because treatment effects are concentrated in the top risk quartiles, this argues against selection on take-up being the sole driver of pathway treatment heterogeneity.

Another possibility is that variation in treatment dosage drives treatment heterogeneity, and dosage is positively correlated with predicted risk. To assess this possibility, the numbers above each marker in the bottom right panel show participation decisions on the intensive margin (hours worked during the 20-month outcome period conditional on participating). Interestingly, while outreach referrals were more likely to come from the lower predicted risk quartiles that had more hours worked, outreach participants did not consistently work more hours after starting than did algorithm participants in the same quartile.³⁷ Nor did outreach participants in the top two risk quartiles consistently work more hours. So it does not appear that the intensive margin of participation explains the patterns of treatment heterogeneity on its own, either.

[Online Appendix A.6.5](#) provides some evidence that the unobservables correlated with pathway, rather than dosage itself, matter for treatment heterogeneity.³⁸ The fact that algorithm participants often worked as much as outreach participants in the same

35. Another possibility is that outreach workers are selecting on observables, but in a different way than is captured in the algorithmic risk score. [Online Appendix A.6.2](#) and [A.6.3](#) provide evidence that this is not the case. Baseline covariates only weakly predict whom the outreach workers will select ([Online Appendix Table A.XIV](#)), and adjusting algorithm referrals' observables to look like those of outreach referrals does little to close the gap in estimated treatment effects between these groups ([Online Appendix Table A.XV](#)).

36. Note that relatively flat take-up rates across quartiles in each pathway also mean that the patterns for the TOT estimates are similar to the ITTs we show in the top panel.

37. [Online Appendix A.6.4](#) expands on this pattern by showing that pathway explains more of the variation in take-up than the other observables, but little of the variation in hours worked beyond what observables explain.

38. In [Online Appendix A.6.5](#), we adapt the [Abadie, Chingos, and West \(2018\)](#) endogenous stratification exercise to predict dosage among controls, then estimate heterogeneous treatment effects for low- versus high-predicted dosage groups. We find suggestive evidence that treatment effects are larger for the high-predicted dosage group. But when we do the stratification exercise without pathway as a dosage predictor, we do not find the same pattern. This is further support for the idea that the unobservables correlated with pathway rather than dosage itself matter for treatment heterogeneity. It may be that the kinds of people who select into high doses of READI also respond more, but for reasons other than a direct dose-response relationship.

risk quartile is also suggestive evidence that outreach workers did not just generate a more supportive environment for their own referrals, or work harder to keep them in the program. Of course, hours worked is not the only measure of dosage, so we cannot rule out differential dosage across pathways. But we also found little evidence in our interviews and focus groups that outreach workers treated participants differently based on their pathway after take-up.

In the end, although we have some evidence for what outreach workers used in their selection process (unobserved predictors of $Y(0)$ and take-up), we can rule out that these factors alone explain the larger responsiveness among outreach referrals. That leaves us without a clear explanation for the patterns of treatment effects by pathway and predicted risk of gun violence involvement. It is important for future research to further explore why this interaction occurs, as different possibilities have different implications for policy.

For example, it is natural to look at the pattern in the top panel of [Figure III](#) and conclude that while outreach workers are essential, screening out referrals with lower predicted risk might be beneficial. This would be true if those with lower $\hat{Y}(0)$ are genuinely less responsive, including the possibility of a floor effect with too little room for decreases in violence ([Heller et al. 2022](#)). If so, continuing to have outreach workers make referrals, or finding other ways to identify people with the relevant unobservables, would remain crucial for reducing violence. Continuing to make referrals solely from an algorithm trained to predict $\hat{Y}(0)$ would be ineffective. The combination of outreach screening and the algorithm would still be important to limit outreach referrals to those most responsive and to bring those whom outreach workers might not have otherwise considered into the screening process.

Two possibilities might change those policy implications. First, it is possible that READI staff had an easier time finding alternative programming for control group men with lower predicted risk because outreach workers already knew them—even though, as discussed earlier, programming comparable to READI was seldom available. If so, then the apparent lack of responsiveness could be due to an improved counterfactual rather than a lower potential gain from READI. If access to counterfactual programming matters, then screening out people with lower predicted risk would only be productive in settings with enough access to outside programming. Second, outreach workers could have engaged

more successfully with their own referrals or generated a more productive environment for them, but in a way that did not consistently result in outreach participants working more hours than algorithm participants in the same risk quartile. The role of ongoing interaction and relationships with caseworkers seems to play a role in program success in other settings, such as sectoral employment training programs (see [Katz et al. 2022](#)). If such personal engagement is key, it might not be necessary or desirable to exclude higher predicted risk algorithm referrals from programming: efforts could instead focus on improving relationships and increasing engagement with such referrals to raise their responsiveness to treatment.

We emphasize that the findings in this section are exploratory: our sample is too small to determine whether all of these subgroup results differ from each other significantly. On their face, they suggest that the combination of human and algorithmic screening used by READI seemed to outperform either referral mechanism individually, which may provide some guidance for future CVI design and for research. Future work confirming this heterogeneity, and unpacking the reasons for it, would be useful.

VI. SOCIAL BENEFITS AND COSTS

In many settings, the statistical uncertainty about declines in primary outcome components, combined with the lack of improvements in other secondary outcomes, might lead us to question the value of the intervention. But there is a crucial difference about this setting: the outcomes of interest are literally measures of life and death. Shootings and homicides generate such enormous costs to individuals, families, and communities that even limited improvement on a small number of outcomes could generate large benefits to society. Weighting outcome measures by their social value can help inform how to think about the value of the program and can provide a useful alternative to multiple-testing adjustments when policy makers value changes in the different outcomes differently ([Viviano, Wuthrich, and Niehaus 2021](#)).

Here we try to quantify READI's benefits and then compare them with program costs. Our main focus is on the social costs surrounding crime and violence. Our calculations are necessarily a rough approximation because of the uncertainty involved with assigning costs to crime and the loss of life

(Domínguez and Raphael 2015), and because we ignore a range of other difficult-to-measure costs and benefits (e.g., the social value of investing and generating jobs in historically underserved neighborhoods, the value of the work that READI participants do, gains in unmeasured outcomes like mental or physical health, and the opportunity cost of program spending).³⁹ Our intention is not to provide a definitive cost-benefit analysis but to provide a basic sense of how the substantive importance of the outcomes READI affects might shape how we think about the program's effects.

Because of uncertainty about the social costs of crime (including the statistical value of life), and because there are numerous choices about how to translate imperfect measures like arrests into assessments of the true amount of crime, Table VII presents sets of less and more inclusive estimates for how READI affects social costs. Since there is only one outcome and the analysis was prespecified, we do not adjust for multiple testing.

We assign to each arrest and victimization of a READI study member an estimated cost of crime depending on its type. Online Appendix A.7 contains the details of these calculations. The less inclusive estimates use the lower-end estimates of the cost of crime from Cohen and Piquero (2009), and only count the harm from crime when directly observed in the data (i.e., from each observed arrest or victimization of a READI study participant). The more inclusive estimates use the higher willingness-to-pay estimates of the cost of crime (Cohen and Piquero 2009); extrapolate how much crime is likely to have occurred for every arrest, given clearance rates; and scale each victimization up by average reporting rates for each crime type.⁴⁰ When we compare the benefits of reduced crime with the administrative costs of the program, we treat program wages as a transfer from society to participants. The net READI costs are thus the administrative costs less payments to READI participants.

39. Because READI was largely funded by private philanthropy, there was no deadweight loss from raising tax dollars. Future versions of the program might use public funding, and so might benefit from weighing the marginal value of public funds (Hendren and Sprung-Keyser 2020).

40. In particular, for each crime type's clearance rate C , we assume there are $\frac{1}{C}$ crimes committed for each observed arrest, and C offenders arrested for every observed victimization. We use Chicago-specific clearance rates for homicides and clearance rates from Belfield et al. (2006) for other offenses.

TABLE VII
READI'S ESTIMATED EFFECT ON THE SOCIAL COSTS OF CRIME

	Less inclusive estimates			More inclusive estimates		
	CM	ITT	CCM	TOT	CM	ITT
Social cost of victimization						
READI sample victims	\$147,173	-\$22,741 (\$33,363)	\$149,680	-\$41,724 (\$58,848)	\$390,251	-\$56,719 (\$82,482)
READI sample offenders	\$127,058	-\$69,594** (\$29,140)	\$188,796	-\$127,686** (\$51,542)	\$1,061,717	-\$435,114** (\$201,407)
Social cost of punishment						
Legal system costs	\$10,776	-\$4,535** (\$1,916)	\$14,819	-\$8,320** (\$3,389)	\$14,679	-\$5,113** (\$2,065)
Productivity loss from incarceration	\$4,506	-\$2,106** (\$894)	\$6,369	-\$3,864** (\$1,582)	\$6,315	-\$2,376** (\$964)
Total social cost of crime	\$289,513	-\$98,976** (\$45,791)	\$359,665	-\$181,595** (\$80,911)	\$1,472,962	-\$499,321** (\$219,071)
Social cost of program						
Administrative costs		\$33,510		\$61,483		\$33,510
Transfer to participants		-\$5,260		-\$9,651		-\$5,260
Net READI costs		\$28,250		\$51,832		\$28,250
Benefit-cost ratio		3.5:1		3.5:1		17.7:1

Notes. $N = 2,456$. Outcome variables in the top panel are different measures of social costs from crime. Negative point estimates show gains from reduced crime. Crime cost estimates, from [Cohen and Piquero \(2009\)](#) and inflated to 2017 dollars, aim to quantify harm to victims from each crime, the productivity loss to offenders from their involvement in the legal system, and the cost to the government of running the legal system. The less inclusive estimates use bottom-up crime cost estimates and only assign costs to observed arrests and victimizations (ignoring both offending that does not result in arrest and arrests of nonstudy individuals who victimized READI participants). The more inclusive estimates use willingness-to-pay crime cost estimates and extrapolate social harm based on estimates of clearance rates (i.e., for each crime type's clearance rate, C , we assume there are $1/C$ crimes committed for each observed arrest and C arrests made for each observed victimization), as well as victimization reporting rates. Program costs are calculated from spending reported by Heartland Alliance over the 20-month outcome period. CM shows control means; ITT shows intent-to-treat estimates; CCM shows control complier means; and TOT shows treatment-on-the-treated estimates from an IV regression. Regressions include baseline covariates and randomization strata fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table VII reports program impacts on these costs. The first thing to note is how much cost society incurred from the level of crime and violence in READI's absence: between \$360,000 and \$1,870,000 on average, depending on how inclusive the costs are, for each control complier. The second thing to note is that despite the uncertainty about how each element of our primary index responds to READI, the severity of shootings and homicides is so great that their higher weights in the cost of crime calculations generate less uncertainty about program benefits. We estimate that READI saved society a total of between \$182,000 and \$916,000 for each participant ($p = .03$ and $p = .02$, respectively), about a 50% decline. Compared to the cost of offering READI over 20 months (about \$52,000 for each participant), these benefits are at least 4—and perhaps as much as 18—times as large as the program's costs.

VII. DISCUSSION

The READI study was designed to determine whether it is possible to find people at high enough risk of gun violence for a targeted intervention to reduce shootings sufficiently to outweigh program costs; whether they could be engaged; and whether the combination of jobs, CBT, and outreach could reduce their involvement in serious violence. The answer to the first two questions is a definitive yes. Each referral pathway found men at extraordinary risk of being shot, killed, or arrested for a serious violent crime. Moreover, most of these men were willing to engage, despite the risks and barriers they faced. The 55% take-up rate for READI was higher than the take-up rates among high school boys for the in-school program studied by Heller et al. (2017) and among former prisoners for shorter transitional jobs programs (e.g., Redcross et al. 2016). Our interviews with participants suggest that many were initially motivated by the paycheck but stayed due to how the combination of CBT-related skills and caring program staff improved their well-being.

READI's effects on serious violence, however, are mixed. There is no statistically significant decline in our prespecified primary outcome. This means we cannot conclude with certainty that the version of READI evaluated here decreased serious violence. But there is suggestive evidence that READI reduced arrests for shootings and homicides, with the estimated effect being just beyond traditional statistical significance cutoffs.

Moreover, for one subgroup—men referred by outreach workers—the declines in arrests and victimizations for shootings and homicides clearly pass standard statistical significance thresholds.

If program operators are limited to one referral pathway, then using only outreach workers to identify referrals may be the best way to maximize treatment effects. This is important for external validity, since jurisdictions will likely find it easier to implement an outreach referral pathway; many cities already have local organizations working with people at high risk of gun violence. In contrast, it may be more difficult to gain access to detailed police or other administrative data useful for predicting gun violence risk, find the expertise to train an algorithmic model, and convince local providers to trust the results.

Still, our exploratory analysis shows that only the outreach referrals with the highest algorithmic risk predictions respond to READI. So it is possible that a combination of human intelligence and machine-driven risk prediction may more effectively anticipate treatment responsiveness than either method alone. The fact that men with equally high risk of gun violence seem to respond quite differently depending on how they were selected for READI highlights the importance of future research on what explains the interaction between predicted risk and referral pathway: differences in selection by outreach workers, in program experiences by pathway, in responsivity by risk level, or in varying counterfactuals.

In considering external validity, several findings echo patterns in other violence prevention studies. Despite differences in setting, population, and treatment between READI, a CBT program serving former members of armed groups in Liberia, and summer jobs programs across several U.S. cities, both [Blattman et al. \(forthcoming\)](#) and [Heller \(2022\)](#) find larger effects for those facing the largest risk of negative outcomes. A completely different violence-prevention program, which involves information sharing between hospitals, local government, and police, finds a decline in the most lethal outcomes combined with an increase in less lethal violence, similar to our pattern of results ([Florence et al. 2011](#)). It will always be important to attend to context and diagnosis of the conditions driving violence. The blend of therapy and economic engagement seems to work in places like Chicago and Liberia, where a significant fraction of the violence arises from fragmented groups who are responding to perceived slights, injustices, and reputational concerns; it may work less well in

contexts where killings are more rooted in organized crime or instrumental violence (Blattman 2022). Still, commonalities such as variation in responsiveness by risk level (at least among some subgroups) and substitution from more to less lethal violence raise the possibility that similar kinds of behavior change are possible across settings.

Our estimates do not capture READI's complete impact on participants owing to several measurement challenges. For example, low clearance rates and underreporting of less severe victimizations results in there being many more reported shooting and homicide victimizations in our data than arrests for the far more common crimes of aggravated assault and robbery. Arrests outside of Chicago also remain unobserved. Administrative data will always yield an incomplete view of a program's impact. As one example, our qualitative data collection made clear that the investments in the community (the urban renewal work participants performed and the READI-induced hiring at local organizations) were valuable to those who took part in READI. At least some participants reported unmeasured effects on relationship quality, self-confidence, and community integration.

From a scientific perspective, the statistically insignificant decline in our primary outcome merits caution. Science is rightfully conservative about overturning null hypotheses based on a single imprecise finding. From a policy perspective, however, the binary conclusion from one hypothesis test is not dispositive; $p > .1$ is not the same as a true zero. As a number of economists have emphasized (Ziliak and McCloskey 2008; Manski 2019; Imbens 2021), policy makers should attend to the range of plausible magnitudes consistent with a point estimate, along with the availability of other credible evidence and the benefits and costs of the status quo versus new investments. In this vein, it is worth emphasizing that about 74% of the confidence interval for READI's estimated impact on the serious violence index is below zero; 87% of the confidence interval for shooting and homicide arrests is negative. Weighting arrests and victimizations by their social costs shows that READI's estimated benefits to society were at least four times its cost.

For policy makers trying to reduce the enormous harms of gun violence, and given the absence of other convincing evidence about how to reduce shootings without the collateral costs that often accompany aggressive law enforcement, the findings we

report here should still be useful inputs into decision making, despite the need for more evidence to generate firm scientific conclusions. READI is a demonstration that investments of this scale are possible, even among a population deeply disconnected from every institution other than the criminal legal system, with a potentially large social payoff given the concentration of socially costly outcomes among a relatively small group of people. Future research could usefully experiment with refinements to the program model, targeting strategies, and scaling challenges, and other ways to reduce the extraordinarily high risk of gun violence among a small and underserved group.

UNIVERSITY OF CHICAGO CRIME LAB AND EDUCATION LAB,
UNITED STATES

UNIVERSITY OF MICHIGAN AND NATIONAL BUREAU OF ECONOMIC
RESEARCH, UNITED STATES

CORNELL UNIVERSITY, UNITED STATES

BOOTH SCHOOL OF BUSINESS, UNIVERSITY OF CHICAGO AND
NATIONAL BUREAU OF ECONOMIC RESEARCH, UNITED STATES;
AND CENTRE FOR ECONOMIC POLICY RESEARCH, UNITED
KINGDOM

HARRIS SCHOOL OF PUBLIC POLICY, UNIVERSITY OF CHICAGO,
AND NATIONAL BUREAU OF ECONOMIC RESEARCH, UNITED
STATES

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at
The Quarterly Journal of Economics online.

DATA AVAILABILITY

The code underlying this article are available in the Harvard
Dataverse, <https://doi.org/10.7910/DVN/LHH29P> (Bhatt et al.
2023).

REFERENCES

- Abadie, Alberto, “Semiparametric Instrumental Variable Estimation of Treatment Response Models,” *Journal of Econometrics*, 113 (2003), 231–263. [https://doi.org/10.1016/S0304-4076\(02\)00201-4](https://doi.org/10.1016/S0304-4076(02)00201-4)
- Abadie, Alberto, Matthew M. Chingos, and Martin R. West, “Endogenous Stratification in Randomized Experiments,” *Review of Economics and Statistics*, 100 (2018), 567–580. https://doi.org/10.1162/rest_a_00732

- Abt, Thomas, *Bleeding Out: The Devastating Consequences of Urban Violence—and a Bold New Plan for Peace in the Streets* (London: Hachette, 2019).
- Anderson, Elijah, *Code of the Street: Decency, Violence, and the Moral Life of the Inner City* (New York: WW Norton, 1999).
- Anderson, Michael L., “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American Statistical Association*, 103 (2008), 1481–1495. <https://doi.org/10.1198/016214508000000841>
- Ang, Desmond, “The Effects of Police Violence on Inner-City Students,” *Quarterly Journal of Economics*, 136 (2021), 115–168. <https://doi.org/10.1093/qje/qjaa027>
- Arbour, William, “Can Recidivism Be Prevented from Behind Bars? Evidence From a Behavioral Program,” University of Montreal Working Paper, 2022.
- Aspholm, Roberto, *Views from the Streets: The Transformation of Gangs and Violence on Chicago’s South Side* (New York: Columbia University Press, 2020).
- Beck, Aaron T., *Cognitive Therapy of Depression* (New York: Guilford Press, 1979).
- Beck, Aaron T., and David J. A. Dozois, “Cognitive Therapy: Current Status and Future Directions,” *Annual Review of Medicine*, 62 (2011), 397–409. <https://doi.org/10.1146/annurev-med-052209-100032>
- Belfield, Clive R., Milagros Nores, Steve Barnett, and Lawrence Schweinhart, “The High/Scope Perry Preschool Program: Cost–Benefit Analysis Using Data from the Age-40 Followup,” *Journal of Human Resources*, 41 (2006), 162–190. <https://doi.org/10.3368/jhr.XLI.1.162>
- Benjamini, Yoav, and Yosef Hochberg, “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing,” *Journal of the Royal Statistical Society: Series B (Methodological)*, 57 (1995), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berk, Richard, Lawrence Sherman, Geoffrey Barnes, Ellen Kurtz, and Lindsay Ahlman, “Forecasting Murder within a Population of Probationers and Parolees: A High Stakes Application of Statistical Learning,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172 (2009), 191–211. <https://doi.org/10.1111/j.1467-985X.2008.00556.x>
- Bhatt, Monica P., Sara B. Heller, Max Kapustin, Marianne Bertrand, and Christopher Blattman, “Replication Data for: ‘Predicting and Preventing Gun Violence: An Experimental Evaluation of READI Chicago,’” (2023), Harvard Dataverse, <https://doi.org/10.7910/DVN/LHH29P>
- Blattman, Christopher, *Why We Fight: The Roots of War and the Paths to Peace* (New York: Viking, 2022).
- Blattman, Christopher, Sebastian Chaskel, Julian C. Jamison, and Margaret Sheridan, “Cognitive Behavior Therapy Reduces Crime and Violence Over 10 Years: Experimental Evidence,” *American Economic Review: Insights*, forthcoming.
- Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan, “Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia,” *American Economic Review*, 107 (2017), 1165–1206. <https://doi.org/10.1257/aer.20150503>
- Braga, Anthony A., “Serious Youth Gun Offenders and the Epidemic of Youth Violence in Boston,” *Journal of Quantitative Criminology*, 19 (2003), 33–54. <https://doi.org/10.1023/A:1022566628159>
- Braga, Anthony A., and Philip J. Cook, *Policing Gun Violence: Strategic Reforms for Controlling Our Most Pressing Crime Problem* (Oxford: Oxford University Press, 2023). <https://doi.org/10.1093/oso/9780199929283.001.0001>
- Braga, Anthony A., David M. Kennedy, Elin J. Waring, and Anne Morrison Piehl, “Problem-Oriented Policing, Deterrence, and Youth Violence: An Evaluation of Boston’s Operation Ceasefire,” *Journal of Research in Crime and Delinquency*, 38 (2001), 195–225. <https://doi.org/10.1177/0022427801038003001>
- Braga, Anthony A., Andrew V. Papachristos, and David M. Hureau, “The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and

- Meta-Analysis," *Justice Quarterly*, 31 (2014), 633–663. <https://doi.org/10.1080/07418825.2012.673632>
- Braga, Anthony A., David Weisburd, and Brandon Turchan, "Focused Deterrence Strategies and Crime Control: An Updated Systematic Review and Meta-Analysis of the Empirical Evidence," *Criminology & Public Policy*, 17 (2018), 205–250. <https://doi.org/10.1111/1745-9133.12353>
- Butts, Jeffrey A., Caterina Gouvis Roman, Lindsay Bostwick, and Jeremy R. Porter, "Cure Violence: A Public Health Model to Reduce Gun Violence," *Annual Review of Public Health*, 36 (2015), 39–53. <https://doi.org/10.1146/annurev-publhealth-031914-122509>
- Carr, Jillian, and Jennifer L. Doleac, "The Geography, Incidence, and Under-reporting of Gun Violence: New Evidence Using ShotSpotter Data," SSRN Working Paper, 2016. <https://doi.org/10.2139/ssrn.2770506>
- CDC, "WISQARS (Web-based Injury Statistics Query and Reporting System)—Injury Center," Centers for Disease Control and Prevention, 2020.
- Chalfin, Aaron, Benjamin Hansen, Emily K. Weisburd, and Morgan C. Williams Jr., "Police Force Size and Civilian Race," *American Economic Review: Insights*, 4 (2022), 139–158. <https://doi.org/10.1257/aeri.20200792>
- Chandler, Dana, Steven D. Levitt, and John A. List, "Predicting and Preventing Shootings among At-Risk Youth," *American Economic Review: Papers and Proceedings*, 101 (2011), 288–292. <https://doi.org/10.1257/aer.101.3.288>
- Cohen, Mark A., and Alex R. Piquero, "New Evidence on the Monetary Value of Saving a High Risk Youth," *Journal of Quantitative Criminology*, 25 (2009), 25–49. <https://doi.org/10.1007/s10940-008-9057-3>
- Craig, Ashley, Sara B. Heller, and Nikhil Rao, "A Preliminary Analysis of Spillovers in READI Chicago—Early Results from 'Using Network Data to Measure Social Returns and Improve Targeting of Crime-Reduction Interventions'," University of Michigan Working Paper, 2022.
- , "Using Network Data to Measure Social Returns and Improve Targeting of Crime-Reduction Interventions" (in progress).
- Cummings, Danielle, and Dan Bloom, "Can Subsidized Employment Programs Help Disadvantaged Job Seekers? A Synthesis of Findings from Evaluations of 13 Programs," OPRE Report 2020-23, 2020.
- Dinarte, Lelys, and Pablo Egaña del Sol, "Preventing Violence in the Most Violent Contexts: Behavioral and Neurophysiological Evidence," World Bank Policy Research Working Paper, 2019. <https://doi.org/10.1596/1813-9450-8862>
- Domínguez, Patricio, and Steven Raphael, "The Role of the Cost-of-Crime Literature in Bridging the Gap between Social Science Research and Policy Making: Potentials and Limitations," *Criminology and Public Policy*, 14 (2015), 589–632. <https://doi.org/10.1111/1745-9133.12148>
- Fagan, Jeffrey, and Deanna L. Wilkinson, "Guns, Youth Violence, and Social Identity in Inner Cities," *Crime and Justice*, 24 (1998), 105–188. <https://doi.org/10.1086/449279>
- Farrell, Albert D., David Henry, Catherine Bradshaw, and Thomas Reischl, "Designs for Evaluating the Community-Level Impact of Comprehensive Prevention Programs: Examples from the CDC Centers of Excellence in Youth Violence Prevention," *Journal of Primary Prevention*, 37 (2016), 165–188. <https://doi.org/10.1007/s10935-016-0425-8>
- Farrington, David P., Jeremy W. Coid, Louise Harnett, Darrick Jolliffe, Nadine Soteriou, Richard Turner, and Donald J. West, *Criminal Careers up to Age 50 and Life Success up to Age 48: New Findings from the Cambridge Study in Delinquent Development* (London: Home Office Research, Development and Statistics Directorate, 2006).
- Florence, Curtis, Jonathan Shepherd, Iain Brennan, and Thomas Simon, "Effectiveness of Anonymised Information Sharing and Use in Health Service, Police, and Local Government Partnership for Preventing Violence Related Injury: Experimental Study and Time Series Analysis," *BMJ*, 342 (2011). <https://doi.org/10.1136/bmj.d3313>

- Geller, Amanda, Jeffrey Fagan, Tom Tyler, and Bruce G. Link, "Aggressive Policing and the Mental Health of Young Urban Men," *American Journal of Public Health*, 104 (2014), 2321–2327. <https://doi.org/10.2105/AJPH.2014.302046>
- Green, Ben, Thibaut Horel, and Andrew V. Papachristos, "Modeling Contagion through Social Networks to Explain and Predict Gunshot Violence in Chicago, 2006 to 2014," *JAMA Internal Medicine*, 177 (2017), 326–333. <https://doi.org/10.1001/jamainternmed.2016.8245>
- Harcourt, Bernard E., *Illusion of Order: The False Promise of Broken Windows Policing* (Cambridge, MA: Harvard University Press, 2005).
- Haveman, Robert, Rebecca Blank, Robert Moffitt, Timothy Smeeding, and Geoffrey Wallace, "The War on Poverty: Measurement, Trends, and Policy," *Journal of Policy Analysis and Management*, 34 (2015), 593–638. <https://doi.org/10.1002/pam.21846>
- Heller, Sara B., "When Scale and Replication Work: Learning from Summer Youth Employment Experiments," *Journal of Public Economics*, 209 (2022), 104617. <https://doi.org/10.1016/j.jpubeco.2022.104617>
- Heller, Sara B., Benjamin Jakubowski, Zubin Jelveh, and Max Kapustin, "Machine Learning Predicts Shooting Victimization Well Enough to Help Prevent It," NBER Working Paper no. 30170, 2022. <https://doi.org/10.3386/w30170>
- Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack, "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago," *Quarterly Journal of Economics*, 132 (2017), 1–54. <https://doi.org/10.1093/qje/qjw033>
- Hendren, Nathaniel, and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," *Quarterly Journal of Economics*, 135 (2020), 1209–1318. <https://doi.org/10.1093/qje/qjaa006>
- Imbens, Guido W., "Statistical Significance, p-Values, and the Reporting of Uncertainty," *Journal of Economic Perspectives*, 35 (2021), 157–174. <https://doi.org/10.1257/jep.35.3.157>
- Jones, Nikki, "The Regular Routine: Proactive Policing and Adolescent Development among Young, Poor Black Men," *New Directions for Child and Adolescent Development*, 143 (2014), 33–54. <https://doi.org/10.1002/cad.20053>
- Kapustin, Max, Jens Ludwig, Marc Punkay, Kimberley Smith, Lauren Speigel, and David Welgus, "Gun Violence in Chicago, 2016," University of Chicago Crime Lab, 2017.
- Katz, Lawrence F., Jonathan Roth, Richard Hendra, and Kelsey Schaberg, "Why Do Sectoral Employment Programs Work? Lessons from WorkAdvance," *Journal of Labor Economics*, 40 (2022), S249–S291. <https://doi.org/10.1086/717932>
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75 (2007), 83–119. <https://doi.org/10.1111/j.1468-0262.2007.00733.x>
- Lipsey, Mark W., Nana A. Landenberger, and Sandra J. Wilson, "Effects of Cognitive-Behavioral Programs for Criminal Offenders," *Campbell Systematic Reviews*, 3 (2007), 1–27. <https://doi.org/10.4073/csr.2007.6>
- Loftin, Colin, and David McDowall, "The Use of Official Records to Measure Crime and Delinquency," *Journal of Quantitative Criminology*, 26 (2010), 527–532. <https://doi.org/10.1007/s10940-010-9120-8>
- Manski, Charles F., "Treatment Choice with Trial Data: Statistical Decision Theory Should Supplant Hypothesis Testing," *American Statistician*, 73 (2019), 296–304. <https://doi.org/10.1080/00031305.2018.1513377>
- MDRC, "Building Knowledge about Successful Prisoner Reentry Strategies," MDRC, 2013.
- Pattillo, Mary, Bruce Western, and David Weiman, *Imprisoning America: The Social Effects of Mass Incarceration* (New York: Russell Sage Foundation, 2004).
- Raphael, Steven, and Michael A. Stoll, *Why Are So Many Americans in Prison?* (New York: Russell Sage Foundation, 2013).
- Redcross, Cindy, Bret Barden, Dan Bloom, Joseph Broadus, Jennifer Thompson, Sonya Williams, Sam Elkin, Randall Juras, Janaé Bonsu, and Ada Tso et al.,

- "The Enhanced Transitional Jobs Demonstration: Implementation and Early Impacts of the Next Generation of Subsidized Employment Programs," MDRC and the Employment and Training Administration, US Department of Labor, 2016.
- Roman, Caterina G., Hannah J. Klein, and Kevin T. Wolff, "Quasi-Experimental Designs for Community-Level Public Health Violence Reduction Interventions: A Case Study in the Challenges of Selecting the Counterfactual," *Journal of Experimental Criminology*, 14 (2018), 155–185. <https://doi.org/10.1007/s11292-017-9308-0>
- Sherman, Lawrence W., and Dennis P. Rogan, "Effects of Gun Seizures on Gun Violence: 'Hot Spots' Patrol in Kansas City," *Justice Quarterly*, 12 (1995), 673–693. <https://doi.org/10.1080/07418829500096241>
- Skogan, Wesley G., Susan M. Hartnett, Natalie Bump, Jill Dubois, and et al., "Evaluation of Ceasefire-Chicago," Northwestern University, 2008.
- Viviano, Davide, Kaspar Wüthrich, and Paul Niehaus, "(When) Should You Adjust Inferences for Multiple Hypothesis Testing?," arXiv Working Paper no. 2104.13367, 2021. <https://doi.org/10.48550/arXiv.2104.13367>
- Westfall, Peter H., and S. Stanley Young, *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment* (Hoboken, NJ: Wiley-Interscience, 1993).
- Wheeler, Andrew P., Robert E. Worden, and Jason R. Silver, "The Accuracy of the Violent Offender Identification Directive Tool to Predict Future Gun Violence," *Criminal Justice and Behavior*, 46 (2019), 770–788. <https://doi.org/10.1177/0093854818824378>
- Williams, J. Patrick, "Emergent Themes," *Sage Encyclopedia of Qualitative Research Methods*, 1 (2008), 248–249.
- Wilson, David B., Leana Allen Bouffard, and Doris L. MacKenzie, "A Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders," *Criminal Justice and Behavior*, 32 (2005), 172–204. <https://doi.org/10.1177/0093854804272889>
- Wolfgang, Marvin E., and Paul E. Tracy, *The 1945 and 1958 Birth Cohorts: A Comparison of the Prevalence, Incidence, and Severity of Delinquent Behavior* (Philadelphia: Center for Studies in Criminology and Criminal Law, University of Pennsylvania, 1982).
- Ziliak, Steve, and Deirdre N. McCloskey, *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives* (Ann Arbor: University of Michigan Press, 2008).