Impact Evaluation of the LAPD Community Safety Partnership*

Sydney Kahmann † Erin Hartman ‡
Jorja Leap § P. Jeffrey Brantingham ¶

This version: September 28, 2020

Abstract

In 2011, the Los Angeles Police Department (LAPD), in conjunction with other governmental and nonprofit groups, launched the Community Safety Partnership (CSP) in several communities long impacted by multi-generational gangs, violent crime and a heavy-handed approach to crime suppression. Following a relationship-based policing model, officers were assigned to work collaboratively with community members to reduce crime and build trust. However, evaluating the causal impact of this policy intervention is difficult given the unique nature of the units and time period where CSP was implemented. In this paper, we use a novel data set based on the LAPD's reported crime incidents and calls-for-service to evaluate the effectiveness of this program via augmented synthetic control models, a cutting-edge method for policy evaluation. We perform falsification analyses to evaluate the robustness of the results. In the public housing developments where it was first deployed, CSP reduced reported violent crime incidents, shots fired and violent crime calls, and Part I reported crime incidents. We do not find evidence of crime displacement from CSP regions to neighboring control regions. These results are promising for policy-makers interested in policing reform.

Keywords: gang violence, community policing, time-series cross-sectional data, augmented synthetic control method, policy evaluation, causal inference

---

*We thank Chad Hazlett and the UCLA Causality Reading Group for thoughtful comments and discussion. We thank LAPD Chief of Police Michel Moore, the CSP Research and Evaluation Advisory Committee, and the community residents and institutional partners that make CSP possible. This research was funded by the Balmer Group, The California Endowment, Caruso, Cindy Miscikowski, The Smidt Foundation, the Weingart Foundation, and an anonymous donor. We acknowledge Connie Rice as a co-creator of the CSP public safety model and a visionary regarding relationship-based policing.

†Corresponding Author: skahmann@ucla.edu. Department of Statistics, UCLA. SK was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1650604.

‡Departments of Statistics and Political Science, UCLA.

§Department of Social Welfare, UCLA.

¶Department of Anthropology, UCLA. PJB serves on the board of PredPol.

1 Introduction

In 2011, the Los Angeles Police Department (LAPD) partnered with the Housing Authority of the City of Los Angeles, and the Urban Peace Institute to launch the Community Safety Partnership (CSP) (Rice and Lee, 2015). CSP was designed to address high levels of violent crime and the corrosive effects of multi-generational violent gangs entrenched in several Los Angeles public housing developments (PHDs). In Jordan Downs and Nickerson Gardens, located in the Watts neighborhood of South Los Angeles, gangs dominated the physical spaces and controlled the activities of daily life. Long-standing rivalries between gangs in the PHDs brought near daily conflict to the communities and allowed violent and property crime to thrive. Residents in PHDs in Los Angeles had largely resigned themselves to a life of insecurity and instability, vulnerabilities reinforced by high levels of concentrated disadvantage.

At its core, CSP was a rethinking of how to approach the problems of violent crime and gangs built around a relationship-based policing model. Two decades of heavy-handed crime suppression had succeeded in cultivating widespread distrust of police, but appeared to do little to blunt the control of gangs. CSP saw arrest-based crime suppression as a last resort and rather sought to have "police officers and residents work in mutually respectful partnership to identify and prevent crime" (LA Office of the Mayor, 2017). Funded by the LAPD and the Housing Authority of the City of Los Angeles, CSP recruited and trained a special group of officers for five-year assignments in the PHDs. CSP officers were tasked with "support[ing] community and youth programs, address[ing] quality of life issues and develop[ing] programs to address and reduce violent crimes" (LAPD News Release, 2015). In effect, CSP officers were to become invested members of the community, rather than simply an intervening force that shows up when there are problems.

The communities participating in CSP exhibited reduced violent crime and improved relations with the police almost immediately, a fact that attracted considerable media attention (e.g. Siegler, 2013; Blackstone, 2014; Streeter, 2014). Though encouraging, it is critical to recognize that inferring causality is challenging under the circumstances. CSP was implemented in highly targeted communities and developed at a time when Los Angeles was experiencing an unprecedented decline in crime. Crime peaked in Los Angeles in 1992, a year which saw nearly 1,100 murders citywide. Los Angeles then experienced nearly two decades of falling violent and property crimes, while simultaneously adding a half-million new residents. At its lowest point in 2013, there were 251 homicides citywide, a 335% reduction. Given the overall decrease in crime throughout the city, it is possible that crime would have fallen in CSP areas, even if the program had not existed. Given the significance surrounding policing reform in America, it is important to carefully consider the causal impact of community-policing programs such as the CSP.

To determine whether CSP actually reduced crime, we require knowledge of what crime and disorder *would have been like* in these communities had CSP not been implemented, i.e., the counterfactual. For example, in Jordan Downs and Nickerson Gardens, two CSP PHDs located in South Los Angeles, average violent crime dropped from roughly 31 incidents pre-CSP to roughly 21 incidents post-CSP per PHD per semester (six-month period). In order to attribute this whole reduction to CSP, one would have to assume crime would have remained constant throughout the period of study had CSP not been implemented, an assumption almost certainly violated given trends seen city-wide. Therefore, we wish to know how much of that reduction is likely due to the impact of CSP versus how much is attributable to the overall reduction of crime in Los Angeles during this period.

One common approach, the difference-in-differences method, assumes that underlying patterns in CSP regions are similar to other Los Angeles communities which did not participate in CSP. Average violent crime in census tracts in Los Angeles, excluding those eligible for CSP, saw a reduction from roughly 8.2 to roughly 7.6 incidents per census tract per semester, indicating the counterfactual reduction had CSP not been implemented would have been less than 1 violent crime per housing development per semester. Under this strong "parallel trends" assumption—that crime patterns would have had a similar trajectory in the CSP units as in the greater South Los Angeles area—the impact of CSP was a reduction of roughly 10 violent crimes per housing development per semester.

However, the ecological uniqueness of the public housing developments in which CSP was implemented, for example, in terms of housing density and the wide range of observable and unobservable characteristics that describe these neighborhoods, strains the credibility of the parallel trends assumption. It also does not account for the temporal and spatial richness of the data available for analyzing the impact of CSP. Using the same daily crime data for South Los Angeles, an alternative to the difference-in-differences approach is the synthetic control method (SCM, Abadie, Diamond, and Hainmueller, 2010). In lieu of assuming the naturally-occurring group of control units is comparable in its average level or trends in crime, the synthetic control method weights the control units to match the observed pre-treatment outcome among the treated units. This combination of control units, the *synthetic* control, "often does a better job of reproducing the characteristics of the unit or units representing the case of interest than any single comparison unit alone" (Abadie, Diamond, and Hainmueller, 2015, pg. 496). By projecting out this model over the post-treatment period, we can estimate a counterfactual for the treated unit and therefore estimate an effect of treatment. Due to concerns about in-exact balance and extrapolation, we apply a derivative of the SCM method, the augmented synthetic control method (ASCM, Ben-Michael, Feller, and Rothstein, 2019), to evaluate the impact of the CSP. This recent approach provides a unique avenue for estimating the impact of CSP given the challenges for causal identification in this data set. We test observable implications of the identifying assumptions necessary for the ASCM by conducting a series of placebo tests, including both spatial and temporal tests, to investigate the credibility of the resulting estimates.

We evaluate the impact of CSP on crime in Nickerson Gardens and Jordan Downs, two public housing developments in South Los Angeles, for the six-year period between 2012-2017 using a novel data set based on the LAPD's reported crime incidents and calls-for-service data. Specifically, we estimate the average number of violent crime incidents, Part I crime incidents, and violent crime calls-for-service prevented per semester per public housing development. We find that, on average, CSP prevented 6.55 violent crimes and reduced violent crime calls-for-service by 2.96 per semester per housing development between January 1, 2012 and December 31, 2017. Additionally, we find that, on average, CSP led to an reduction of 4.04 Part I reported crime incidents per semester per housing development between January 1, 2012 and December 31, 2017.

The remainder of this paper proceeds as follows. In Section 2, we describe CSP and how it sought to address well-known challenges in "community policing." In Section 3, we describe the data used in model balancing and testing. Section 4 introduces augmented synthetic control methods and underlying assumptions. Section 5 turns to model evaluation using placebo tests. Section 6 presents results on the impact of CSP in Jordan Downs and Nickerson Gardens PHDs. Section 7 discusses the implications of the results for CSP and relationship-based policing more broadly.

2 Background

The Community Safety Partnership was launched in four public housing developments in Los Angeles in late 2011 (Leap, 2020). It is one part of a comprehensive approach to violence reduction in places that have long suffered under the control of powerful street gangs and the corrosive effects of crime suppression tactics, persistent neglect by city officials, and concentrated social and economic disadvantage. Gang prevention and intervention efforts, as well as broad community engagement projects, are spearheaded by the Mayor's Office of Gang Reduction and Youth Development (Tremblay et al., 2020). Infrastructure improvement in PHDs is the responsibility of the Housing Authority of the City of Los Angeles. CSP is responsible for establishing and sustaining basic security and safety in the targeted communities. The approach of CSP is to concentrate on building long-lasting relationships between police and members of the community, leveraging those relationships for collaborative problem solving (Leap, 2020; Rice and Lee, 2015).

CSP extends a long history of community policing efforts in both the United States and United Kingdom, while seeking to correct for deficiencies in these past efforts. The Anglo-American model of policing places service to the public and crime prevention at its core (Reisig, 2010). These core principles for policing, arising during the founding of the London Metropolitan Police Service in 1829, have often been set aside in the drive for modernization and professionalization of police organizations around a paramilitary model. The professional model of policing was envisioned as a means for strictly regulating the legitimate use of force, limiting police discretion, and preventing corruption from within the ranks and political influence from outside. However, the quest for professionalism and efficiency also encouraged the impersonal delivery of police services and use of law enforcement tactics decoupled from the social contexts in which crime and disorder occur. In essence, citation, arrest, and the use of force, even if legitimate in a legal sense and necessary to ensure safety, can appear arbitrary and capricious if police have neither the time, inclination or incentive to understand the context in which such actions are taken. A lack of trust in police is an understandable outcome.

The first serious moves towards community policing in the late 1960s sought to reestablish the importance of context, while continuing along the path of police modernization and professionalization.[1] The President's Commission on Law Enforcement and Administration of Justice laid bare the many failings of policing in mid-Century America (Blumstein, 2018; Katzenbach, 1967). The Commission's report recommended a series of reforms designed to improve police-community relationships including geographic-based commands to deal with local (especially minority) community needs, citizen-advisory groups to convey what those needs are, and attention to recruiting and promoting minority officers. Team policing initiatives in the early 1970s, though short-lived, were a direct response to these recommendations (Sherman et al., 1973).

In spite of promising early steps, community policing today is perhaps best described as a general orientation adopted by policing organizations. Indeed, the US Department of Justice defines community policing as "a philosophy that promotes organizational strategies which support the systematic use of partnerships and problem-solving techniques, to proactively address the immediate conditions that give rise to public safety issues such as crime, social disorder, and fear of crime" (DOJ, 2009, pg.1). The most persistent criticism of community policing is that it is too amorphous (Cordner, 1997). A wide range of policing strategies and tactics may qualify as community policing. Thus, evaluation and generalization of individual community policing efforts is challenging.

CSP in Los Angeles was designed with these weaknesses of past community policing efforts in mind. It is a deliberate model of police recruitment, training, deployment, strategic and tactical orientation, and command oversight (Leap, 2020). CSP started with a year of planning prior to launch in 2011. A detailed selection process was established to recruit officers with existing orientations towards problem solving. Selected officers underwent training aimed at building understanding of the interrelated cultural, demographic, and economic factors that impact public safety in CSP sites. Training was designed and

[1]Incidents of excessive use of force and corruption continue to arise with unfortunate regularity (Fryer Jr, 2016), indicating that there is still room for significant police reform.

delivered by the Urban Peace Institute, a community-based civilian organization. Officers were trained on techniques for defusing community-wide dangers without over-relying on traditional suppression tactics such as arrest.

The CSP model also recognized that alternatives to suppression require trust and a network of community relationships that could be called on to solve immediate, local problems. Since building reliable social networks requires both time and stability of effort, CSP established long-term deployments for officers, lasting five-years in each community. The deployments also had a separate command structure allowing for greater autonomy and discretion of officers. Officers were provided unique incentives (promotion and pay) to reward community-engaged behaviors not captured by traditional metrics such as crime and arrest statistics. The ultimate goal of CSP was not only to build trust in policing, but also to provide the basic security and safety necessary for normal social and economic activity. While it is clear that CSP has not been immune to many of the well-known challenges facing community policing, such as ambiguity about how to balance enforcing laws against relationship building (Leap, 2020), even partial adherence to the CSP model may be expected to have an impact. Our purpose is to evaluate whether CSP succeeded in providing basic security and safety in the PHDs where it was deployed.

3 Data

Outcomes of interest for evaluating CSP come from two data sources: LAPD reported crime incidents data and LAPD calls-for-service data. Reported crimes typically originate from a call to the police by a member of the public. However, because reported crimes also undergo a verification process, they filter out much of the noise associated with calls (Klinger and Bridges, 1997). Calls-for-service are thus viewed as an aggregate indicator of police demand, fear of crime, and victimization (Porter et al., 2019). Reported crimes are viewed primarily as an indicator of victimization. The observed reported crime incidents available for use in this study occurred between January 1, 2006 and December 31, 2017, while observed calls for service incidents occurred between June 5, 2007 to May 31, 2019. We only include observations with valid geospatial coordinates so the location corresponding to treatment status (CSP or non-CSP) can be attributed. We exclude crime incidents and calls recorded with geospatial coordinates corresponding to police stations.[2] The final dataset is restricted to reported incidents and calls-for-service from July 1, 2007 to December 31, 2017, the period of overlap for the two data sources which fully captures as many semesters as possible.

Six PHDs received treatment under CSP within LAPD's South Bureau: Jordan Downs, Nickerson Gardens, Imperial Courts, Avalon Gardens, Gonzaque Village, and Harvard Park[3] (listed in order of implementation, earliest to latest). CSP was implemented in the first three developments in November 2011, the next two in July 2016, and the last in October 2017. CSP was also deployed in the Ramona Gardens PHD in East Los Angeles in late 2011. Ramona Gardens was not considered in our analyses. The remainder of this paper is focused on Jordan Downs and Nickerson Gardens, where enough time has passed since CSPs initial implementation to evaluate the impact.[4] Analyses for the latter three developments are not included in the main body of this paper as not enough time has passed to follow post-treatment behavior. See Sections S6 and S7 of the Supplementary Material for these preliminary results.

Due to the sensitive nature of the data, we are not able to provide the raw data as part of the publication. Substantially similar open-source data is available at `https://data.lacity.org/`. In addition, individuals may request data for research purposes by contacting the LAPD directly. Instructions are available at `https://www.lapdonline.org/inside_the_lapd/content_basic_view/9136`.

3.1 Control Units

As PHDs are unique in their community structure, we aim to find naturally-occurring control units that are of similar size and lie within a similar geographic region. We use

[2]It is common practice to use a police station address when the location of the crime is unknown. This exclusion removes approximately two percent of the data.

[3]Harvard Park is the only treated unit which is not a PHD. Instead, Harvard Park is a traditional residential neighborhood. Therefore, CSP may act differently in this region due to confounding factors as compared to CSP PHDs (e.g. strength of community ties, building composition, local housing association/ governance structures).

[4]The results are qualitatively very similar when the Imperial Courts PHD is included, as shown in the Supplementary Material.

Census boundaries, and their associated shapefiles, to construct control units for analysis.

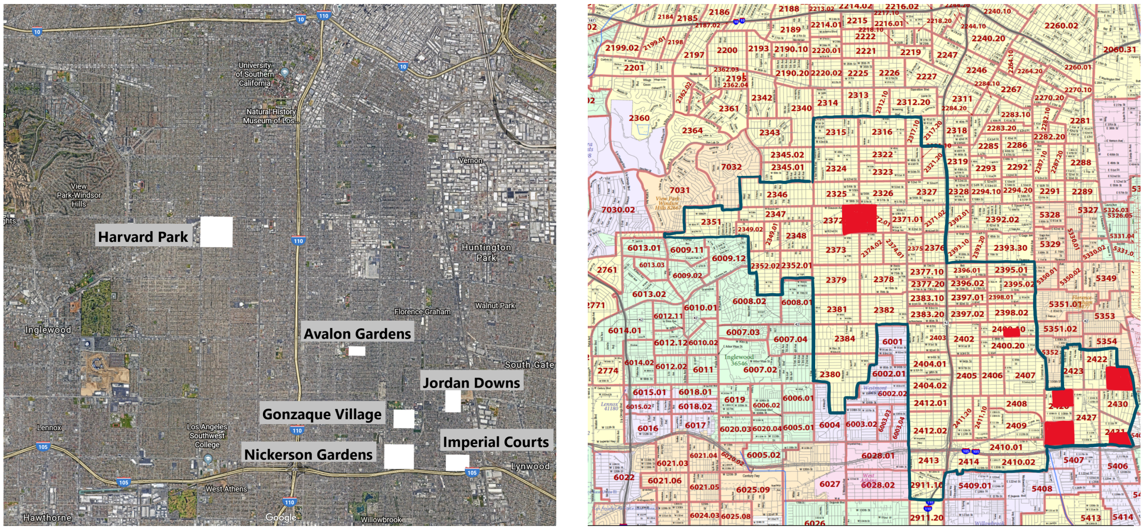

Figure 1: Left: a Google Earth view of the region of interest with CSP PHDs labelled. Right: South Bureau in terms of Census Tract boundaries to the southernmost cut-point of Tract 2911.10. Region of study outlined in blue. CSP public housing developments in red. The outlined region of interest is a subset of South Bureau including most of the LAPD's Southeast and 77th Street Divisions.

To construct control units, we aggregate raw reported crime incident and calls-for-service observations across space and time using Census geographies. Spatially, events are grouped by Census boundaries by *block group*. Temporally, events are grouped by fiscal *semester*, the second semester of every year begins on July 1.

We focus on *block group*-level control units, rather than *block*-level control units (the smallest Census geography unit) as the former is of a similar size to public housing units. All CSP PHDs are comprised of multiple Census blocks. [5]

To spatially define the control units, we first restrict the data to tracts within the 77th Street and Southeast Divisions of LAPD's South Bureau. Southeast Division is then reduced further, with the southernmost tract in the region being Tract 2911.10. We aim to

[5] Data from the Census Bureau's Decennial Census, which records various information on demographics, socioeconomic status, etc., are only available for *blocks* and *block groups*. Therefore, defining the region of interest in terms of Census boundaries allows us to use Census population data to consider the effect of CSP in terms of per capita outcomes. These analyses are included in Section S3 of the Supplementary Material.

avoid interpolation bias by restricting the pool of control units to a small geography around the treated units, rather than using the whole of the LAPD area (Abadie, Diamond, and Hainmueller, 2015). A Google Earth view of the final region and the resulting breakdown into Census tracts is shown in Figure 1. In total, the pool of control units consists of 234 *block groups* where all units that eventually receive treatment are excluded.

Temporally, *semesters* are defined as events from January 1 to June 30 and from July 1 to December 31 within a given year. *Semester* is chosen as the smallest time-level specification as it is better able to demonstrate seasonal trends than *year* and is less noisy than *quarter*-level data.[6]

3.2 Treated Units

The PHDs that receive CSP define our treated units. While bigger than a census *block*, they do not always consist of an entire *block group* region. Therefore, all *blocks* within a PHD are aggregated and recoded as a single, new *block group* unit per PHD. *Blocks* assigned to a PHD are then removed from the set of control *block groups*. Census *block group* boundaries across the region of study are shown in Figure 1. For figures demonstrating the allocation of Census units to PHD boundaries for each PHD, see Section S1 of the Supplementary Material. In our analyses, Jordan Downs encompasses seven *blocks* and Nickerson Gardens encompasses 13 *blocks*. Jordan Downs and Nickerson Gardens appear of similar geographical size to the average control unit in our study: the average number of *blocks* within a control *block group* in our study is 10.41.

CSP was implemented in Jordan Downs and Nickerson Gardens in November 2011. Therefore, at the semester level, treatment implementation is approximated with the beginning of 2012. Under this approximation, the final pre-treatment semester contains approximately two months of CSP conditions.[7]

[6]The *semester-level* results are substantively similar to the *quarter*-level results, as shown in Section S3 of the Supplementary Material.

[7]If we reassign the treatment periods to be exact (i.e. recode incidents in November and December 2011 to have occurred during the first semester of 2012), our results are comparable. For *violent crime*, the ATT estimates are -6.55 and -6.26 under approximated treatment assignment and exact treatment assignment, respectively. For *Part I crime*, the ATT estimates are -4.04 and -4.52 under approximated

3.3 Outcomes

In the body of this paper, we are primarily interested in crime and disorder in Jordan Downs and Nickerson Gardens. We focus on *violent crime* incidents, *shots fired and violent crime* calls-for-service, and *Part I crime* incidents. The US Federal Bureau of Investigation tracks two broad categories of crimes on an annual basis: Part I and Part II crimes. The so-called Part I crimes are serious crimes that occur with sufficient frequency and visibility to warrant statistical tracking. The Part I crimes include homicide, aggravated assault (assault with a deadly weapon), robbery, rape, burglary, car theft and arson. It excludes crimes such as kidnapping, which is serious, but rare. The so-called Part II crimes are lesser crimes that range from simple assault to vandalism. In addition to violent crime outcomes, we examine the Part I crimes excluding arson. We exclude arson from the standard list of Part I crimes due to irregular reporting in police data.

The outcomes are defined as follows (LAPD Consolidated Crime Analysis Database (CCAD) code in parentheses): *violent crime* is defined as homicide (110), assault with a deadly weapon/ attempted homicide (230), and robbery (210); *shots fired and violent crime* calls-for-service is defined as shots fired (246), robbery (211), assault with a deadly weapon (245), and murder (187) calls. *Part I crime* incidents is defined as homicide (110), assault with a deadly weapon/ attempted homicide (230), robbery (210), rape (121), burglary (310), and stolen vehicle (510). [8]

4 Methodology

In this section we introduce the augmented synthetic control method (ASCM, Ben-Michael, Feller, and Rothstein, 2019). We also introduce identification assumptions for this method with special attention given to possible spillover effect violations related to the Stable Unit

treatment assignment and exact treatment assignment, respectively. For *shots fired and violent crime*, the ATT estimates are -2.96 and -2.80 under approximated treatment assignment and exact treatment assignment, respectively.

[8] Additional outcomes (reported crime incidents: residential burglary; calls for service: quality of life) were considered and found to have poor synthetic control fit. Placebo results included in Section S5 of the Supplementary Materials.

Treatment Value Assumption (SUTVA) under the potential outcomes framework (Rubin, 1974; Neyman, 1923). We do not find evidence of spillover effects.

4.1 Notation and Estimands

Consider a case with N total units, indexed by $i \in (1, \ldots, N)$, and time periods indexed by $t \in (1, \ldots, T)$. Define a treatment indicator, D_i, such that units receiving treatment are denoted by $D_i = 1$ and units who do not receive treatment, i.e. those in control, have $D_i = 0$. Treatment is implemented at time T_0. The first n_t units, $i \leq n_t$, are treated at time T_0 and the rest, $i > n_t$, are never exposed to treatment. Under the potential outcomes framework (Rubin, 1974), define unit i's potential outcomes at time t as $\{Y_{it}(0), Y_{it}(1)\}$ under control and treatment, respectively, where $Y_{it}(1)$ is only defined for $t \geq T_0$. Similarly, the treatment effect for unit i at time t is defined as

$$\tau_{it} = Y_{it}(1) - Y_{it}(0). \tag{1}$$

The fundamental problem of causal inference is that the potential outcomes for unit i and time t are never jointly observable, making τ_{it} unobservable (Holland, 1986). Yet, this quantity is of primary interest in policy problems such as the CSP evaluation, where we wish to know the impact of CSP for, specifically, the units who received treatment. To evaluate the CSP policy, our primary quantity of interest, then, is the average treatment effect on the treated (ATT), a policy-relevant estimand that captures the impact of the CSP for the PHDs that participated in the program. We can also evaluate the average treatment effect for the treated units at a given time period, called the ATT_t:

$$ATT_t = \mathbb{E}[Y_{it}(1) - Y_{it}(0) | D_i = 1] \tag{2}$$

$$ATT = \mathbb{E}[ATT_t], T_0 < t < T. \tag{3}$$

Substantively, the ATT_t (Equation (2)) is the average difference between the crime rates in treated regions post-treatment and *what the crime rate would have been* in those regions during the post-treatment years had CSP not been implemented for a given post-treatment time period t. Equation (3), the average of the ATT_t over all post-treatment time periods, captures the overall impact of CSP.

Under the potential outcomes framework, we require the SUTVA (Rubin, 1974) that there be no interference between units, meaning the treatment status of unit i cannot affect the potential outcomes of unit j, for $j \neq i$. While the credibility of SUTVA for this application is more thoroughly discussed in Section S2.3 of the Supplementary Material, the concern is there may be interference between treated units and their adjacent control units. Crime displays patterns of local contagion (Mohler et al., 2011) and hot spot policing experiments have shown diffusion of benefits over relatively short spatial distances (Bowers et al., 2011). Therefore we evaluate potential crime displacement by estimating the effect of CSP using pseudo-treatment units neighboring the Jordan Downs PHD. We do not find evidence of crime displacement in these regions. Section S2.3.1 of the Supplementary Material describes our procedure and the relevant displacement result.

4.2 Synthetic Control Method (SCM)

We seek to know the impact of the CSP on the units that actually participated in the program, a quantity captured by the ATT. The ATT is composed of an observable quantity, the average crime rate at time t among the treated units, and an unobservable quantity, the average crime rate at time t for treated units *had they not participated in CSP*. The goal of synthetic control methods is to find weights w_i^* for each control unit, $i > n_t$, to construct an estimate for the counterfactual (unobserved) quantity $Y_{it}(0)$ for treated units (Abadie, Diamond, and Hainmueller, 2010). This can be used to estimate the treatment effect for treated units $i \leq n_t$ at time t with

$$\widehat{ATT}_t = \frac{1}{n_t} \sum_{i \leq n_t} Y_{it}(1) - \sum_{i > n_t} w_i^* Y_{it}(0). \tag{4}$$

Each weight w_i^* is determined so that the sum of the weighted pre-treatment control units "balances," i.e. makes summary statistics such as the mean match for, the treated units and weighted control units across the pre-treatment outcomes for each time period t. Weights are constrained to sum to one. We can test an observable implication of the assumptions by comparing the weighted pre-treatment control outcomes to the observed treated outcomes in the pre-treatment period. In the post-treatment period, we estimate the counterfactual potential outcome for treated units as

$$\frac{1}{n_t}\sum_{i\leq n_t} Y_{it}(0) \approx \sum_{i>n_t} w_i^* Y_{it}(0) := \widehat{Y_t^{SCM}} \tag{5}$$

where we define $\widehat{Y_t^{SCM}}$ as the synthetic control estimate at each time t. Using SCM, the ATT can be estimated as

$$\widehat{ATT} = \sum_{t>T_0}\left(\frac{1}{n_t}\sum_{i\leq n_t} Y_{it}(1) - \widehat{Y_t^{SCM}}\right) \tag{6}$$

where the average outcomes are estimated over the post-treatment time period $t > T_0$. SCM is only able to construct a valid synthetic control if the set of treated units is contained within the convex hull defined by the pre-treatment outcomes of the control units. There has to exist a weighted sum of control units that resembles the treated units on pre-treatment outcomes to reduce the potential for large pre-treatment imbalances under the model. As discussed in Section S2.1 of the Supplementary Material, the raw count *violent crime* outcome does not fit the convex hull criterion. Therefore, the augmented synthetic control model which adjusts for such pre-treatment imbalances is a natural choice for analysis in the case of a likely convex hull violation.

When considering crime and disorder outcomes, SCM approaches have been previously used to investigate the effect of policy interventions (i.e. right-to-carry laws (Donohue, Aneja, and Weber, 2019); permit-to-purchase laws (Rudolph et al., 2015)) and also to investigate causal relationships between crime and ongoing or naturally-occurring phenomena (i.e. drought (Goin, Rudolph, and Ahern, 2017); sporting events (Pyun, 2019); organized crime (Pinotti, 2011; Becker and Klößner, 2017)). Many of these applications focus on state or city-level effects. Our work on the CSP intervention is more closely related to neighborhood-based crime policy interventions (Saunders et al., 2015; Robbins, Saunders, and Kilmer, 2017; Rydberg et al., 2018).

4.3 Augmented Synthetic Control Method (ASCM)

There is a robust and developing time-series cross-sectional literature that include methods that relax the assumptions required for the SCM (e.g. Hazlett and Xu, 2018; Athey et al., 2018; Imai, Kim, and Wang, 2018; Brodersen et al., 2015). We use the ASCM (Ben-Michael, Feller, and Rothstein, 2019), a derivative of the SCM that uses a model-based adjustment to account for bias introduced by inexact balance between the treated and control units and for extrapolation outside the convex hull. As discussed in Ben-Michael, Feller, and Rothstein, 2019, the bias between the pre-treatment fit of the treated units' observed pre-treatment state and the estimated synthetic control is expressed as $\frac{1}{n_t}\sum_{i\leq n_t} Y_{it}(0) - \sum_{i>n_t} w_i^* Y_{it}(0)$ for each time t. Because any potential outcome $Y_{it}(0)$ can be decomposed into the sum of an outcome model, m, and an error term, ϵ, then the bias under imbalanced SCM can be re-written as:

$$\text{Bias from SCM model} = [\sum_{i\leq n_t} m(X_i) - \sum_{i>n_t} w_i^* m(X_i)] + \mathbb{E}[\sum_{i\leq n_t} \epsilon_i - \sum_{i>n_t} w_i^* \epsilon_i] \tag{7}$$

for pre-treatment outcomes X_i. This bias is estimated using an outcome model, $\hat{m}$. Therefore, ASCM adds a bias term to the traditional SCM estimator:

$$Y_t^{\widehat{ASCM}} = \textstyle\sum_{i>n_t} w_i^* Y_{it}(0) + \left(\sum_{i\leq n_t} \hat{m}(X_i) - \sum_{i>n_t} w_i^* \hat{m}(X_i)\right) \tag{8}$$

$$Y_t^{\widehat{ASCM}} = Y_t^{\widehat{SCM}} + \left(\textstyle\sum_{i\leq n_t} \hat{m}(X_i) - \sum_{i>n_t} w_i^* \hat{m}(X_i)\right) \tag{9}$$

In our analyses we use Ridge regression (Tibshirani, 1996) as the outcome model. When the pre-treatment SCM achieves exact balance, Ridge ASCM and traditional SCM will return the same weights. In this context, the treated units are outside the convex hull of control units. Therefore, the Ridge ASCM improves the SCM pre-treatment balance while penalizing the inherent extrapolation.

Defining X_1 as the matrix of pre-treatment outcomes for treated units and X_0 as the matrix of pre-treatment outcomes for control units, the SCM weighting criterion is adjusted so the vector of weights, w, is constrained to be non-negative and sum to one:

$$min_w ||X_1 - X_0' w||_2^2 + \zeta \sum_{i>n_t} f(w_i^*) \tag{10}$$

This re-formulation, Equation 10, allows for imperfect pre-treatment fit between the lagged outcomes while penalizing the dispersion of the weights via some penalizing function, f, and the hyperparameter, ζ. This dispersion penalty regulates the eventual ASCM weights which allow for extrapolation in the case of imperfect pre-treatment fit.

Standard error estimation is an ongoing field of research within the synthetic control context. Confidence bounds in this paper are estimated using the conformal inference method (Chernozhukov, Wuthrich, and Zhu, 2017). Under this framework, confidence interval upper and lower bounds are estimated for each post-treatment time period by leveraging the time-series component of the SCM counterfactual prediction problem, assuming the errors are weakly dependent and stationary across time. Each confidence interval can be interpreted as estimating a null hypothesis of no effect at the given time point. Given these pointwise estimates of confidence interval upper and lower bounds, we report the ATT estimates in text with the p-value for the joint null hypothesis of no effect provided in parentheses.

SCM assumes the outcome can be re-written as a linear factor model of the pre-treatment outcomes, time-related terms, and error term (Abadie, Diamond, and Hainmueller, 2010). Under this framework, unobserved influences on the outcome must be detected within the pre-period to be accounted for in the post-period. Any shocks to the treated unit must be detected in some subset of the control units. A violation of the linear factor model assumption may lead to large interpolation biases (Abadie, Diamond, and Hainmueller, 2010). Under this assumption, the ASCM model estimation error can be broken down into bias from underfitting the model due to imbalance and bias from overfitting the model to noise (Ben-Michael, Feller, and Rothstein, 2019). The augmentation step aims to reduce the bias from imbalanced fit without overfitting. In Section S2.2 of the Supplementary Material we present the ASCM's estimated bias component and discuss the potential bias-reduction from the ASCM augmentation step.

5 Model Evaluation

Evaluating the credibility of SCM methods is difficult, particularly in light of the fact that we are estimating the treatment effect for a very small number of units, in this case only two PHDs. Traditionally, the most common way to evaluate the fit of synthetic control methods has been to use placebo tests (Abadie, Diamond, and Hainmueller, 2015), which rely on the idea that we should not find evidence of treatment effects where none should exist, i.e. before treatment has been implemented or among the control units. Estimation of nonzero placebo effects would undermine the credibility of the final results. The more evidence that passes the scrutiny of the placebo tests, the more credible the resulting analysis.

These placebo tests, suggested by the literature (Heckman and Hotz, 1989; Abadie, Diamond, and Hainmueller, 2010; Abadie, Diamond, and Hainmueller, 2015), serve as the primary method of evaluating the appropriateness of SCM methods. In particular, we assess the fit of the ASCM for each outcome separately, using the following approaches: (1) estimating placebo (i.e. null) effects for the pre-treatment period to assess the model fit; (2) assessing potential confounding events or anticipation effects by running models with pseudo pre-treatment implementation dates; (3) evaluating the range of placebo treatment effects by estimating the placebo effect for each control unit (Section S4 of the Supplementary Material).

Taken holistically, these checks are used to evaluate the credibility and confidence in each outcome model. *Violent crime* incidents, *Part I crime* incidents, and *shots fired and violent crime* calls-for-service outcomes had strong results across the suite of evaluation approaches and are discussed in depth.

5.1 Model Specification

Before evaluating the impact of CSP in the post-treatment periods, we assess the ability of the ASCM to balance the trajectory of the pre-treatment outcome for the treated units and the synthetic control. During the pre-treatment period, CSP was not implemented and therefore the ASCM models should not detect a treatment effect for any outcome. Non-zero effects would indicate possible remaining confounding, calling into question causal inferences on that outcome. Any observed imbalance could also indicate the potential scale of bias in the estimated impact of CSP.

To ensure the placebo ASCM models are only evaluated on the pre-treatment period, we split this period into training and testing sets using a 2/3 : 1/3 rule. Under this rule, the placebo treatment implementation dates are assigned at the 2/3 marker of the pre-treatment period. For the Jordan Downs and Nickerson Gardens PHDs, where treatment was implemented in 2012, the psuedo-implementation date is 2010.5[9] and therefore three

[9] 0.5 years denotes treatment start at the second semester of the given year.

periods (those starting 2010.5, 2011, 2011.5) are considered post-treatment for evaluation of the placebo models.

While the results of this paper focus on the violent crime incidents, Part I crime incidents, and violent crime calls-for-service outcomes for the Jordan Downs and Nickerson Gardens PHDs, the model specification approach described here severely limits the time frame for these regions to 2007-2011. Besides containing very few temporal units at the semester level, this period is unusual due to the economic strain of the 2008 stock market crash and resulting recession through 2009. Therefore, the placebo tests for other PHDs are included in the Supplementary Material to evaluate our overall ability to find a suitable synthetic control for these outcomes. Overall good pre-treatment model fit for the three sets of PHDs in terms of these outcomes, see results in Sections S6 and S7, bolsters our confidence in the final models for the Jordan Downs and Nickerson Gardens PHDs.

The results for this simple placebo in-time are shown in Table 1. The placebo effect for *violent crime* is estimated as -2.55. For reference, the pre-treatment semester average for this outcome in Jordan Downs and Nickerson Gardens is 31.28 incidents. Therefore the ratio of the placebo estimate to the pre-treatment average is approximately 8 %. This suggests confidence in our ability to fit a synthetic control model for this outcome and these PHDs, as the placebo effect is small relative to the scale of the outcome.

Outcome	Pre-Treatment Average	ATT Estimate
Violent Crime Incidents	31.28	-2.55
Part I Crime Incidents	58.06	2.70
Shots Fired and Violent Crime Calls-for-Service	43.06	0.48

Table 1: Estimated placebo impact of CSP for the Jordan Downs and Nickerson Gardens PHDs. To assess scale of the estimates, the pre-treatment semester average is provided for each outcome.

Part I crime incidents has an estimated psuedo-ATT of 2.70, approximately 5% of the pre-treatment semester average of 58.06. The *shots fired and violent crime* calls-for-service estimate is 0.48, approximately 1% of the pre-treatment semester average of 43.06. We consider the proportionally small estimated bias in these outcomes a good indicator of our ability to construct appropriate synthetic control models for the *Part I crime* incidents and *shots fired and violent crime* calls outcomes.

The psuedo-ATT point estimates for *violent crime*, *Part I crime*, and *shots fired and violent crime* calls-for-service are roughly 39%, 67%, and 16% of the estimated impact of CSP in final ATT estimates for these outcomes, respectively. This could indicate the potential scale of bias in the final results.

5.2 Confounding Across Time

To assess potential anticipation effects or confounding events in the pre-treatment period, we conduct a placebo check in time. More specifically, potential anticipation effects for each PHD are evaluated by pseudo-"assigning" the treatment start date as two years, one year, and one semester earlier than the actual implementation date. If there are no other unobserved or uncontrolled factors influencing crime outcomes during the pre-treatment period, the resulting ATT_t estimates using the pseudo-implementation date should follow a similar trend as the true ATT_t estimates.

As demonstrated in Figure 2, count results for the Jordan Downs and Nickerson Gardens PHDs largely follow a consistent trend: the psuedo-treatment ATT_t (red) follows the estimated ATT_t (black with shaded conformal inference bounds for the 95 % confidence interval). Taking both the model specification and temporal results into consideration, we are most confident in the synthetic control for the count outcomes of *violent crime* incidents, *Part I crime* incidents, and *shots fired and violent crime* calls-for-service for the Jordan Downs and Nickerson Gardens PHDs.

6 Results

In this section, we evaluate the treatment effect of CSP on violent crime and Part I crime outcomes at the Jordan Downs and Nickerson Gardens PHDs, the outcomes for which the placebo tests from Section 5 are strongest.

As CSP was largely motivated by a desire to reduce violent crime in PHDs, we are substantively interested in the effect of CSP on the reported crime incidents *violent crime* outcome and the calls-for-service *shots fired and violent crime* outcome. From the previous

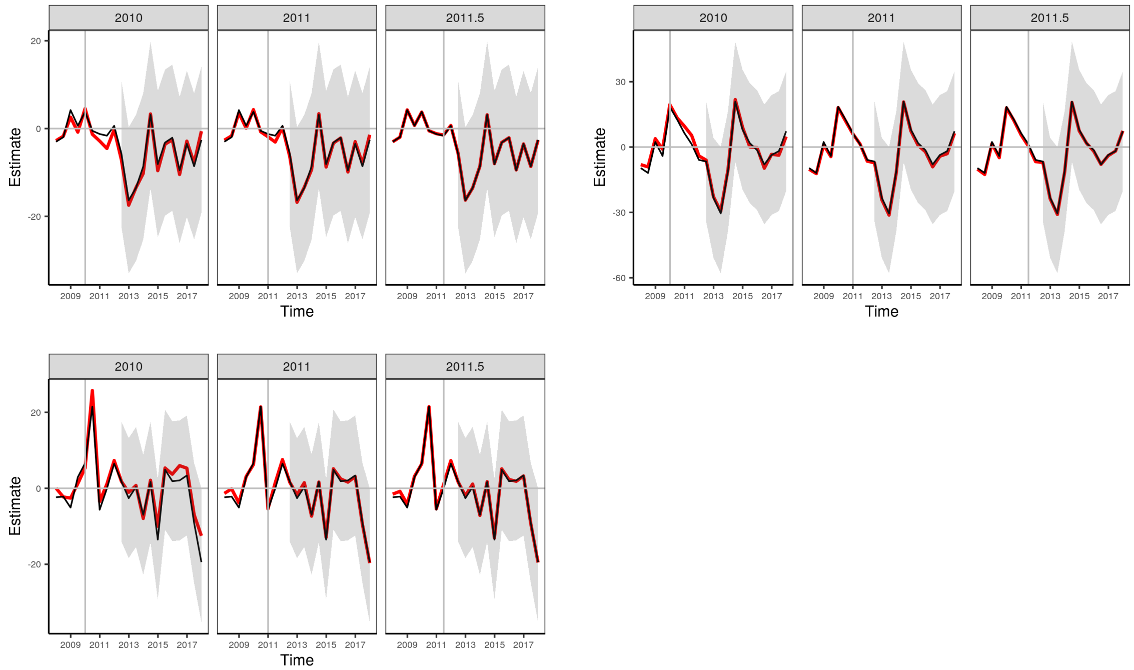

Figure 2: We evaluate robustness across time for Jordan Downs and Nickerson Gardens models of violent crime incidents, Part I crime incidents, and shots fired and violent crime calls (from left to right, top to bottom) by comparing psuedo-treatment ATT_t estimates (red) to the ATT_t estimated using the true implementation date of 2012 (black line with shaded conformal inference bounds). For each panel, the psuedo-implementation model is marked in the panel title. In the absence of confounding events during the pre-treatment period, we would expect to see the pseudo-implementation ATT_t estimates closely follow the ATT_t estimated from the true treatment period.

section, we find evidence that these results pass several placebo tests, bolstering confidence in the fit of the ASCM models for these outcomes of interest. As presented in Table 2, CSP has an estimated impact of 6.55 fewer reported *violent crime* incidents and 2.96 fewer *shots fired and violent crime* calls-for-service per semester per housing development during the post-treatment period. Compared to the pre-treatment semester averages, these PHDs experienced an average decrease of 21% and 7% in reported *violent crime* incidents and *shots fired and violent crime* calls per semester, respectively. For the Jordan Downs and Nickerson Gardens PHDs combined, the ATT estimates correspond to 13.1 fewer reported *violent crime* incidents and 5.92 fewer reported *shots fired and violent crime* calls-for-service per semester over the post-treatment period.

Additionally, CSP has an estimated reduction of 4.04 reported *Part I crime* incidents per semester per housing development during the post-treatment period, a decrease of 7% compared to the pre-treatment semester average. For the Jordan Downs and Nickerson Gardens PHDs combined, this ATT estimate corresponds to 8.08 fewer reported *Part I crime* incidents per semester over the post-treatment period.

Outcome	ATT Estimate
Violent Crime Incidents	-6.55
Part I Crime Incidents	-4.04
Shots Fired and Violent Crime Calls-for-Service	-2.96

Table 2: Estimated impact of CSP for the Jordan Downs and Nickerson Gardens PHDs.

6.1 CSP Reduced Violent Crime and Disorder, Part I Crime and Disorder in Deployed Areas

Figure 3 (left) plots the trajectories of the reported violent crime incidents in the Jordan Downs and Nickerson Gardens PHDs and the corresponding synthetic control model. One final model evaluation measure is evident in the pre-treatment period, the observed reported crime incidents in the treated units and the estimated synthetic control are very similar, as evidenced by similar trends in Figure 3 (left) and by ATT_t estimates near zero in Figure 3 (right).

Figure 3 (right) visualizes the changing *violent crime* trend across time while contextualizing the scale of the effect estimate. During the pre-treatment period, the Jordan Downs and Nickerson Gardens PHDs (solid line) each experienced approximately 30 violent crimes on average per semester. The synthetic control model (dashed line) closely tracks the mean trend in the Jordan Downs and Nickerson Gardens PHDs. The solid vertical line shows the CSP implementation at the end of 2011. In the post-treatment period, the observed behavior in Jordan Downs and Nickerson Gardens (solid) is consistently beneath the estimated synthetic control (dashed). The post-treatment treated and synthetic control units stabilize at a reduction of approximately 5 violent crimes on average per semester. The

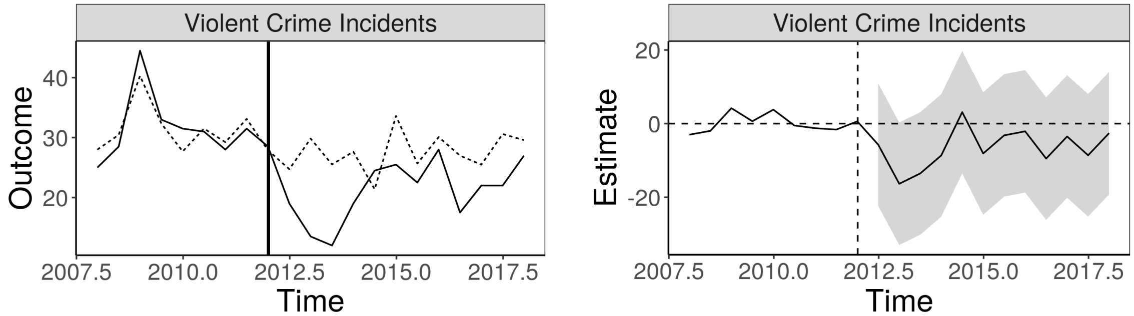

Figure 3: Left: reported violent crime incidents trajectories for the observed treated units (solid) versus the estimated synthetic control units (dashed). Right: the ATT_t estimates are provided with shaded conformal inference bounds. 2012 treatment implementation date is denoted with vertical line.

joint null p-value for no effect is 0.45. This result is perhaps unsurprising given the wide conformal inference bounds and inherent uncertainty in estimating synthetic controls.

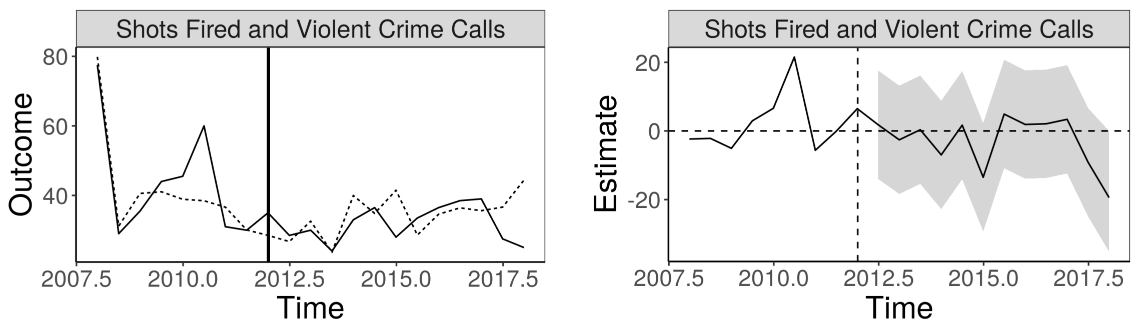

Figure 4: Left: reported shots fired and violent crime calls-for-service trajectories for the observed treated units (solid) versus the estimated synthetic control units (dashed). Right: the ATT_t estimates are provided with shaded conformal inference bounds. 2012 treatment implementation date is denoted with vertical line.

The *shots fired and violent crime* calls-for-service models, Figure 4, exhibit good pre-treatment fit with a relatively large discrepancy between the two models in 2010. Overall, the observed post-treatment behavior (solid line) in Figure 4 (left) closely tracks the estimated synthetic control (dashed line). The ATT plot, Figure 4 (right), suggests a null effect of CSP on *shots fired and violent crime* calls. This conclusion of a null effect is supported by the joint null p-value of 0.79.

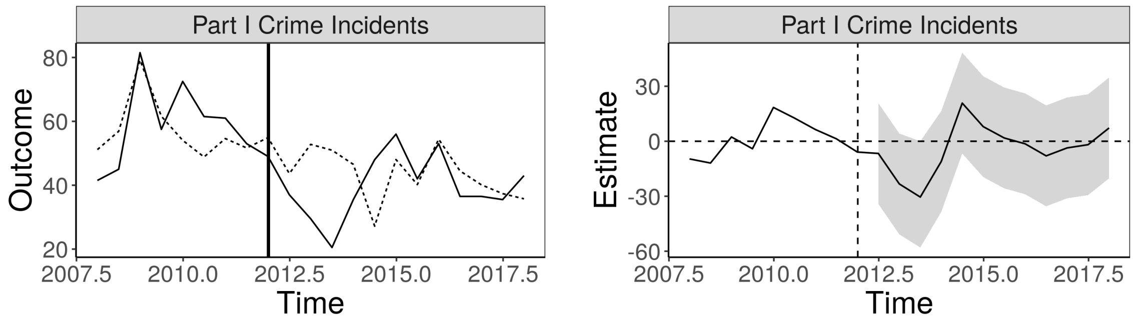

Figure 5: Left: reported Part I crime incidents trajectories for the observed treated units (solid) versus the estimated synthetic control units (dashed). Right: the ATT_t estimates are provided with shaded conformal inference bounds. 2012 treatment implementation date is denoted with vertical line.

The Ridge ASCM model for *Part I crime* incidents exhibits good pre-treatment balance in Figure 5. In the ATT plot (Figure 5, right), the ATT_t demonstrates an initial reduction in reported Part I crime incidents before stabilizing to a null effect in later years. As consistent with previous results, the joint null p-value is insignificant with a value of 0.5.

7 Discussion and Conclusions

This paper seeks to quantify the effect of the Community Safety Partnership on reported crime incidents and calls-for-service in Jordan Downs and Nickerson Gardens, two public housing developments in South Los Angeles. Both housing developments suffered from high levels of violent crime, entrenched multi-generational gangs, and policing focused on crime suppression before treatment implementation in 2011. Using the augmented synthetic control method, which allows us to construct counterfactuals for communities in which CSP was implemented; we found CSP led to an average reduction of 6.55 fewer violent crimes and 2.96 fewer violent crime calls-for-service per semester per housing development between January 1, 2012 and December 31, 2017. The estimates reflect reductions of 21% and 7% in reported crime incidents and calls-for-service, respectively, compared to pre-intervention means. CSP led to an average decrease of 4.04 Part I crime incidents per semester per housing development during this same period, corresponding to an average decrease of 7% in the Part I crime rate. Confidence bounds for these estimates are large, perhaps unsurprisingly, due to the nature of the synthetic control framework, which has few treated units, as well as the noisiness of the data. Furthermore, preliminary analyses suggest CSP did not simply displace crime from Jordan Downs to neighboring regions, as discussed in Sections S2.3 and S2.3.1.

We report results using raw counts of crime and disorder outcomes. However, the population density of the Jordan Downs and Nickerson Gardens PHDs is above that of the average control unit. The average population for the control units is 1278 in terms of 2010 Census population counts.[10] For comparison, the combined Jordan Downs and Nickerson Gardens region was reported to have 6719 residents: 2714 residents from Jordan Downs and 4005 residents from Nickerson Gardens. This increased density may contribute to a violation of the convex hull assumption, placing more emphasis on modeling to alleviate potential bias. To investigate sensitivity to this assumption, we run the analyses for per-capita outcomes defined as crime counts per 1000 residents. The Census population vector is recorded by *block* so we can construct an exact estimate of population, in terms of perfectly matched spatial boundaries, for both the treated and control units. Ultimately, as seen in Section S3 of the Supplementary Materials, the per-capita results are substantively similar to our analyses.

These results offer several benchmarks for cities considering developing relationship-based policing programs like CSP. First, relationship-based policing does appear to have a significant impact on violent crime above and beyond other factors. We present results per semester per housing development. The aggregate effect over the lifetime of the deployment (i.e. two developments over 12 semesters) is substantial. The impact appears to be substantial early on in the deployment then is attenuated at a modest stable reduction. These features of the treatment effects are important to know given the commitment in time and resources needed to mount a relationship-based policing program. CSP officers are dedicated to small areas, which removes them from general patrol or other investigative duties in the community at large. That this commitment yields a reduction demand for service and crime may more than compensates for the costs.

[10]US Census Bureau: 2010 Census, Summary File 1, Table P1: Total Population.

Finally, it is important to recognize that our results here do not speak to the other major component of CSP, which was to restore trust and build lasting relationships (Leap, 2020; Rice and Lee, 2015). Clearly ensuring the safety and security of community members is a necessary component of such a process, but there is more involved than simply low crime numbers. Future work will need to integrate evidence from the qualitative impact of CSP on people's lives as the counterpart to this quantitative story.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program". In: *Journal of the American Statistical Association* 105.490, pp. 493–505.

— (2015). "Comparative politics and the synthetic control method". In: *American Journal of Political Science* 59.2, pp. 495–510.

Athey, Susan et al. (2018). *Matrix completion methods for causal panel data models.* Tech. rep. National Bureau of Economic Research.

Becker, Martin and Stefan Klößner (2017). "Estimating the economic costs of organized crime by synthetic control methods". In: *Journal of Applied Econometrics* 32.7, pp. 1367–1369.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein (2019). "The Augmented Synthetic Control Method". In:

Blackstone, John (Sept. 8, 2014). "A change of tune for LAPD as community policing makes gains". In: *CBS*. URL: `https://www.cbsnews.com/news/a-change-of-tune-for-lapd-as-community-policing-makes-gains/` (visited on 10/22/2019).

Blumstein, Alfred (2018). "Science and technology and the President's Crime Commission: Past and future". In: *Criminology & Public Policy* 17.2, pp. 271–282.

Bowers, Kate et al. (2011). "Spatial displacement and diffusion of benefits among geographically focused policing initiatives". In: *Campbell Systematic Reviews* 7.1, pp. 1–144.

Brodersen, Kay H et al. (2015). "Inferring causal impact using Bayesian structural time-series models". In: *The Annals of Applied Statistics* 9.1, pp. 247–274.

Chernozhukov, Victor, Kaspar Wuthrich, and Yinchu Zhu (2017). "An exact and robust conformal inference method for counterfactual and synthetic controls". In: *arXiv preprint arXiv:1712.09089*.

Cordner, Gary W (1997). "Community policing: Elements and effects". In: *Critical issues in policing: Contemporary readings* 5, pp. 401–418.

DOJ, US (2009). "Community Policing Defined". In: *Office of Community Oriented Policing Services (COPS)*.

Donohue, John J, Abhay Aneja, and Kyle D Weber (2019). "Right-to-carry laws and violent crime: a comprehensive assessment using panel data and a state-level synthetic control analysis". In: *Journal of Empirical Legal Studies* 16.2, pp. 198–247.

Fryer Jr, Roland G (2016). *An empirical analysis of racial differences in police use of force.* Tech. rep. National Bureau of Economic Research.

Goin, Dana E, Kara E Rudolph, and Jennifer Ahern (2017). "Impact of drought on crime in California: A synthetic control approach". In: *PLoS one* 12.10.

Hazlett, Chad and Yiqing Xu (2018). "Trajectory balancing: A general reweighting approach to causal inference with time-series cross-sectional data". In:

Heckman, James J and V Joseph Hotz (1989). "Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training". In: *Journal of the American Statistical Association* 84.408, pp. 862–874.

Holland, Paul W (1986). "Statistics and causal inference". In: *Journal of the American Statistical Association* 81.396, pp. 945–960.

Imai, Kosuke, In Song Kim, and Erik Wang (2018). "Matching methods for causal inference with time-series cross-section data". In: *Princeton University* 1.

Katzenbach, Nicholas deBelleville (1967). *The challenge of crime in a free society: A report by the President's Commission on Law Enforcement and Administration of Justice.* United States Government Printing Office.

Klinger, David A. and George S. Bridges (1997). "Measurement Error in Calls-for-Service as an Indicator of Crime". In: *Criminology* 35.4, pp. 705–726. ISSN: 1745-9125. DOI: 10.1111/j.1745-9125.1997.tb01236.x. URL: http://dx.doi.org/10.1111/j.1745-9125.1997.tb01236.x.

LA Office of the Mayor (2017). *Mayor Garcetti announces new expansion of Community Safety Partnership.* URL: https://www.lamayor.org/mayor-garcetti-announces-new-expansion-community-safety-partnership (visited on 10/22/2019).

LAPD News Release (2015). *LAPD's Community Safety Partnership Program NR15021SF.* URL: http://www.lapdonline.org/home/news_view/57887 (visited on 10/22/2019).

Leap, J. (2020). *Evaluation of the LAPD Community Safety Partnership.* Los Angeles: UCLA Luskin School of Public Affairs.

Mohler, George O et al. (2011). "Self-exciting point process modeling of crime". In: *Journal of the American Statistical Association* 106.493, pp. 100–108.

Neyman, Jerzy S (1923). "On the application of probability theory to agricultural experiments." In: *Annals of Agricultural Sciences* 10, pp. 1–51.

Pinotti, Paolo (2011). "The economic consequences of organized crime: Evidence from Southern Italy". In: *Bank of Italy*.

Porter, Lauren C et al. (2019). "Where's the Crime? Exploring Divergences Between Call Data and Perceptions of Local Crime". In: *The British Journal of Criminology* 60.2, pp. 444–467. ISSN: 0007-0955. DOI: `10.1093/bjc/azz059`. URL: `https://doi.org/10.1093/bjc/azz059`.

Pyun, Hyunwoong (2019). "Exploring causal relationship between Major League Baseball games and crime: a synthetic control analysis". In: *Empirical Economics* 57.1, pp. 365–383.

Reisig, Michael D (2010). "Community and problem-oriented policing". In: *Crime and justice* 39.1, pp. 1–53.

Rice, Constance and Susan K Lee (2015). "Relationship-based policing achieving safety in Watts". In: *Washington, DC: Advancement Project*.

Robbins, Michael W, Jessica Saunders, and Beau Kilmer (2017). "A framework for synthetic control methods with high-dimensional, micro-level data: evaluating a neighborhood-specific crime intervention". In: *Journal of the American Statistical Association* 112.517, pp. 109–126.

Rubin, Donald B (1974). "Estimating causal effects of treatments in randomized and non-randomized studies." In: *Journal of Educational Psychology* 66.5, p. 688.

Rudolph, Kara E et al. (2015). "Association between Connecticut's permit-to-purchase handgun law and homicides". In: *American Journal of Public Health* 105.8, e49–e54.

Rydberg, Jason et al. (2018). "A quasi-experimental synthetic control evaluation of a place-based police-directed patrol intervention on violent crime". In: *Journal of Experimental Criminology* 14.1, pp. 83–109.

Saunders, Jessica et al. (2015). "A synthetic control approach to evaluating place-based crime interventions". In: *Journal of Quantitative Criminology* 31.3, pp. 413–434.

Sherman, Lawrence W et al. (1973). *Team policing: Seven case studies*. Police Foundation Washington, DC.

Siegler, Kirk (July 25, 2013). "After Years Of Violence, L.A.'s Watts Sees Crime Subside". In: *NPR*. URL: https://www.npr.org/sections/codeswitch/2013/07/25/205198028/Once-Crime-Ridden-South-L-A-s-Watts-Sees-Violence-Drop (visited on 10/22/2019).

Streeter, Kurt (Feb. 16, 2014). "In Jordan Downs housing project, police are forging a new relationship". In: *Los Angeles Times*. URL: https://www.latimes.com/local/la-me-jordan-downs-cops-20140217-story.html (visited on 10/22/2019).

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288.

Tremblay, Anne et al. (2020). "The Los Angeles Mayor's Office of Gang Reduction and Youth Development Comprehensive Strategy". In: *GRYD Research Brief No. 1*.

Supplementary Material
Impact Evaluation of the LAPD Community Safety Partnership

Sydney Kahmann [*] Erin Hartman [†]

Jorja Leap [‡] P. Jeffrey Brantingham [§]

September 28, 2020

S1 Census Units to PHD allocation

In allocating Census tracts, block groups, and blocks to PHD units, there are cases where Census boundaries do not exactly reflect natural public housing development (PHD) boundaries. Figure S1 shows the block allocations for each treated unit, regions that do not contain public housing are highlighted in red. As shown in the figure, we err on the side of over-estimation of the region of interest for conservative effect estimates. Of note for our main analyses, Nickerson Gardens (Tract 2426, top middle) is a total of 13 blocks where three of those contain some control regions. Jordan Downs (Tract 2421, top right) includes the entire tract, seven blocks, where three blocks contain some control regions. The large

[*]Corresponding Author: skahmann@ucla.edu. Department of Statistics, UCLA. SK was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1650604.

[†]Departments of Statistics and Political Science, UCLA.

[‡]Department of Social Welfare, UCLA.

[§]Department of Anthropology, UCLA. PJB serves on the board of PredPol.

untreated portion of Jordan Downs in Block 1003 is comprised of a high school, charter school, and several companies.

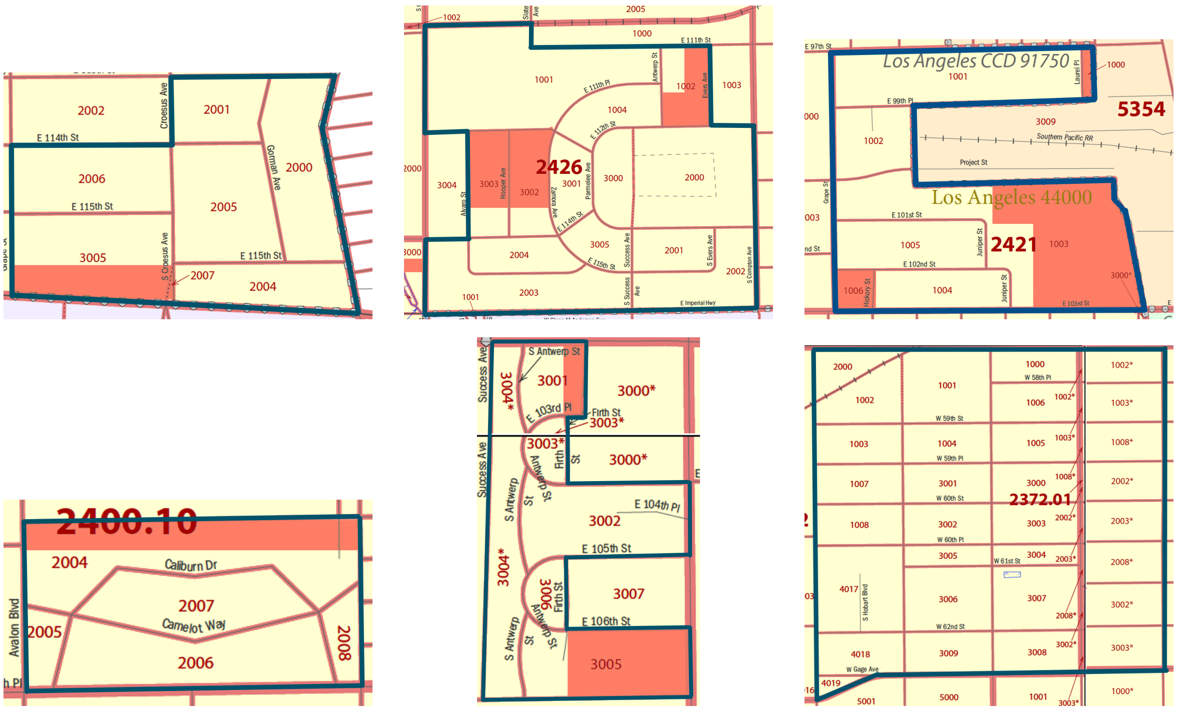

Figure S1: This paper uses block allocations for treated units as outlined in blue. The red highlighted regions do not contain public housing but are included under these boundary definitions. By including the control regions, we may slightly over-estimate the overall effect of CSP. Top, left to right: Imperial Courts, Nickerson Gardens, Jordan Downs. Bottom, left to right: Avalon Gardens, Gonzaque Village, Harvard Park.

S2 Evaluation of Model Assumptions

In this section, we more thoroughly evaluate the identifying assumptions required under this framework and 1) demonstrate the need to use a synthetic control method that accounts for the likely violations of the convex hull assumption, 2) discuss the linear factor model assumption and the potential for bias under the augmented synthetic control method (ASCM, 2019), and 3) fail to find evidence of displacement by failing to find evidence that the Stable Unit Treatment Value Assumption (SUTVA, 1974) does not hold.

S2.1 Assumptions: Convex Hull

The synthetic control method (SCM, 2010, Abadie, Diamond, and Hainmueller, 2015) uses a convex combination of the control units to construct an estimate of the treated unit counterfactual. Therefore, the SCM is only able to construct a valid synthetic control if the set of treated units is contained within the convex hull of the pre-treatment outcomes of the control units. There has to exist a weighted sum of control units that resembles the treated units on pre-treatment outcomes. If the treated units are too different from the controls, then SCM will not be able to find a set of weights that satisfies the weighting constraints and achieves balance between the pre-treatment outcomes for the treated units and the synthetic control. As shown in Figure S2, the raw count data does not fit the convex hull assumption, as PHDs tend to experience a higher volume of crime than our control units, likely partially due to the higher population density of PHDs. The treated outcomes tend to exceed the control outcomes across all time periods at the chosen *block group* unit specification. As a consequence, the traditional synthetic control estimator would be a biased estimator of our data. Crime rate outcomes calculated per capita do lie within the convex hull of pre-treatment control outcomes. For this reason, we check per capita results in Section S3 and show the results are robust to this alternative specification.

S2.2 Assumptions: Linear Factor Model

SCM assumes the outcome can be re-written as a linear factor model of the pre-treatment outcomes, time-related terms, and error term (Abadie, Diamond, and Hainmueller, 2010). A violation of the linear factor model assumption may lead to large interpolation biases (Abadie, Diamond, and Hainmueller, 2010). Under this assumption, in ASCM, the model estimation error can be broken down into bias from underfitting the model due to imbalance and bias from overfitting the model to noise (Ben-Michael, Feller, and Rothstein,

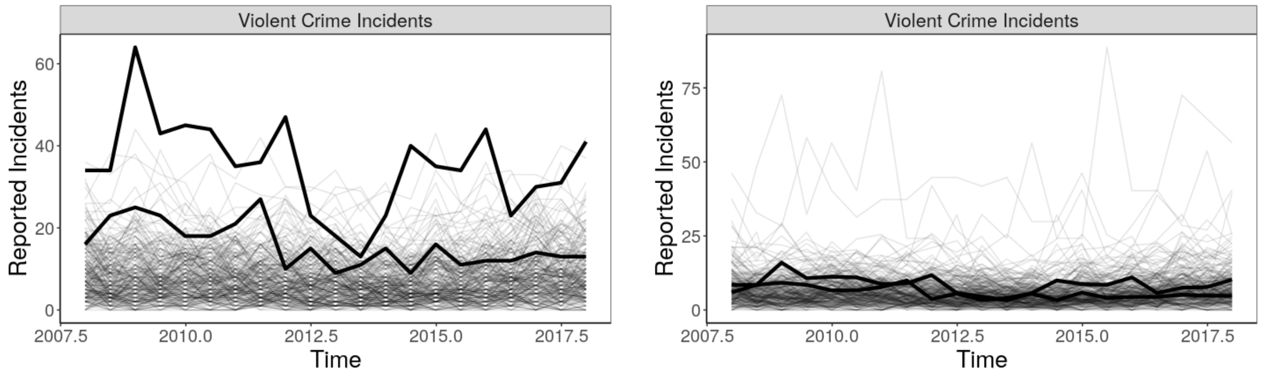

Figure S2: Violent crime outcomes for the treated units (bolded) versus control units (light gray) for the raw count outcome (left) vs per capita outcome (right). The violent crime count outcome shows a convex hull violation that is resolved when considering the rate of crime per 1000 persons.

2019). Ideally, the augmentation step would reduce the bias from imbalanced fit without overfitting.

Table S1 presents the estimated bias component to determine the possible bias-reduction from the modeling step in ASCM. First, we provide the average amount that ASCM's augmentation, or bias-correction, step changes the original SCM estimate (i.e. estimated bias). To grasp the scale of the bias relative to the final effect estimates, we provide the ratio of the average bias to the final point estimate, denoted as a percentage, in parentheses. Assuming the ASCM augmentation is improving model fit as opposed to fitting to noise, the imbalances from traditional SCM are substantial relative to the size of the estimated effects.

	Violent Crime	Part I Crime	Shots Fired and Violent Crime Calls
Mean Imbalance	0.16 (2.5%)	0.5 (12.5%)	1.01 (34.1%)

Table S1: Estimated average imbalance between the SCM and ASCM models for the *violent crime*, *Part I crime*, and *shots fired and violent crime* count outcomes. For reference, the ratio of the imbalance to the ASCM effect estimate is provided as a percentage in parentheses.

S2.3 Assumption: Stable Unit Treatment Value Assumption (SUTVA)

Under the potential outcomes framework, we require the SUTVA (Rubin, 1974) that there be no interference between units, meaning the treatment status of unit i cannot affect the potential outcomes of unit j, for $j \neq i$. In the present case, no interference between treated units and non-adjacent control units is reasonable to assume (Ridgeway, Grogger, Moyer, and MacDonald, 2019). Contagious spread of the benefits of deterrence, or of crime events (e.g., via displacement) are thought to be very local processes (Weisburd, Wyckoff, Ready, Eck, Hinkle, and Gajewski, 2006; Loeffler and Flaxman, 2017). We cannot exclude, however, interference between treated units, or between treated units and their adjacent controls. First, despite the spatial separation of the PHDs, CSP deployments in Nickerson Gardens and Jordan Downs shared a common LAPD command structure. Officers and staff could have discussed operations across PHDs, causing joint treatment effects across both units. We must assume this effect is negligible.

Second, and of primary concern, treated units may interfere with neighboring, non-CSP regions. Crime displays patterns of local contagion (Mohler, Short, Brantingham, Schoenberg, and Tita, 2011) and hot spot policing experiments have shown diffusion of benefits over relatively short spatial distances (Bowers, Johnson, Guerette, Summers, and Poynton, 2011). Therefore, we evaluate potential crime displacement by estimating the effect of CSP on control units neighboring the Jordan Downs PHD.

These displacement analyses are restricted by the available data. First, some neighbors of Jordan Downs and Nickerson Gardens are under LA Sheriff's Department jurisdiction. Our analyses are limited to only those spillover regions within LAPD jurisdiction. Second, Nickerson Gardens is a short distance from Gonzaque Village and therefore the two PHDs share potential spillover regions. Due to these shared potential spillover regions, we would not be able to attribute any observed effect to Nickerson Gardens as opposed to Gonzaque Village. To address this, the investigation of displacement effects is solely focused on Jordan Downs and should be considered exploratory in regards to the overall displacement effect of CSP.

To test displacement from Jordan Downs, we consider neighboring regions of Jordan Downs that are within two distance thresholds, approximately 890 ft (270 m) and approximately 1,640 ft (500 m), from the borders of the PHD. For each distance, we redefine neighborhoods within the threshold as a new, psuedo-treated *block group* unit. True treated regions are removed from consideration as either treated or control units. For a map of the Jordan Downs displacement regions, see Figure S3. We then estimate the treatment effects of these pseudo-treatment regions, which neighbor the treated PHD, but were not, themselves, directly treated. Evidence of an effect in these buffer regions would indicate possible spillover effects of the CSP.

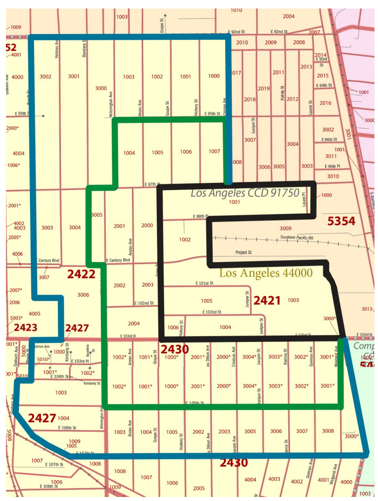

Figure S3: Displacement regions for Jordan Downs (black) are considered for Census regions within approximately 270 m (green) and approximately 500 m (blue). The orange and green shaded regions to the right of Jordan Downs are not within LAPD jurisdiction.

S2.3.1 CSP Does Not Displace Crime to Neighboring Regions

In terms of displacement, we are interested in evaluating whether the observed reduction of violent crime and disorder in the Jordan Downs and Nickerson Gardens PHDs is simply transferred to neighboring regions with a less intense police presence.

Type	Distance	Violent Crime	Part I Crime	Shots Fired and Violent Crime Calls
Raw	270 m	-2.96	-1.96	2.34
Per Capita*	270 m	-3.47	-3.71	-2.44
Raw	500 m	6.62	23.21	14.34
Per Capita*	500 m	-2.08	-1.63	-2.25

Table S2: Estimated ATT for regions within each distance threshold (270, 500 m) from Jordan Downs. To compare the per capita results with the count results, we approximate the count estimates with row "Per Capita*" where we multiply the crime per capita (by 1000 residents) by 2.714 (for the 2714 residents in Jordan Downs in 2010 census counts).

Figures S4 and S5 provide the estimated crime and calls trajectories for the 270 m and 500 m distance thresholds, respectively. Using the available LAPD data, we do not find evidence of crime displacement from Jordan Downs to neighboring regions within 270 and 500 m. Preliminary evidence suggests that *violent crime* and *Part I crime* incidents may have declined slightly, compared to synthetic controls, in areas within 270 m around Jordan Downs. In terms of raw counts and per capita counts, respectively, there were 2.96 fewer or 3.47 fewer reported *violent crime* incidents in the displacement region 270 m from Jordan Downs and 1.96 or 3.71 fewer reported *Part I crime* incidents in that same region. The *shots fired and violent crime* outcomes indicate a null effect in the 270 m buffer region. Therefore, we conclude there is no evidence for crime displacement and tentative evidence for a small diffusion of benefits in *violent crime* and *Part I crime* to areas within 270 m of Jordan Downs.

The 500 m displacement results in Figure S5 are inconclusive. Synthetic control models for those outcomes do not achieve suitable pre-treatment balance before the implementation date of 2012. Instead of finding practically negative effects of crime displacement (increase in crime in neighboring areas), these results could suggest no effect of CSP in neighboring regions. Evidence for potential diffusion of benefits as a result of CSP is only available at the 270 m distance threshold. ATT estimates for the effect of CSP in these regions are provided in Table S2.

S3 Robustness to Alternative Specifications

In this section, we evaluate the robustness of our results to per capita outcomes (per 1000 persons in 2010 census counts[1]) and the *quarter* time specification. This approach is compared to the raw count outcomes and *semester* time specification chosen in the main paper body. For ease of comparison, we adjust the *quarter*-level to the *semester*-level and adjust the per capita by 1000 persons to the raw population counts. These adjustments are denoted with the "*" in Table S3 and the following discussion.

As shown in Table S3, the results from these two robustness checks are largely consistent with our original estimates. The raw count *violent crime* estimates for *semester* and *quarter** indicate an average semester reduction in violent crime of roughly seven reported incidents; the raw count *Part I crime* estimates indicate an average decrease of roughly six incidents; the raw count *shots fired and violent crime* units indicate an average reduction of roughly three calls-for-service.

The per capita estimates for the three outcomes indicate substantively stronger positive effects, a stronger reduction in crime. For *violent crime* incidents, the *per capita* quarter** analyses indicate an average semester reduction of 11.19 incidents. The *per capita* semester* indicates an average semester reduction of 17.66 incidents. The *Part I crime* incidents per capita models estimate effects of 29.23 and 27.98 for *per capita* quarter** and *per capita* semester*, respectively. Similarly, the *shots fired and violent crime* calls models indicate average reductions of 4.66 and 10.92 for *per capita* quarter** and *per capita* semester*, respectively.

It should be noted the per capita models have weaker placebo results than the raw count models, hence the focus on raw count models in the main text. However it is encouraging that all model specifications indicate consistently negative, substantively positive, effects

[1] US Census Bureau: 2010 Census, Summary File 1, Table P1: Total Population.

of CSP.

Type	Time	Violent Crime	Part I Crime	Shots Fired and Violent Crime Calls
Raw	Semester	-6.55	-4.04	-2.96
Raw	Quarter*	-6.79	-7.16	-2.34
Per Capita*	Quarter*	-11.19	-29.23	-4.66
Per Capita*	Semester	-17.66	-27.98	-10.92

Table S3: Count and per capita ATT estimates for the Jordan Downs and Nickerson Gardens PHDs. To compare the per capita results with the count results, we approximate the count estimates with row "Per Capita*" where we multiply the crime per capita (by 1000 residents) by 6.719 (for the 6719 residents in these PHDs in 2010 census counts). We adjust the quarter results to the semester level with "Quarter*" where quarter estimates are multiplied by two.

S4 Spatial Analyses

As introduced in (Abadie, Diamond, and Hainmueller, 2010) and (Abadie, Diamond, and Hainmueller, 2015), constructing a set of placebo effects across space can also be informative in contextualizing the results. For each control unit, we assign treatment to the control and shift the original treated unit to the donor pool. We then estimate the effect of CSP on the psuedo-treated control unit during the period of study. The control units were never exposed to CSP treatment during this time period, and as such the ASCM models should not detect an effect of CSP.

By estimating placebo effects for the control units, we construct a distribution of placebo effects with which to compare the observed treatment effect and construct a p-value. If the effect of CSP on the treated units is outside of the distribution of placebo effects, we have more confidence in the observed treatment effect estimates.

In calculating the p-value, the root mean squared prediction error (RMSPE) is a commonly used test statistic, loosely defined as the ratio of the average psuedo-treated control fit over the post-treatment period to the average psuedo-treated control fit during the pre-treatment period:

$$RMSPE_j = \frac{\sum_{t=T_0+1}^{T}(Y_{jt} - \widehat{Y_{jt}(0)})^2/(T - T_0)}{\sum_{t=1}^{T_0}(Y_{jt} - \widehat{Y_{jt}(0)})^2/T_0} \tag{1}$$

As noted in (Abadie, Diamond, and Hainmueller, 2010), models with "poor" pre-treatment fit as compared to that of the treated unit should be removed from the analyses and p-value calculation. A poorly fitted model indicates an inability to construct an appropriate synthetic control with the available data. Therefore, after calculating the RMSPE and before calculating the p-value, psuedo-treated control models with "poor" pre-treatment fit are removed.

For our data, we define "poor" fit as an RMSPE that is five times greater than the observed RMSPE for the treated.[2] The p-value is then the proportion of control models with higher RMSPEs than the treated model using the reduced set of control models. This approach is akin to testing a null hypothesis of no effect on the outcome where the p-value is calculated by giving equal weight to the included control models (2018). This assumption may not be reasonable given the uniqueness of the PHD units as compared to the control units.

The results of the spatial robustness check are contained in Figure S6 and Table S4. From the figure, the effect for *violent crime* appears significant. The ATT_t estimates of the Jordan Downs and Nickerson Gardens PHDs exceeds many of the control ATT_t estimates across time for both of these outcomes. The resulting p-value for the reduced set is 0.39 ($\frac{61}{157}$) for *violent crime* under the adjusted control group. For comparison, the p-value constructed from the full set of controls for *violent crime* is 0.59 ($\frac{138}{234}$). The discrepancy between the figure and table may be due to the convex hull violation. The scale of the outcome for the treated units is more extreme than that of the controls. The effects on

[2]Regardless of the specification, the p-values derived from the ratios are not significant. We provide the unadjusted and adjusted visualizations and p-values to demonstrate this.

Part I crime and *shots fired and violent crime* appear insignificant both visually and in terms of the p-values.

Type	Violent Crime	Part I Crime	Shots Fired and Violent Crime Calls
Unreduced	0.59 ($\frac{138}{234}$)	0.96 ($\frac{225}{234}$)	1 ($\frac{234}{234}$)
Reduced	0.39 ($\frac{61}{157}$)	0.88 ($\frac{67}{76}$)	1 ($\frac{31}{31}$)

Table S4: The p-values, the proportion of control models with higher RMSPEs than the treated models, are provided for the *violent crime*, *Part I crime*, and *shots fired and violent crime* outcomes for the Jordan Downs and Nickerson Gardens PHDs.

S5 Additional Outcomes: Jordan Downs and Nickerson Gardens

In this section we introduce the *burglary* reported crime incidents and *quality of life* calls-for-service outcomes and provide the corresponding placebo results for the Jordan Downs and Nickerson Gardens PHDs.

The outcomes are defined as follows (LAPD Consolidated Crime Analysis Database (CCAD) code in parentheses): *burglary* incidents are defined as burglary (310); *quality of life* calls-for service are defined as intoxication (390), disturbance (415), minor disturbance (507), vandalism (594), dispute (620), and screaming (930).

S5.1 Model Evaluation

As shown in Table S5, the placebo models estimate effects of 5.11 and -6.02 for the *burglary* incidents and *quality of life* calls-for-service outcomes, respectively. These placebo effects correspond to 33% and 3% of the pre-treatment semester averages of *burglary* and *quality of life*, respectively. These results indicate the Ridge ASCMs may achieve proper synthetic control fit for the *quality of life* outcome but not the *burglary* outcome.

Outcome	Pre-Treatment Average	ATT Estimate
Burglary Incidents	15.56	5.11
Quality of Life Calls	194.94	-6.02

Table S5: Estimated placebo impact of CSP for the Jordan Downs and Nickerson Gardens PHDs. To assess scale of the estimates, the pre-treatment semester average is provided for each outcome.

S5.2 Temporal

The temporal placebos for these additional outcomes indicate a large shift from the estimated ATT_t for the *quality of life* outcome at the 2010 psuedo-implementation date (Figure S7). Taken with the model evaluation results, we lack confidence in the synthetic control fit of both the *burglary* reported crime incidents and *quality of life* calls-for-service outcomes.

S6 Jordan Downs, Nickerson Gardens, and Imperial Courts

In this section, results for the alternative PHD specification of Jordan Downs, Nickerson Gardens, and Imperial Courts are provided. These PHDs were all implemented in late 2011, treatment start is approximated with the start of 2012.

S6.1 Model Evaluation

As shown in Table S6, the *violent crime* outcome placebo estimate is -3.72, approximately 14% of the pre-treatment semester average; the *Part I crime* outcome placebo estimate is 2.60, approximately five % of the pre-treatment semester average; the *shots fired and violent crime* estimate is -0.45, approximately one % of the pre-treatment semester average. Similarly, the *burglary* outcome placebo estimate is 2.40, approximately 18% of the pre-treatment semester average; the *quality of life* outcome placebo estimate is -21.01, approximately 12% of the pre-treatment semester average.

These results suggest the *violent crime*, *Part I crime*, *shots fired and violent crime*, and *quality of life* models estimate placebo effects on a relatively low scale as compared to their pre-treatment semester averages.

Outcome	Pre-Treatment Average	ATT Estimate
Violent Crime Incidents	26.19	-3.72
Burglary Incidents	13.04	2.40
Part I Crime Incidents	48.22	2.60
Shots Fired and Violent Crime Calls	38.33	-0.45
Quality of Life Calls	172.81	-21.01

Table S6: Estimated placebo impact of CSP for the Jordan Downs, Nickerson Gardens, and Imperial Courts PHDs. To assess scale of the estimates, the pre-treatment semester average is provided for each outcome.

S6.2 Temporal

The Jordan Downs, Nickerson Gardens, and Imperial Courts temporal analyses presented in Figures S8 and S9 largely do not indicate strong confounding events in the pre-treatment period. The *quality of life* outcome is the exception with a large deviance in the 2010 psuedo-implementation model that is still present in the 2011 model. For the remaining outcomes, the psuedo-treatment ATT_t estimates closely follow the actual ATT_t estimate for both outcomes.

As in the count results, we focus on *violent crime*, *Part I crime*, and *shots fired and violent crime* in the remainder of the analyses due to their consistent and strong placebo results.

S6.3 Distribution of Placebo Effects Across Space

Results for the spatial robustness check for Jordan Downs, Nickerson Gardens and Imperial Courts are contained in Figure S10 and Table S7. From the figure, the effect on *violent crime* appears moderately significant relative to the scale of the psuedo effects. Conversely, the p-values are insignificant for all of the outcomes with p-values of 0.21 ($\frac{37}{180}$), 0.85 ($\frac{77}{91}$), and 0.91 ($\frac{51}{66}$) for the *violent crime*, *Part I crime*, and *shots fired and violent crime* outcomes reduced sets, respectively. Therefore, we fail to reject null hypotheses of no effect for both outcomes.

Type	Violent Crime	Part I Crime	Shots Fired and Violent Crime Calls
Unreduced	0.39 ($\frac{91}{234}$)	0.94 ($\frac{220}{234}$)	0.98 ($\frac{229}{234}$)
Reduced	0.21 ($\frac{37}{180}$)	0.85 ($\frac{77}{91}$)	0.91 ($\frac{51}{66}$)

Table S7: The p-values, the proportion of control models with higher RMSPEs than the treated models, are provided for the *violent crime*, *Part I crime*, *shots fired and violent crime* outcomes for the Jordan Downs, Nickerson Gardens, and Imperial Courts PHDs.

S6.4 Results

The final estimates for Jordan Downs, Nickerson Gardens, and Imperial Courts appear similar to the final estimates for Jordan Downs and Nickerson Gardens. To compare against the original estimates of -6.55, -4.04, and -2.96 for *violent crime*, *Part I crime*, and *shots fired and violent crime*, the updated estimates for the three PHDs are -6.79, -7.56, and -4.45 for the three outcomes, respectively (Table S8). The ATT_t estimates (Figure S11) also follow a similar trend as compared to the Jordan Downs and Nickerson Gardens PHDs estimates. The joint null p-values are 0.76, 0.61, and 0.84 for the *violent crime*, *Part I crime*, and *shots fired and violent crime* outcomes, respectively.

Outcome Type	Violent Crime	Part I Crime	Shots Fired and Violent Crime Calls
Raw	-6.79	-7.56	-4.45

Table S8: Estimated impact of CSP for the Jordan Downs, Nickerson Gardens, and Imperial Courts PHDs.

S7 Avalon Gardens and Gonzaque Village

In this section, we present placebo analyses for the Avalon Gardens and Gonzaque Village PHDs. CSP was implemented in these PHDs in mid-2016, therefore the post-treatment period consists of three periods using the available data. Final results for these analyses are preliminary in nature due to the short post-treatment period and not formally presented in this document.

S7.1 Model Evaluation

Semester averages for *violent crime*, *Part I crime*, and *shots fired and violent crime* in the pre-treatment period for Avalon Gardens and Gonzaque Village are substantially lower than those of the previous PHDs (Table S9). Therefore, the model evaluation placebo estimates of -2.49 for *violent crime* and -2.97 for *Part I crime* are large compared to the pre-treatment semester averages of 4.32 and 9.24, respectively. The *shots fired and violent crime* model appears well-balanced, the estimated placebo effect is eight percent of the pre-treatment semester average.

We now consider the additional outcomes of *burglary* reported crime incidents and *quality of life* calls-for-service (Table S9). The model evaluation ASCMs estimate placebo effects of -1.58 and -0.64 for *burglary* and *quality of life*, respectively. These placebo effects are 65% and 1.5% of the corresponding pre-treatment semester averages for *burglary* and *quality of life*, respectively. This indicates the *quality of life model* may be able to construct a suitable synthetic control.

Outcome	Pre-Treatment Average	ATT Estimate
Violent Crime Incidents	4.32	-2.49
Burglary Incidents	2.42	-1.58
Part I Crime Incidents	9.24	-2.97
Shots Fired and Violent Crime Calls	8.71	-0.68
Quality of Life Calls	43.63	-0.64

Table S9: Estimated placebo impact of CSP for the Avalon Gardens and Gonzaque Village PHDs. To assess scale of the estimates, the pre-treatment semester average is provided for each outcome.

S7.2 Temporal

The Avalon Gardens and Gonzaque Village PHDs temporal placebos closely follow the final ATT_t in the pre-treatment period (Figures S12 and Figure S13). The estimated pre-treatment ATT_t estimates for *violent crime*, *Part I crime*, *shots fired and violent crime*, *burglary*, and *quality of life* appear well-balanced across the pre-treatment period, closely following zero.

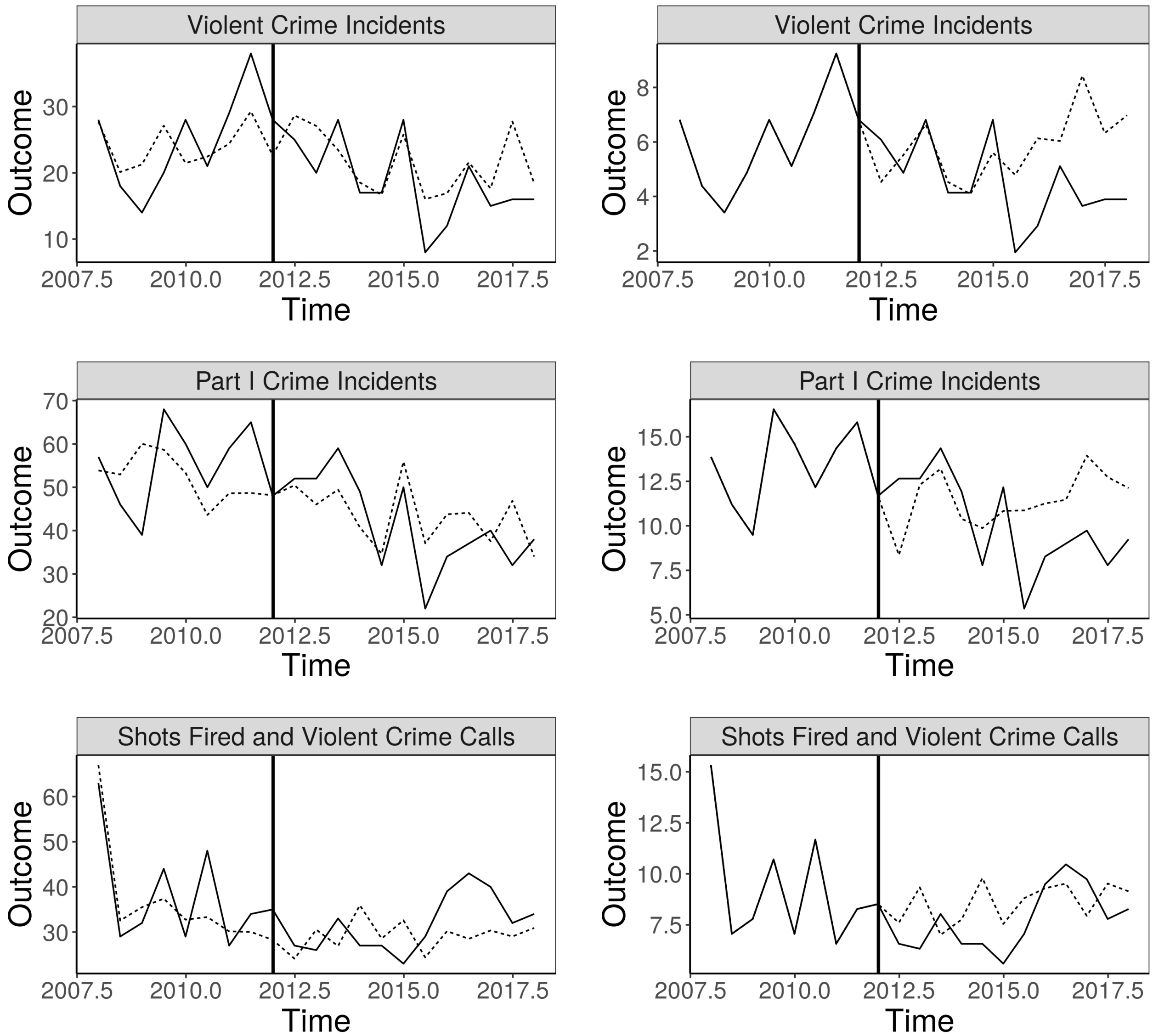

Figure S4: Crime outcome trajectories for Jordan Downs (solid) versus the estimated synthetic control (dashed) do not indicate crime displacement between Jordan Downs and adjacent controls within the 270 m distance threshold. Left to right: count outcomes, per capita outcomes.

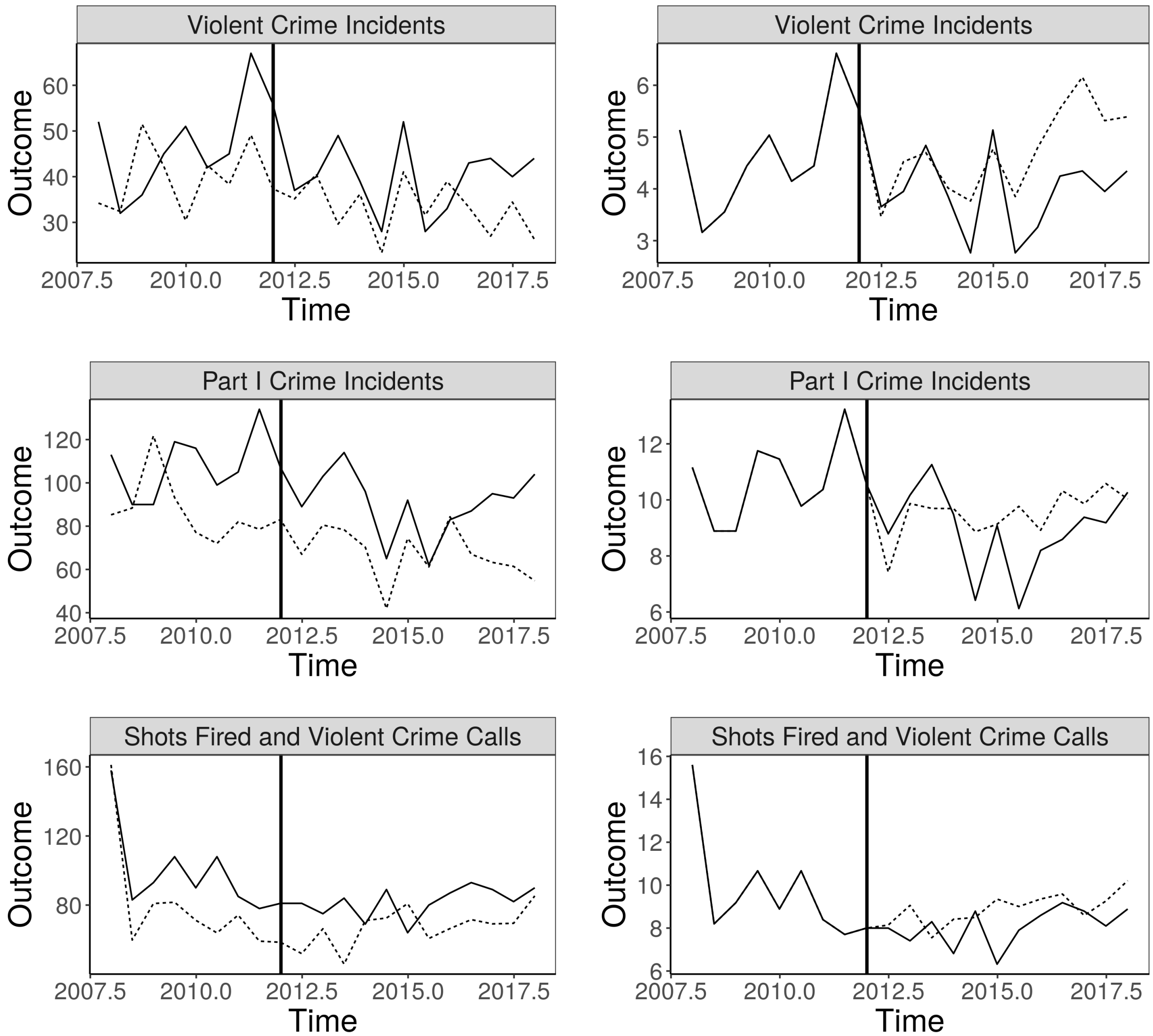

Figure S5: Crime outcome trajectories for Jordan Downs (solid) versus the estimated synthetic control (dashed) do not indicate crime displacement between Jordan Downs and adjacent controls within the 500 m distance threshold. Left to right: count outcomes, per capita outcomes.

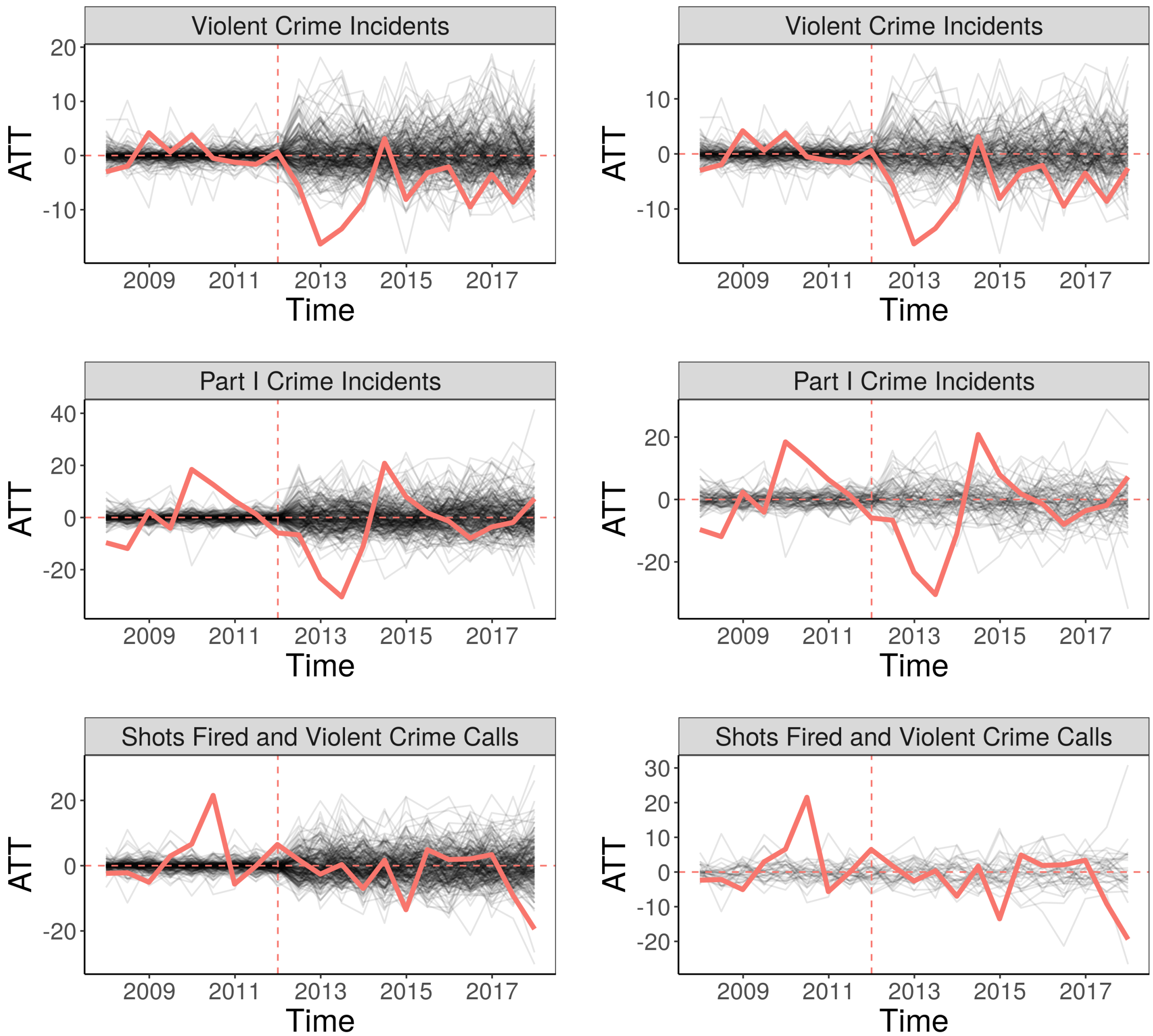

Figure S6: Spatial placebos for the Jordan Downs and Nickerson Gardens PHDs for the count outcomes. The comparative scale of the ATT_t effect appears moderately significant for the violent crime incidents outcome, while the ratio measure of the post-pre effect is insignificant. Left: unreduced control group. Right: reduced control group.

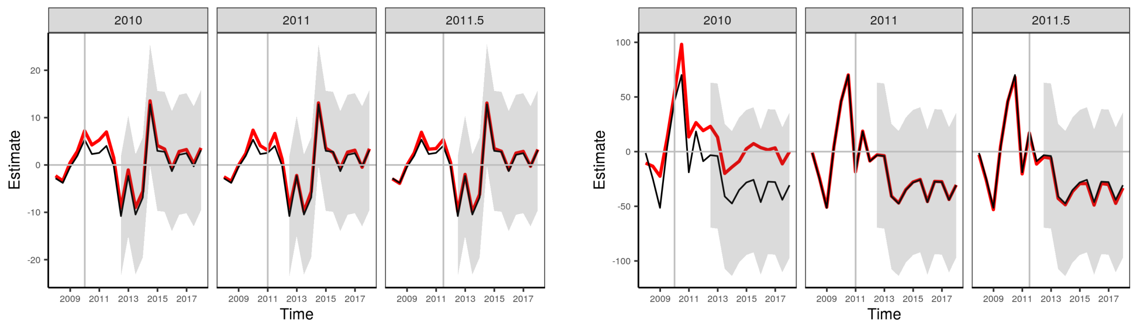

Figure S7: We evaluate robustness across time for Jordan Downs and Nickerson Gardens models of burglary and quality of life (from left to right) by comparing psuedo-treatment ATT_t estimates (red) to the ATT_t estimated using the true implementation date of 2012 (black line with shaded conformal inference bounds). For each panel, the psuedo-implementation model is marked in the panel title. In the absence of confounding events during the pre-treatment period, we would expect to see the pseudo-implementation ATT_t estimates closely follow the ATT_t estimated from the true treatment period.

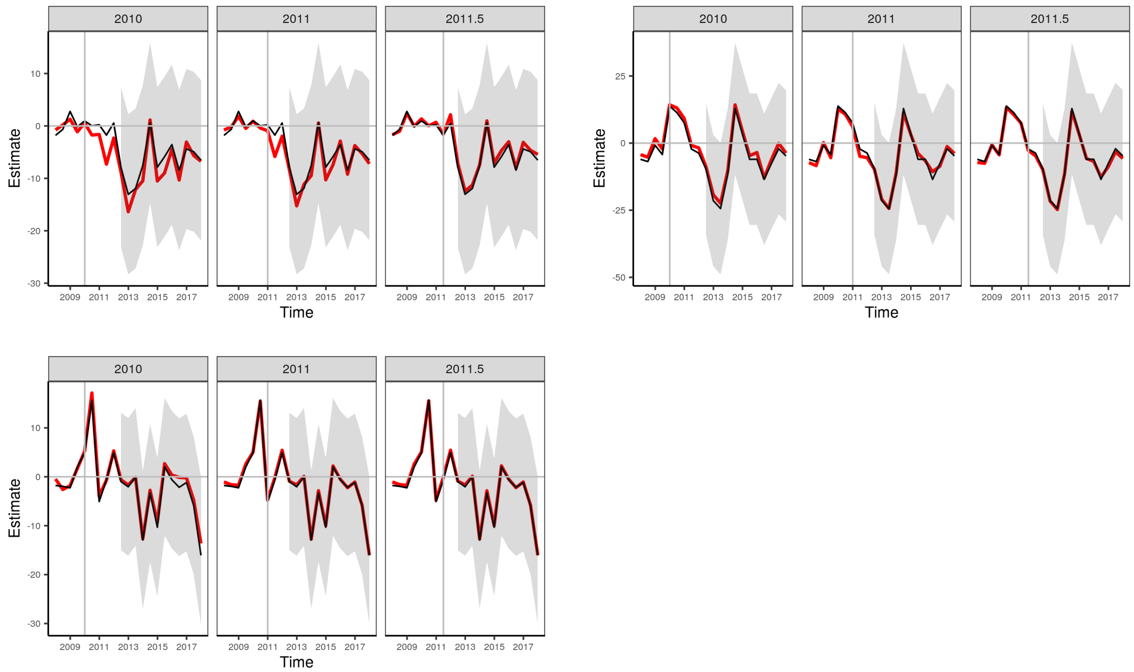

Figure S8: We evaluate robustness across time for Jordan Downs, Nickerson Gardens, and Imperial Courts models of violent crime incidents, Part I crime incidents, and shots fired and violent crime calls (from left to right, top to bottom) by comparing psuedo-treatment ATT_t estimates (red) to the ATT_t estimated using the true implementation date of 2012 (black line with shaded conformal inference bounds). For each panel, the psuedo-implementation model is marked in the panel title. In the absence of confounding events during the pre-treatment period, we would expect to see the pseudo-implementation ATT_t estimates closely follow the ATT_t estimated from the true treatment period.

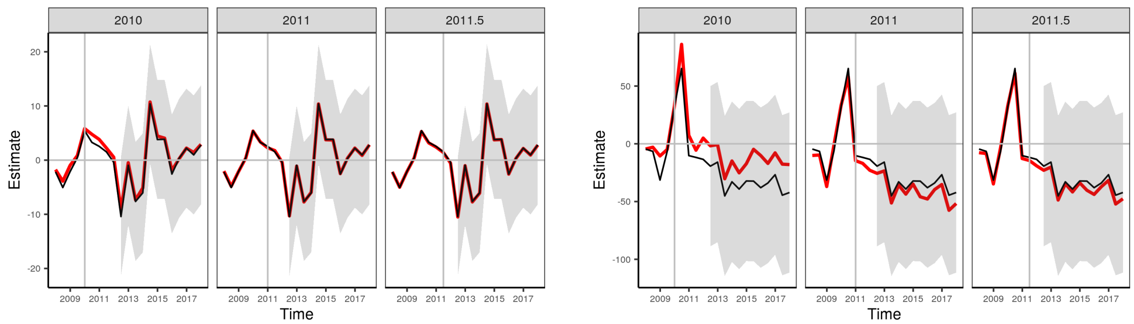

Figure S9: We evaluate robustness across time for Jordan Downs, Nickerson Gardens, and Imperial Courts models of burglary and quality of life (left to right) by comparing psuedo-treatment ATT_t estimates (red) to the ATT_t estimated using the true implementation date of 2012 (black line with shaded conformal inference bounds). For each panel, the psuedo-implementation model is marked in the panel title. In the absence of confounding events during the pre-treatment period, we would expect to see the pseudo-implementation ATT_t estimates closely follow the ATT_t estimated from the true treatment period.

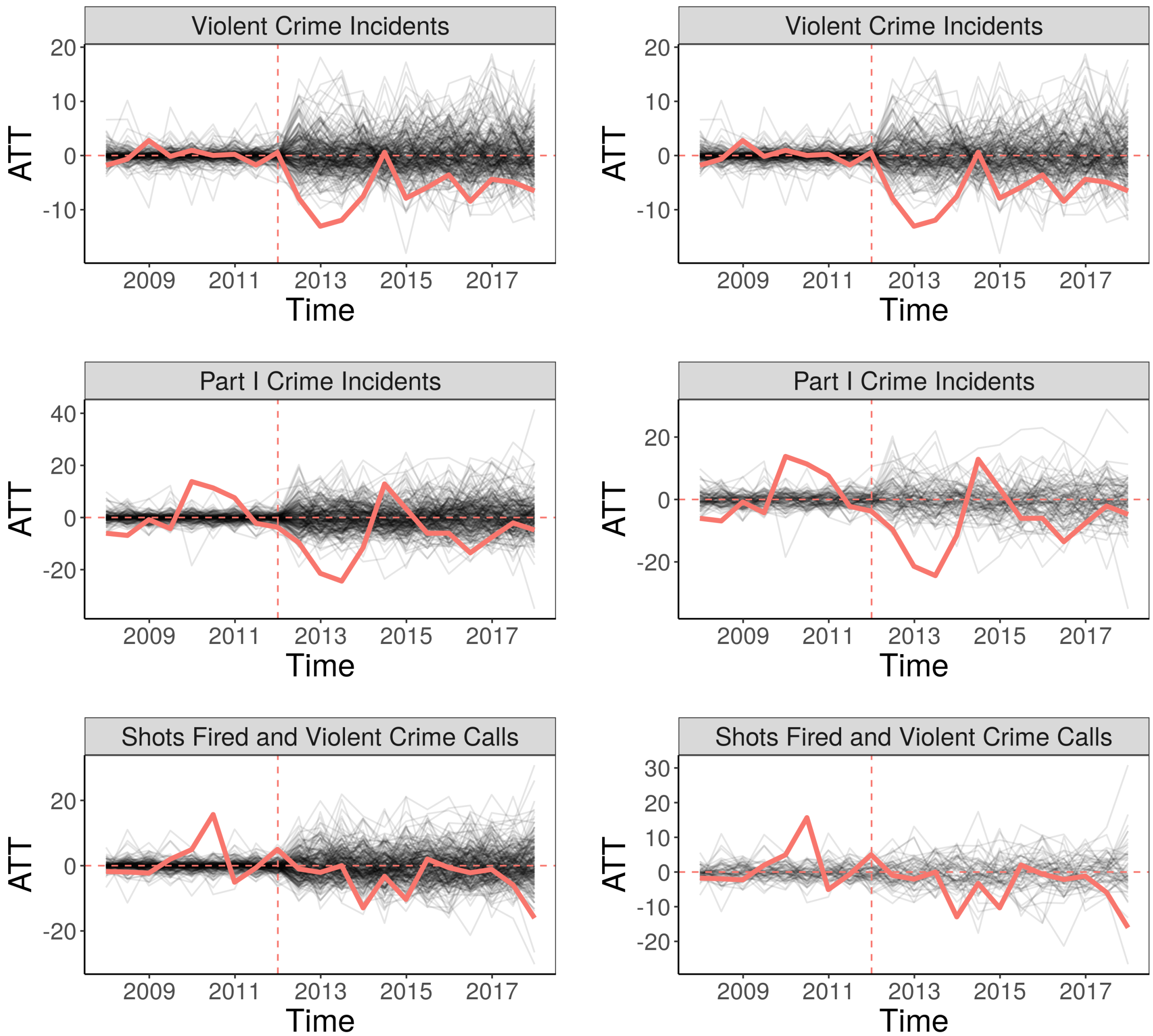

Figure S10: Spatial placebos for the Jordan Downs, Nickerson Gardens, and Imperial Courts count outcomes. The comparative scale of the ATT effect appears moderately significant for the violent crime outcome, while the ratio measure of the post-pre effect is insignificant. Left: unreduced control group. Right: reduced control group.

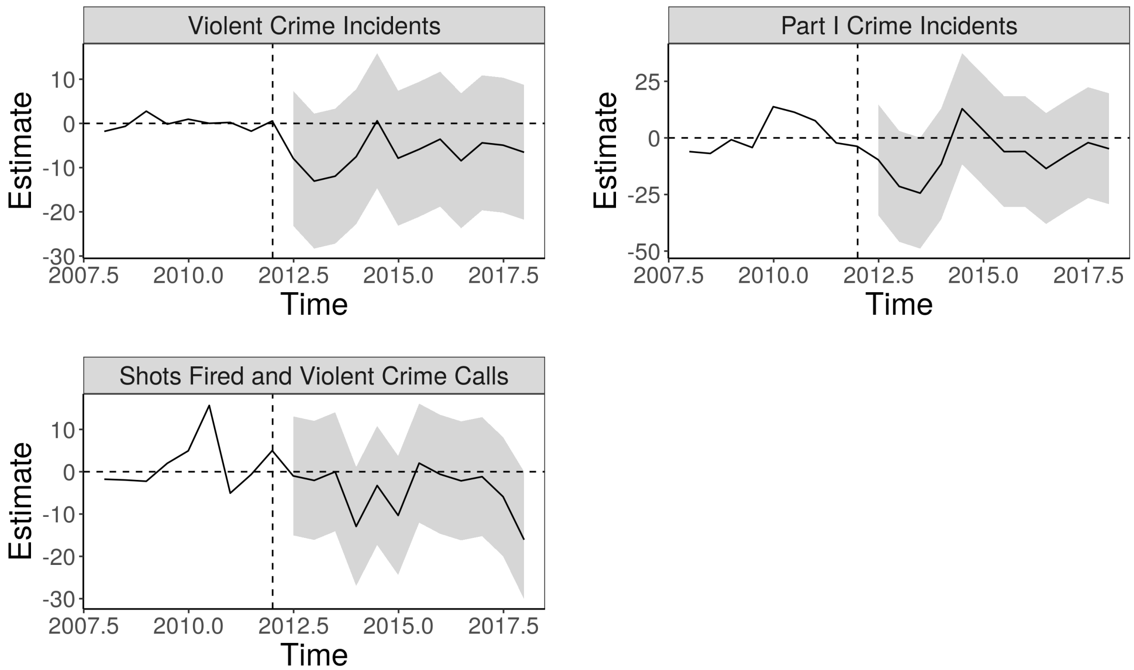

Figure S11: The ATT_t estimates for the Jordan Downs, Nickerson Gardens, and Imperial Courts PHDs are provided with shaded conformal inference bounds. 2012 treatment implementation date is denoted with vertical line. Left to right, top to bottom: *violent crime*, *Part I crime*, *shots fired and violent crime*.

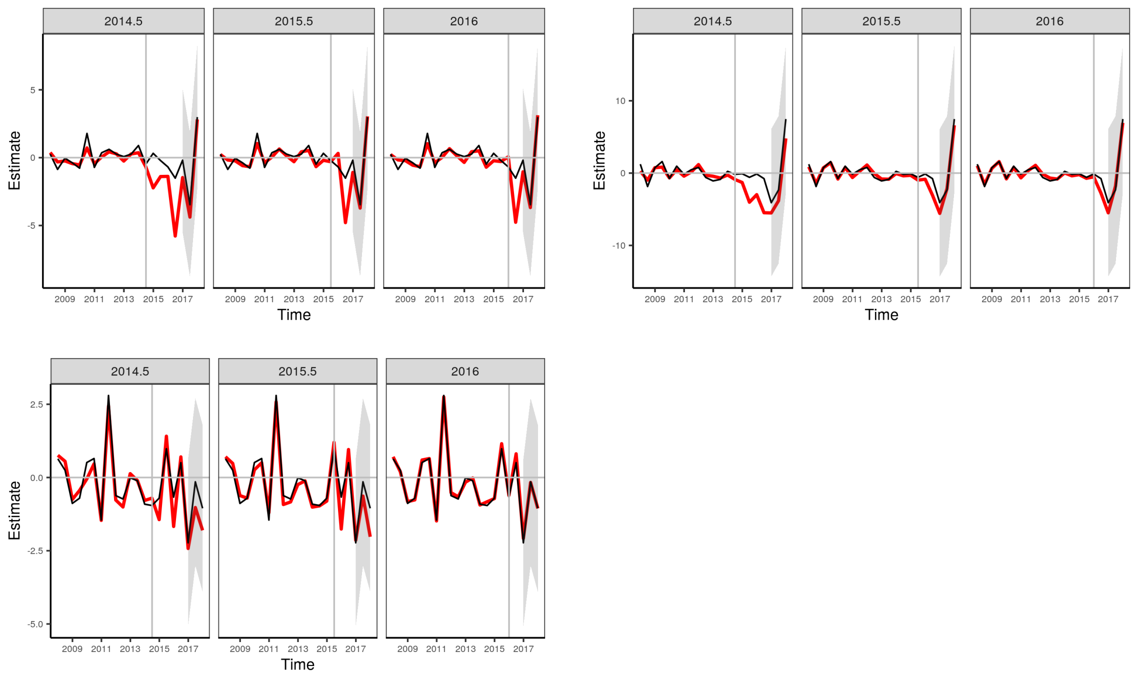

Figure S12: We evaluate robustness across time for Avalon Gardens and Gonzaque Village models of violent crime incidents, Part I crime incidents, and shots fired and violent crime calls (from left to right, top to bottom) by comparing psuedo-treatment ATT_t estimates (red) to the ATT_t estimated using the true implementation date of 2016.5 (black line with shaded conformal inference bounds). For each panel, the psuedo-implementation model is marked in the panel title. In the absence of confounding events during the pre-treatment period, we would expect to see the pseudo-implementation ATT_t estimates closely follow the ATT_t estimated from the true treatment period.

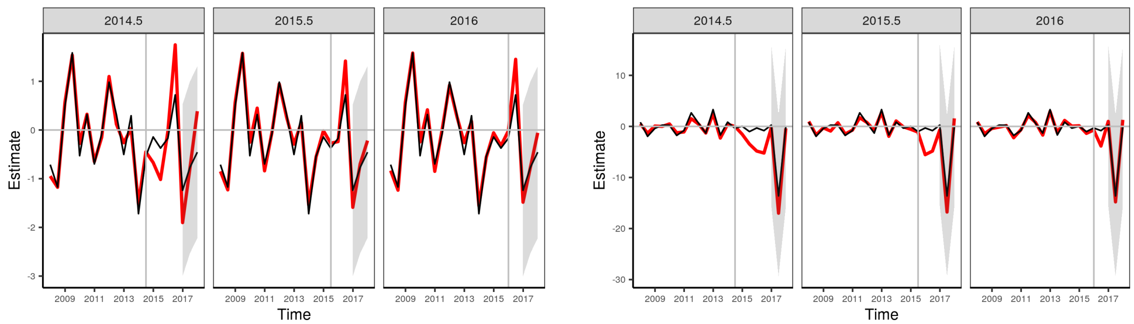

Figure S13: We evaluate robustness across time for Avalon Gardens and Gonzaque Village models of burglary and quality of life (left to right) by comparing psuedo-treatment ATT_t estimates (red) to the ATT_t estimated using the true implementation date of 2016.5 (black line with shaded conformal inference bounds). For each panel, the psuedo-implementation model is marked in the panel title. In the absence of confounding events during the pre-treatment period, we would expect to see the pseudo-implementation ATT_t estimates closely follow the ATT_t estimated from the true treatment period.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program". In: *Journal of the American Statistical Association* 105.490, pp. 493–505.

— (2015). "Comparative politics and the synthetic control method". In: *American Journal of Political Science* 59.2, pp. 495–510.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein (2019). "The Augmented Synthetic Control Method". In:

Bowers, Kate et al. (2011). "Spatial displacement and diffusion of benefits among geographically focused policing initiatives". In: *Campbell Systematic Reviews* 7.1, pp. 1–144.

Firpo, Sergio and Vitor Possebom (2018). "Synthetic control method: Inference, sensitivity analysis and confidence sets". In: *Journal of Causal Inference* 6.2.

Loeffler, Charles and Seth Flaxman (2017). "Is gun violence contagious? A spatiotemporal test". In: *Journal of Quantitative Criminology*, pp. 1–19. ISSN: 0748-4518.

Mohler, George O et al. (2011). "Self-exciting point process modeling of crime". In: *Journal of the American Statistical Association* 106.493, pp. 100–108.

Ridgeway, Greg et al. (2019). "Effect of gang injunctions on crime: A study of Los Angeles from 1988–2014". In: *Journal of Quantitative Criminology* 35.3, pp. 517–541.

Rubin, Donald B (1974). "Estimating causal effects of treatments in randomized and nonrandomized studies." In: *Journal of Educational Psychology* 66.5, p. 688.

Weisburd, D. et al. (2006). "Does Crime Just Move Around the Corner? A Controlled Study of Spatial Displacement and Diffusion of Crime Control Benefits". In: *Criminology* 44.3, pp. 549–592.