The Philadelphia Foot Patrol Experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots

Final draft of:

Originating with the Newark foot patrol experiment, research has found police foot patrols improve community perception of the police and reduce fear of crime, but are generally unable to reduce the incidence of crime. Previous tests of foot patrol have, however, suffered from statistical and measurement issues and have not fully explored potential dynamics of deterrence within micro-spatial settings. In this paper we report on the efforts of over 200 foot patrol officers during the summer of 2009 in Philadelphia. GIS analysis was the basis for a randomized controlled trial of police effectiveness across 60 violent crime hotspots. Results identified a significant reduction in the level of treatment area violent crime after 12 weeks. A linear regression model with separate slopes fitted for treatment and control groups clarified the relationship further. Even after accounting for natural regression to the mean, target areas in the top 40% on pre-treatment violent crime counts had significantly less violent crime during the operational period. Target areas outperformed the control sites by 23 percent, resulting in a total net effect (once displacement was considered) of 53 violent crimes prevented. The results suggest that targeted foot patrols in violent crime hotspots can significantly reduce violent crime levels as long as a threshold level of violence exists initially. The findings contribute to a growing body of evidence on the contribution of hotspots and place-based policing to the reduction of crime, and especially violent crime, a significant public health threat in the United States. We suggest that intensive foot patrol efforts in violent hotspots may achieve deterrence at a micro-spatial level, primarily by increasing the certainty of disruption, apprehension and arrest. The theoretical and practical implications for violence reduction are discussed.

For most of the history of American policing, foot patrol’s role in public safety has been almost mythical. The growth of the night and rattle watches of the 1700s was the consequence of the assumed deterrence abilities of a patrolling, uniformed authority carrying the explicit threat of government intervention should social order unravel. To this day, there is a consistent public demand for foot patrols as a ‘proactive, non-threatening, community-oriented approach to local policing’ (Wakefield, 2007: 343). Key questions yet remain. For example, do foot patrols achieve more than simply providing reassurance to the public? Does the enhanced visibility of officers on foot, instead of in cars, serve a significant and measurable deterrent effect? The evidence to date on these questions has been mixed, despite the fact that the police have long been assumed to provide a deterrence function. This assumption can be traced back to the writings of both Bentham and Beccaria who argued for the need to influence the calculus of would-be criminals, for society to ensure that the costs of committing a crime would be outweighed by any potential benefits. Beccaria argued that the central mechanisms for adjusting this calculus are
certainty of detection, severity of punishment, and celerity (or swiftness of punishment) (see Nagin and Pogarsky, 2001). The very origins of the police institution rest on this view. Sir Robert Peel established his police in London as a means of providing an ‘unremitting watch’ (Shearing, 1996: 74) through visible patrol. Citizens would be deterred through this system of surveillance, knowing that their chances of being caught and punished would be high.

Yet, despite the longevity of the deterrence doctrine, the evidence on whether the practice of foot patrol actually deters crime has been weak. Following on from the Kansas City Preventative Patrol Experiment finding that vehicle-based patrol had no significant impact on crime rates (Kelling et al., 1974), the Newark Foot Patrol Experiment did much to cement the view among many criminologists that varying the dosage of uniformed patrol has no quantifiable impact on crime (Kelling, 1981). Varying foot patrol levels across 12 Newark, New Jersey beats resulted in no significant differences between treatment and control beats for recorded crime or arrest rates, though treatment areas did show improvements in community fear of crime (Pate, 1986).

Further studies followed, ranging in magnitude and scope. For example, four foot patrol officers in a business district of Asheville, NC had the same, apparently negligible, impact on recorded crime as the 300 officers moved to foot patrol as part of the Boston Police Department’s 1983 Patrol Reallocation Plan (Bowers and Hirsch, 1987; Esbensen, 1987). Notwithstanding this lack of evidence, foot patrol became “the most popular and widely implemented component of community policing” (Rosenbaum and Lurigo, 1994: 303) even if many police departments adopted foot beats more to address community relations and fear of crime than for any direct crime deterrence benefits (Cordner, 1986; Jim, Mitchell, and Kent, 2006). The National Research Council review of police policy and practices summarized foot patrol as an unfocused community policing strategy with only weak to moderate evidence of effectiveness in reducing fear of crime (National Research Council, 2004).

Since these early foot patrol studies, criminologists have gained a more nuanced understanding of criminal behavior within spatial and temporal contexts. For instance, both Routine Activity Theory (Felson, 1987) and Crime Pattern Theory (Brantingham and Brantingham, 1984) identify place as a fundamental component of the requirements of a crime, the centrality of which environmental criminologists have adopted as a potential avenue along which to promote crime control opportunities. It is now widely understood that crime clusters within highly specific geographic locations, commonly termed ‘hotspots’¹. A crime hotspot is the accepted term for what was originally described as a cluster of addresses (Sherman and Weisburd, 1995), widened to include the possibility of street intersections and public space (Buerger, Cohn, and Petrosino, 1995). The term is now generally defined as a “geographical area of higher than average crime... an area of crime concentration, relative to the distribution of crime across the whole region of interest” (Chainey and Ratcliffe, 2005: 145-6). With the growth of crime mapping, crime hotspots have become significant loci for focused police activity.

With a refocusing on place, location-specific crime prevention can add to general offender deterrence with options to prevent potential offenders committing crime at a specific location. Nagin (2010: 313) recently pointed out that effective deterrence stems from a tangible and direct prospect of detection, and that focused policing at crime hotspots “is probably effective because it tangibly and directly increases apprehension risk at the hot spot by substantially increasing police presence.”

While the National Research Council’s (2004) review lamented the paucity of quality studies on the benefit of proactive police activity such as field interrogations and traffic enforcement, there has long

¹ The research and professional literature refers to both ‘hotspots’ and ‘hot spots’. We use the former throughout simply for purposes of consistency.
been general support from Wilson and Boland’s study of 35 cities to suggest that even some unfocused proactive police activity\(^2\) can have a reductive effect on robbery (Wilson and Boland, 1978). A more extensive study of 171 American cities and the proactive drink/drive and disorder activities of police again found a similar dampening effect on robbery (Sampson and Cohen, 1988), and a recent update with a more fully specified statistical model again found a significant negative association between robbery rates and proactive policing across a similar number of US cities (Kubrin et al., 2010). Focusing on gun violence, studies including the Kansas City gun intervention (Sherman, Shaw & Rogan, 1995) and the Indianapolis directed patrol project (McGarrell, Chermak, & Weiss, 2002) lead Koper and Mayo-Wilson (2006) to conclude that directed patrols targeted to the carrying of illegal weapons had a suppressive effect on gun violence at high-risk places and times.

A strong evidence base has similarly emerged in relation to the positive effects of the related strategy, hotspots policing. Echoing the findings of previous studies (such as Braga and Bond, 2008; Sherman, Gartin, and Buerger, 1989; Weisburd and Braga, 2006), both the National Research Council (2004) and Braga’s (2007) systematic review concludes that focused hotspots policing works. Previous hotspots experiments have to date examined problem-oriented policing rather than foot patrol per se (Braga et al., 1999), or where foot patrol strategies were mixed with other interventions such as vehicle patrols (Sherman and Weisburd, 1995; Weisburd and Green, 1995). A rare exception is the British study in Hull, Humberside, where additional foot patrols in the city center reduced personal robbery over the course of a year by 16 percent while regional and national rates increased (Jones and Tilley, 2004). Given this new evidence, we suggest it is timely to re-examine the question of whether foot patrol, as a specific hotspots intervention holds promise as an approach to reducing crime, and especially violent crime, a leading cause of death and injury in the United States (see Miller, Cohen, and Rossman, 1993).

In light of the theoretical advances above, and the development of new techniques in spatial analysis, one can re-visit the research designs of earlier foot patrol studies with fresh eyes. Sherman and Weisburd (1995) have already pointed out that many of these early studies suffered from statistical and measurement problems; namely, a statistical bias across area-based studies toward the null hypothesis, and the measurement issue of an often inappropriate study area. The latter problem addresses the question of whether to organize a project by police districts, police beats or other areas. Even if hotspots policing was part of the lexicon at the time, the ability to achieve a micro-spatial focus has traditionally been hampered by the need to measure and organize police resources by larger administrative regions.

This issue has to some extent been resolved with the development of Geographic Information Systems (GIS) and the accompanying field of Geographic Information Science (GISc), though as Rengert and Lockwood point out, many crime analysts simply accept the “bounded space that is available to them rather than construct their own boundaries” (Rengert and Lockwood, 2009: 110). GIS and GISc together provide both a tool and an analytical regime to approach spatially-customized target areas for crime prevention activities. This has enabled more recent police effectiveness research projects to concentrate on crime hotspots.

The ability over time to move down through the cone of resolution (Brantingham, Dyreson, and Brantingham, 1976) from studying large administrative areas to smaller and smaller spatial units, has enabled crime researchers to now explore crime hotspots at micro units of place, defined as addresses, street segments or clusters of these micro spatial units (Weisburd, Bernasco, and Bruinsma, 2009: 4). A

---

\(^2\) Defined broadly in this context as police activity that is not reactive to calls for service from the public but rather indicative of a police decision to take action where such activity could be considered discretionary; for example, to initiate a traffic stop for a minor traffic violation, to conduct a field interview, or to undertake surveillance of a known offender.
focus on smaller places can address dosage concerns; concerns that foot patrol officers assigned to replace vehicle-based patrol in large geographic areas will be spread too thin thereby diminishing any deterrence effect that could have been created by their presence.

Spatially-oriented crime control programs have actively addressed the re-designation of places that provide crime opportunities, looking to a locational focus to create constraints on criminality. Weisburd and Green (1995) employed a randomized control design to examine a seven month operation to reduce drug activity at drug hotspots in Jersey City, NJ, and found “consistent, strong effects of the experimental strategy on disorder-related emergency calls for service” (Weisburd and Green, 1995: 731). Taking the cone of resolution to individual properties and corners, Green (1995) found an Oakland, CA, program that combined traditional enforcement with third party interventions targeted at nuisance drug locations not only reduced drug problems but also demonstrated a diffusion of benefits to nearby locations.

The potential diffusion impact of crime prevention strategies at specific locations raises the question of how interventions such as foot patrols can prevent crime. General and specific deterrence may occur if the presence of a police officer is sufficient to increase an offender’s perceived risk of apprehension (Nagin, 2010). A second potential mechanism is ‘proactive policing’ (Kubrin et al, 2010), whereby the activity of a police officer such as stopping and questioning suspects, performing a stop-and-frisk (also known as a Terry stop), or (with probable cause) conducting a full search of a suspect, may increase the chances that police will identify a fugitive or find illegal weapons or items and increase the arrest rate. The visible enforcement of minor infractions and disorder offenses may be seen by offenders as indicative of a change in the apprehension risk, according to Sampson and Cohen (1988). Therefore deterrence can potentially occur through officer presence, or where specific activities of police officers either increase the arrest-offense ratio or the perception that it has increased (Kubrin et al, 2010).

Spatial diffusion of benefits may occur if offenders perceive that officers patrolling a nearby hotspot may be able to intervene quickly should the alarm be raised about a crime, or if patrol boundaries are not known to offenders. A spatial diffusion could also occur if deterrence can serve to discourage the carrying of crime-enabling items, a change that can affect the offender both in and outside the target area.

Conversely, place-based interventions can theoretically displace crime to nearby areas if officers never patrol nearby areas, and if the boundaries of the target area are known to local offenders. Yet even in these scenarios, displacement may be beneficial. Offenders may move to spaces that are less inviting or less familiar to them, resulting in a reduction of their activity. Specific behaviors like drug market activity could be displaced to less public spaces, away from children, recovering drug addicts and everyday people such that these groups are less exposed to the harms associated with dealing and selling (Caulkins and Reuter, 2009). The social harm outcomes of proactive police activity can therefore be theoretically beneficial in either a diffusion or displacement regime. Displacement can move criminal activity to less optimal (Taniguchi, Rengert, & McCord, 2009) or less public locations, while a diffusion of benefits could mean reduced exposure to violence overall, a crime reduction outcome which has also been associated with improved public health outcomes (Guerra, Huesmann, and Spindler, 2003).

At the outset we should note that disentangling specific deterrence effects of officer presence versus officer (proactive) activity is beyond the reach of this paper; however, within the broad research literature outlined above our current study of officers walking patrol areas concentrated at crime hotspots can be characterized as both a study of foot patrol as well as hotspots policing. The remainder of this paper reports on what the authors believe is the first large-scale randomized, controlled experiment of the effectiveness of foot patrol to reduce violence in crime hotspots.
Experimental design

BACKGROUND TO THE PHILADELPHIA EXPERIMENT

Philadelphia is the fourth largest police department in America, with over 6,600 police officers. These officers police a city of nearly 1.5 million people, recently ranked the 30th most dangerous in the US (Morgan, Morgan, and Boba, 2010). Violence, recognized as one of the worst public health threats both nationally and locally (CDC, 2010), remains a problem in the city. In 2008 (the year before this study’s intervention) there were 331 homicides in the city, and since the year 2002, Philadelphia has experienced over 100 shootings per month (Ratcliffe and Rengert, 2008).

Although the city had seen a gradual reduction in violent crime levels for a couple of years, there has been a noticeable and consistent seasonal cycle of violent crime increasing during summer months (figure 1). A pilot study of 43 foot beats patrolled during the summer of 2008 indicated a modest reduction in violence in the target areas, with a slight diffusion of benefits to a buffer area of approximately 1,000 feet around target sites (Ratcliffe and Taniguchi, 2008).

Figure 1. Weekly Violent Crime Counts, 2006 to October 2009, Philadelphia, PA.

With the availability of two waves of new recruits emerging from the police academy in March and late June of 2009, there was an opportunity to conduct a larger study. Police Commissioner Charles Ramsey expressed a desire to focus the new recruits emerging from the police academy towards small, targeted foot patrols in high violent crime areas primarily in order to reduce summer violent crime.

SELECTION OF RANDOM ASSIGNMENT HOTSPOTS

We followed a multi-step process to identify the most dangerous places in Philadelphia. During January and February of 2009, violent crime reports were drawn from the incident (INCT) database of the
Philadelphia Police Department for 2006, 2007 and 2008\(^3\). Violent crime was defined as homicide, aggravated assault, and robberies not occurring indoors (the outdoor selection of offenses being in line with the approach of Sherman and Weisburd, 1995). These categories of serious violent crime are not typically affected by issues with crime reporting or police discretion (Gove, Hughes, and Geerken, 1985). Crime events were weighted so events from 2008 counted 1.0, 2007 crimes counted 0.5, and 2006 crime events counted 0.25. In this way, more recent events had greater relevance in the creation of the target locations for 2009, but the areas values could still retain a portion of the long-term hotspot component, given many urban locations have long-term crime trajectories (Weisburd et al., 2004). These weighted values were summed for homicide, aggravated assault and robbery, and then these events were mapped and aggregated to spatial units called Thiessen polygons to create a Voronoi network of spatial units. A Voronoi network consists of areal units created by using lines to divide a plane into areas closest to each of a set of points (in our case street intersections) such that the space within each polygon is closer to the specific point within than to any other point (Chrisman, 2002). For points, we chose the nearly 22,000 intersections in the city. The Voronoi network as a unit of analysis is very similar to the ‘epicenter’ (Sherman and Weisburd, 1995) and ‘intersection area’ approaches (Braga et al., 1999; Weisburd and Green, 1995) used in previous place-based experiments. Those experiments either included the entire blocks associated with an intersection (Braga et al., 1999; Weisburd and Green, 1995) or a more subjective measure of “as far as the eye could see from sidewalk corners” in each direction (Sherman and Weisburd, 1995: 633).

Polygons greater than one million square feet were excluded\(^4\), and a map of weighted violent crime totals for each polygon were presented to the two Philadelphia Regional Operations Commanders (ROC North and ROC South), with the top 220 violent crime corner polygons highlighted. This top one percent of corners (approximately, based on the three year weighted values) contained 15% of the 2008 robberies, 13% of 2008 aggravated assaults, and over 10% of all 2008 homicides. The top 5% of corners accounted in 2008 for 39% of robberies, 42% of aggravated assaults, and 33% of homicides.

Police commanders informed us that they would have sufficient personnel to cover 60 foot patrols for 16 hours a day, 5 days a week, so the Regional Operations Commanders were asked to identify at least 120 potential foot patrol areas of roughly equivalent size, where each patrol area must contain at least one of the top 220 violent crime corners in the city. To aid the creation of patrol areas, we mapped the results of a local Moran’s I spatial autocorrelation test (Anselin, 1995; Moran, 1950). Local Moran’s I is one of a range of local indicators of spatial association (LISA) statistics available to crime scientists that can indicate clustering of high crime values (Anselin, 1996; Getis and Ord, 1992; Getis and Ord, 1996; Unwin, 1996). Mapping polygons with high violence counts that were among high violence neighbors enabled police commanders to see where the hottest corners were surrounded by other high crime areas, and from this information construct more effective foot patrol areas.

Commanders drew 129 potential foot beats they felt were the most important to pursue. The authors examined the patrol areas and adjusted some that were overlapping or deemed too large as originally drawn. During this process some of the original foot beats were split and others were combined which left us with a total of 124 foot beats. The final areas were on-screen digitized and a point-in-polygon GIS

---

\(^3\) We recognize that other studies have relied exclusively on calls for service or some combination of calls for service and reported crime. However, in busy metropolitan police departments, the calls for service files are prohibitively large and not routinely used to inform patrol. District and Regional Commanders in Philadelphia are accountable to the number of crime incidents rather than the frequency of calls for service.

\(^4\) Very large polygons were excluded because the focus of the experiment was on foot patrols in violent crime neighborhoods, and these large areas were deemed less suitable for foot patrol operations that were designed to straddle a number of street intersections. Most of these large polygons bordered parks or industrial areas.
operation was used to re-aggregate the weighted crime points from 2006 to 2008 to the new spatial units. The four lowest crime foot beats were dropped from consideration to leave 120 potential foot patrol areas.

To test the intervention of foot patrols, we employed a randomized block design. In some regards, this approach has some comparable components to a complete block design (Braga and Bond, 2008). Block designs have the advantage of minimizing the effects of variability by allowing for the comparison of similar cases (Mazerolle, Kadleck, and Roehl, 1998). An aggregate total of 2006-2008 temporally weighted violent crime (detailed above) was used to rank all 120 areas from highest to lowest. The foot patrol areas were ranked such that the first couple contained the 1st and 2nd highest ranked areas, the second couple contained the 3rd and 4th highest areas, and so on to the sixtieth couple which contained the 119th and 120th ranked locations. A quasi-random number generator was used to assign one member of each couple as a target area (which would receive foot patrol officers) or a control area (which would receive no foot patrol policing)\(^5\). This randomization process was done without regard to the spatial location or proximity of the treatment and control groups, or on similarity of any other characteristic; randomization was solely a function of the temporally weighted violent crime counts for the three years preceding the experiment. In this way, we were able to use data from 2006 to 2008 to generate a group of target areas for the summer of 2009 that we anticipated would be collectively equivalent in terms of crime intensity as an equivalent group of control areas\(^6\). Police district commanders were not provided with detailed information on the control locations\(^7\). The target and control areas are shown in figure 2.

---

\(^5\) **We recognize that another approach would have been to select at random 60 areas from the original 120; however, this may have prevented us from discovering the threshold level of violent crime for foot patrol effectiveness that is mentioned later in the paper and is a key finding. Furthermore, the research was conducted in an operational environment where the method adopted reassured the city police department that at least a portion of the city’s very high crime areas would be actively patrolled. Specifically, the method ensured randomization would not assign all of the most violent areas to control rather than treatment.**

\(^6\) **Police commanders requested that one area be changed from a control to a target area, and this was changed with the target area with the closest temporally weighted violent crime count from the citywide list.**

\(^7\) **Police commanders were involved in the selection of all 129 potential foot patrol areas some months before the start of the experiment, so theoretically they may have been able to recall the location of the control areas through a process of elimination; however, after the initial site selection, no one at the police department was provided with maps or other details of the control sites.**
Data from 2006-08 were employed to determine the hotspot areas because the police required sufficient lead time to set up the officer allocation and assignment orders. As would be expected with a randomized design no difference between treatment and control groups was found\(^8\); however, when

\(^8\) An independent samples t-test was conducted to evaluate the equivalence in long-term violent crime levels between treatment and control groups. Based on the three years of data used to create the study areas, an independent samples t-test indicated no statistically significant difference between treatment (mean = 32.41, SD = 14.20) and control groups (mean = 31.95, SD = 13.96), t(118) = -0.18, p = 0.86 (two tailed). We should note that if the data are examined using a paired samples t-test, then the difference between treatment and control groups, while substantively small, was nevertheless significant, t(118) = -3.63, p < 0.001 (two tailed). However, as Shadish, Cook and Campbell (2002) point out, random assignment may result in classification where observed mean differences between treatment and control groups exist even when matched designs are used. The randomization process negates the possibility that these differences are indicative of systematic bias. The minor differences found between treatment and control groups in their pre-treatment violent crime levels are statistically controlled for in the regression models. There were no significant differences between
the experiment began in late spring of 2009, for currency we used the immediate three months of crime data prior to the start of the experiment for the pre-treatment measure. Independent samples t-test indicated no significant difference between treatment (mean = 5.98; SD = 4.04) and control groups (mean = 4.93; SD = 3.34) on pre-treatment violent crime counts t(118) = -1.55, p > 0.10 (two tailed)\(^9\).

An \textit{a priori} power analysis was conducted to determine the power of the experimental design (Faul et al., 2007). Given a two-tailed test, an alpha level of 0.10, and 60 cases in the treatment group and 60 cases in the control group power was found to be adequate (> 80\%) when the effect size was large (>) 0.80) or medium (0.50). Power was low when effect size was small (0.10) – a problem common to place-based randomized trials (Boruch et al, 2004); but, power would be above conventionally acceptable levels (greater than 0.80) given an effect size greater than 0.4 when using the parameters listed above.

**TREATMENT**

The target and control areas included an average of 14.7 street intersections (SD 5.3) and 1.3 miles of streets (SD 0.4)\(^10\). Each target area was patrolled by two pairs of officers recently graduated from the police academy. They received a one week orientation at the police district of their specific foot patrol location, and then spent an initial period of a few weeks in and around their beat with an experienced officer. Because none of these orientation activities were required to remain in the foot patrol area, the evaluation date started the week following the final orientation. The officer pairs were assigned either a morning (10am to 6pm) or an evening shift (6pm to 2am) that they policed Tuesdays through Saturday nights. The pairs alternated morning and evening shifts every other week. This meant that the areas were not assigned foot patrols from 2am to 10am each day, and from 2am Sunday right through to 10am Tuesday each week.

Officers were assigned from the academy in two phases. Phase 1 commenced on 31\(^{st}\) March 2009 with officers in 24 foot patrol areas, and continued to September. Phase 2 commenced on 7\(^{th}\) July 2009 and lasted for 12 weeks. There were 36 patrolled areas in Phase 2. This theoretically provided for 57,600 hours of foot patrol activity over the initial 12 weeks of both phases. District Captains were instructed to ensure the foot beats were fully staffed over the experimental period. All patrol officers were provided with an initial criminal intelligence brief on their foot patrol area by the criminal intelligence unit, as well as whatever information about the area they gleaned from their initial orientation. They did not receive specific instructions on policing style from Police Headquarters; however, some officers did report being briefed on the expectations of their respective District Commanders (at the rank of Captain in the Philadelphia Police Department).

\(^9\) Here again we see minor differences between the results of an independent samples and a paired samples t-test. A paired samples t-test indicated that the difference between treatment and control groups was significant, t(118) = 2.03, p < 0.05. See footnote 8 for further discussion of this issue. It is worth noting that if one accepts the difference between treatment and control groups as both significant and substantive, then the direction of these differences works against identifying a treatment effect; the results here should be taken as a conservative estimate of the impact of foot patrol.

\(^10\) An independent samples t-test found no significant differences in the amount of area encompassed by treatment (M = 891,953; SD = 305,506) and control groups (M = 833,038; SD = 332,537) t(118) = -1.01, p > 0.10, the length of road (ft) contained within treatment (M = 6,957; SD = 2,212) and control (M = 6,631; SD = 2,084) groups t(118) = -0.83, p > 0.10, or the number of intersections contained within treatment (M = 14.42; SD = 5.21) and control (M = 14.02; SD = 5.38) groups t(118) = -1.45, p > 0.10.
Field observations by trained researchers found considerable variation in activity. Some officers engaged in extensive community-oriented work, speaking to community members and visiting child care centers and juvenile hangouts, while others were more crime oriented, stopping vehicles at stop signs and intersections, and interviewing pedestrians. Some officers reported receiving a considerable level of supervision and interest from their immediate supervisors, while others reported being largely left to their own devices. Field observers reported that only a few foot patrol boundaries were rigidly observed; a number of officers – either through boredom or a perception that they were displacing crime to nearby streets – would stray for a time if they were aware of areas of interest just beyond the foot patrol area.

OUTCOME MEASURE

The outcome measure for the experiment was reported violent crime. The crime data were drawn from the INCT database of the Philadelphia Police Department, a database containing all police incidents occurring in the city. The database records a Uniform Crime Reports (UCR) classification as used by the national reporting mechanism administered by the FBI, and premises and nature codes that indicate the type of location and the origin of the incident. Violent crime is defined here as criminal homicide, all robberies (except cargo theft), and a majority of aggravated assaults. We excluded violent crime incidents that were deemed unlikely that a patrolling officer could be expected to prevent, such as rape (largely an indoor activity) and some aggravated assaults in specific categories such as against a student by a school employee or against a police officer. School assaults would largely take place on school premises, and assaults against police may increase artificially due to the increased presence of police officers. The INCT database incidents were drawn from roughly three months of each phase (the operational period), and the three months immediately preceding each phase (pre-treatment period). INCT records were drawn at the end of the overall experiment period. The Philadelphia system automatically geocodes crime events with a success (hit) rate in excess of 98 percent, well above an empirically derived minimum acceptable geocoding rate of 85 percent (Ratcliffe, 2004). Descriptive statistics are provided in Table 1.

Table 1. Descriptive Statistics for Counts of Violent Events by Time Period, Experimental and Control Areas.

<table>
<thead>
<tr>
<th>Status (Time Period)</th>
<th>Sum</th>
<th>Mean</th>
<th>Median</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target (Before, (t_0))</td>
<td>359</td>
<td>5.98</td>
<td>5.00</td>
<td>4.04</td>
<td>1</td>
<td>18</td>
<td>0.96</td>
</tr>
<tr>
<td>Target (During, (t_1))</td>
<td>306</td>
<td>5.10</td>
<td>5.00</td>
<td>3.08</td>
<td>0</td>
<td>15</td>
<td>0.77</td>
</tr>
<tr>
<td>Control (Before, (t_0))</td>
<td>296</td>
<td>4.93</td>
<td>4.50</td>
<td>3.34</td>
<td>0</td>
<td>14</td>
<td>0.79</td>
</tr>
<tr>
<td>Control (During, (t_1))</td>
<td>327</td>
<td>5.45</td>
<td>5.00</td>
<td>4.26</td>
<td>0</td>
<td>21</td>
<td>1.63</td>
</tr>
</tbody>
</table>

STATISTICAL APPROACH AND MEASUREMENT OF EFFECTS

The central mechanism to determine the impact of the experiment on the data described above was through comparison of crime frequency before and during the operational period, in both the 60 target areas and the 60 control areas. Descriptive findings are reported in terms of an odds ratio value as a way to indicate the overall difference in the ratio of pre-intervention levels of crime in the target area as compared to the during-intervention crime levels in the control areas, as described by Welsh and Farrington (2009: 134-6). To calculate the reduction in crime, the odds ratio was inverted and calculated as:

\[ OR = 1/(a*d/b*c) \]
Where \( a \) is the event count in the target area pre-intervention, \( d \) is the event count in the control area during the intervention, \( b \) is the event count in the target area during the intervention, and \( c \) is the event count in the control area pre-intervention.

A long running discussion surrounds the most appropriate statistical test to assess change over time in randomized controlled trials (Bohrnstedt, 1969; Frison and Pocock, 1992; Twisk and Vente, 2008). Multiple methods were therefore used to investigate the effect of foot patrol implementation. Primarily, we employed a linear regression model in which the crime value of the operational period serves as the dependent variable and the pre-operation crime level serves as a covariate. Importantly, this approach effectively controls for regression to the mean; a common threat to the internal validity of a study (Twisk and Proper, 2004)\(^{11}\). Subsequently the linear regression model outcomes were examined in phases based on percentile levels of pre-intervention violence in order to examine the impact (if any) of the pre-intervention violent crime frequency.

**Results**

Table 1 shows change in reported crime in the 60 control and 60 target areas for the 3 months before and during the implementation dates for the operational phases. The inverted odds ratio for the crime reduction was 0.77 which, when converted to a percentage change for the target areas relative to control sites indicates a relative reduction of 23 percent.

A simple approach to assessing the significance level of the effect of the intervention is to calculate a change score, the difference between \( t_0 \) and \( t_1 \). These scores are then subjected to an independent-samples t-test to determine if the change between pre-operation and operational time periods was significantly different for treatment and control areas. A significant difference between treatment (mean = -0.88, SD = 4.32) and control (mean = 0.52, SD = 3.44) groups was found (\( t(118) = 1.96, p = 0.05 \)) suggesting that treatment areas had significantly lower change scores (indicating a greater reduction of crime or smaller increases) than their control counterparts\(^{12}\). There are, however, substantial limitations in assessing the effects of treatment through a simple change score analysis. Since change scores only measure the relative change from \( t_0 \) and \( t_1 \) they do not properly account for the starting point of each area.

Numerous methods exist to assess the statistical significance of the change between the pre-operational and operational time periods (Twisk and Vente, 2008). Given the randomization process employed in this study the most direct method of evaluating change would normally be to conduct an independent samples t-test to compare the count of events in the treatment and control groups during the operational period\(^{13}\). Unfortunately this approach, much like the change score analysis presented above,

---

11. The dependent variable in this analysis was the count of violent crime occurring during the three month operational period. Typically count data would be analyzed with a count regression model such as a Poisson regression or a negative binomial regression. The dependent variable was assessed with regards to the functional form of its distribution and was found to be normally distributed to an acceptable degree suggesting that an ordinary least squares approach would be most appropriate. Additional count regression models (results omitted) were conducted but these models indicated no noteworthy differences from the standard ordinary least squares approach adopted here.

12. Using a paired samples t-test produces similar results, \( t(118) = -2.04, p < 0.05 \).

13. Problems with regression to the mean caution against employing an independent samples t-test on the during-treatment violent crime count. Nevertheless, in the interest of presenting complete results an independent samples t-test was conducted to determine the differences in the count of violent crime between treatment and control areas. No statistically significant difference between treatment (mean = 5.10, SD = 3.08) and
fails to adequately control for short-term changes in violent crime. Not considering the differences between treatment and control groups creates a situation where regression towards the mean could threaten the internal validity of the study. In other words, failing to account for the starting point of each area (indicated by the ‘before’ crime count) could lead to over- or under-estimation of the treatment effect (Galton, 1886; Twisk and Vente, 2008) where areas with very high or very low crime at \( t_0 \) will naturally migrate towards more moderate crime levels at \( t_f \).

Therefore to explore the impact of foot patrols on violent crime levels while capturing extraneous influences such as regression to the mean, a limitation of the typical t-test, we employed linear regression models (Frison and Pocock, 1992; Twisk and Proper, 2004; Twisk and Vente, 2008). The dependent variable was the count of violent crime during the three month operational phase, and the independent variable was a dummy variable representing treatment or control status. Pre-treatment scores for the three months prior to the intervention were entered as a covariate, effectively controlling for natural regression to the mean. Table 2 (model 1) presents the results of a linear regression model predicting the violent crime count during the operational phase with the violent crime count during the pre-operational phase and a dummy variable representing treatment status.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Err.</td>
<td>( t )</td>
<td>B</td>
</tr>
<tr>
<td>Constant</td>
<td>3.240***</td>
<td>0.594</td>
<td>5.453</td>
<td>1.585*</td>
</tr>
<tr>
<td>Pre-treatment violent crime count</td>
<td>0.448***</td>
<td>0.083</td>
<td>5.404</td>
<td>0.783***</td>
</tr>
<tr>
<td>Treatment status</td>
<td>-0.820</td>
<td>0.616</td>
<td>-1.332</td>
<td>2.209*</td>
</tr>
<tr>
<td>Pre-treatment violent crime count * treatment status</td>
<td></td>
<td></td>
<td>-0.565***</td>
<td>0.161</td>
</tr>
</tbody>
</table>

\( N = 120 \)
* \( p < .05 \); ** \( p < .01 \); *** \( p < .001 \)

There was a strong relationship between pre-treatment violent crime count and the violent crime count during the operational period. Treatment status was found to be non-significant. When differences between the starting violent crime levels in the foot patrol areas were properly accounted for, treatment and control areas showed no significant differences in the violent crime level during the operational period; however, one assumption underlying regression models is that the relationship between the covariate (here the pre-treatment violent crime count) and the dependent variable (the crime level during the operational period) is the same for each group. Put simply, the treatment and control areas are assumed to have a similar relationship between the pre-treatment violent crime count and the violent crime count during the operational period. Exploratory analysis of the regression slopes fitted for each group suggested treatment and control areas had substantially different slope values. This suggested that an interaction term between treatment and the pre-treatment violent crime count would be informative on both theoretical and statistical grounds. Table 2 (model 2) presents the results of a linear regression model including a pre-treatment count and treatment status interaction term.

\( control \) groups (mean = 5.45, \( SD = 4.26 \) was found, \( t(118) = 0.52, p > 0.10 \) (two tailed). A paired samples t-test produced similar results, \( t(118) = -0.78, p > 0.10 \) (two tailed). Power analysis was re-specified using the conditions set out in this model and was determined to be 0.68 suggesting that the possibility of finding a significant effect of treatment in these models was substantially improved but still below the traditional 0.80 threshold.
The significance of the interaction term suggests that it would be inappropriate to refer to the effectiveness of the treatment in reducing violent crime without also specifying the level of pre-treatment violence. That is, the slope of the pre-treatment violent crime level varied by treatment and control groups. Visual inspection of a scatter plot between violence pre-operation and violence during the operation suggested that treatment may have little effect for areas starting and ending with low violent crime counts but may have a larger effect for areas with higher pre- and during-operation counts. To further explore these trends, adjusted mean crime counts for the operational period were calculated for pre-operational violent crime scores corresponding to the 20th, 40th, 60th, 80th, and 90th percentiles. The differences in expected violent crime count for target and control areas were then assessed to determine under what pre-treatment crime levels treatment had a significant impact. These results can be found in table 3.

Table 3. Treatment-Control Differences in Counts of Violent Offenses, by Pre-Treatment Violent Crime Count.

<table>
<thead>
<tr>
<th>Pre-treatment violent crime count (percentile)</th>
<th>Target area estimate during the operational period</th>
<th>Control area estimate during the operational period</th>
<th>Difference (target area – control area)</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (20th)</td>
<td>4.449</td>
<td>3.936</td>
<td>0.514</td>
<td>0.538</td>
<td>0.465</td>
</tr>
<tr>
<td>4.5 (40th)</td>
<td>4.667</td>
<td>4.719</td>
<td>-0.052</td>
<td>0.007</td>
<td>0.935</td>
</tr>
<tr>
<td>6 (60th)</td>
<td>5.104</td>
<td>6.286</td>
<td>-1.182</td>
<td>3.919</td>
<td>0.050</td>
</tr>
<tr>
<td>8 (80th)</td>
<td>5.540</td>
<td>7.852</td>
<td>-2.312</td>
<td>10.148</td>
<td>0.002</td>
</tr>
<tr>
<td>11 (90th)</td>
<td>6.195</td>
<td>10.202</td>
<td>-4.008</td>
<td>13.705</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

These results suggest that there were no differences between treatment and control groups in the 20th and 40th percentiles. At the 60th percentile and higher, target areas had less violent crime than their control counterparts, a finding significant at p < 0.05. This difference became more noteworthy in higher percentiles. It is worth reiterating at this point that because pre-treatment violent crime counts were entered into the regression model the differences seen here represent impacts above and beyond what would be expected based on regression to the mean. In other words, even after accounting for natural regression to the mean, target areas in the top 40% on pre-treatment violent crime counts had significantly less violent crime once the operational period was under way than their control counterparts. This finding has potential implications for deploying scarce resources and is discussed in the following section.

Variables broadly indicative of police activity can illuminate these results. For example, a pedestrian stop is recorded whenever a police officer conducts a field interview, stop-and-frisk, or search of a suspect in the street. Similarly a vehicle stop is recorded when this is conducted with occupants of a vehicle. There are also types of police activity that are largely the result of proactive policing rather than a response to calls from the public. For instance, disturbances can include incidents such as disorderly crowds or small gatherings that can be identified and dispersed by police officers as well as rowdy behavior in and around liquor establishments; narcotics incidents are largely the result of proactive police work; and disorder incidents such as prostitution, public drunkenness, loitering and violation of city ordinances are often largely left to police to initiate, especially in higher crime areas. Arrests are likely a combination of reactive policing (responding to a call from the public) and proactive activity.
Table 4. Treatment-Control Differences in Counts for Various Incident Types by Time Period.

<table>
<thead>
<tr>
<th>Incident type</th>
<th>Control areas</th>
<th>Treatment areas</th>
<th>% Contribution to increase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-operation</td>
<td>During operation</td>
<td>Pre-operation</td>
</tr>
<tr>
<td>Pedestrian stops</td>
<td>5965</td>
<td>5985 (+1%)</td>
<td>7366</td>
</tr>
<tr>
<td>Vehicle stops</td>
<td>5600</td>
<td>4862 (-13%)</td>
<td>5922</td>
</tr>
<tr>
<td>Disturbances</td>
<td>3600</td>
<td>4033 (+12%)</td>
<td>3980</td>
</tr>
<tr>
<td>Narcotics</td>
<td>397</td>
<td>370 (-7%)</td>
<td>464</td>
</tr>
<tr>
<td>Disorder</td>
<td>249</td>
<td>288 (+16%)</td>
<td>336</td>
</tr>
<tr>
<td>Arrests</td>
<td>1395</td>
<td>1361 (-2%)</td>
<td>1684</td>
</tr>
</tbody>
</table>

† Categories noted in the table are not mutually exclusive, with the exception of pedestrian and vehicle stops. For example, when a Philadelphia police department officer conducts a pedestrian stop, it is recorded as a separate incident regardless of the outcome. This is done for managerial purposes. If the stop results in a narcotics arrest, a separate narcotics incident will be created, with a field that shows an arrest was made.

# In some cases the number of incidents dealt with by non-foot patrol officers decreased from t0 to t1, and where indicated the foot patrol officers conducted sufficient activity to offset the reduction and contribute to an overall increase.

Table 4 shows that the frequency of all of these incident types increased during the police operation and that the foot patrol officers (as identified by their radio call signs) contributed substantially to the rise observed in treatment areas. Using the pedestrian stops example, while stops increased less than one percent in control areas they increased by 64 percent in treatment areas. Foot patrol officers conducted 4,282 pedestrian stops, amounting to 35 percent of all pedestrian stops in the treatment areas during the operation, and contributing 90 percent of the increase in the treatment sites. The additional vehicle stops and narcotics incidents handled by foot patrol officers in the treatment areas offset and added to a decrease in these activities by other (vehicle-bound) officers. This is the same situation for the total number of arrests in the treatment areas. Non-foot patrol officer arrests declined slightly in both treatment and control areas; however the additional nearly 400 arrests by foot patrol officers increased the overall arrest count by 13 percent.

When these additional activities are disaggregated further, it can be seen that foot patrol officers in the top 20 percent of highest crime areas were engaged in significantly more work than foot patrol officers in the lower volume crime hotspots. Table 5 shows little substantial difference among average activity levels for crime hotspots at lower percentiles; however, activity across all measures increases considerably for the top 12 foot patrol areas, with for example, officers conducting on average 115 pedestrian stops during the three month operational period compared to only 57 in the lowest pre-

---

Foot patrol officers conducted some official actions in the control areas. Specifically, in the control areas they conducted 108 vehicle stops (representing 2.2% of all vehicle stops in the control areas), 372 pedestrian stops (6%), attended 6 disorder incidents (2%), 5 drug incidents (1.3%), 92 disturbances incidents (2.2%) and made 23 arrests (1.7%). Based on field observations, we believe these activities were due mainly to foot patrol officers’ activity as they were walking to their assigned beats and when they occasionally strayed from their defined patrol boundaries. The officers were not informed of the control area locations, and the proximity of some experimental and control areas likely contributed to this activity. Overall the foot patrol officers contributed approximately 3.6% of the measured activity in the control areas.
treatment crime areas. They also conducted more vehicle stops, dealt with more disturbances and narcotics incidents, and made substantially more arrests.

Table 5. Mean Counts of Incidents Handled by Foot Patrol Officers During Experiment, by Pre-Treatment Violent Crime Count.

<table>
<thead>
<tr>
<th>Pre-treatment percentile</th>
<th>Pedestrian stops</th>
<th>Vehicle stops</th>
<th>Disturbances</th>
<th>Narcotics</th>
<th>Disorder</th>
<th>Arrests</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20</td>
<td>57.3</td>
<td>11.3</td>
<td>18.9</td>
<td>1.1</td>
<td>1.3</td>
<td>3.2</td>
</tr>
<tr>
<td>20-40</td>
<td>61.8</td>
<td>13.7</td>
<td>23.6</td>
<td>1.5</td>
<td>2.7</td>
<td>6.3</td>
</tr>
<tr>
<td>40-60</td>
<td>44.9</td>
<td>7.5</td>
<td>22.3</td>
<td>0.8</td>
<td>3.0</td>
<td>4.7</td>
</tr>
<tr>
<td>60-80</td>
<td>77.4</td>
<td>13.8</td>
<td>22.2</td>
<td>1.9</td>
<td>2.1</td>
<td>6.9</td>
</tr>
<tr>
<td>80-100</td>
<td>115.4</td>
<td>20.3</td>
<td>36.4</td>
<td>4.6</td>
<td>4.1</td>
<td>12.2</td>
</tr>
</tbody>
</table>

**Displacement or Diffusion of Benefits?**

The issue of crime displacement is not only an unrelenting concern of police and the public, but also a frequent topic of academic interest (Eck, 1993; Green, 1995; Hesseling, 1994; Ratcliffe, Taniguchi, and Taylor, 2009; Weisburd et al., 2006). In the minds of many police officers, displacement is an inevitable outcome of spatially targeted crime prevention activity (Barr and Pease, 1990), yet the research evidence suggests that a diffusion of benefits is also a potential, and likely, outcome (Clarke and Weisburd, 1994; Hesseling, 1994). From a theoretical standpoint, environmental criminology suggests displacing crime to another location is likely to result in a reduction in offending as criminals are pushed to commit crime in a less optimal site (assuming they were originally offending in their optimal location).

To examine this issue, we employ the weighted displacement quotient methodology of Bowers and Johnson (2003). This approach compares changes in the ratio of crime in target areas to changes in control areas, to calculate a success measure. If the success measure indicates that crime was reduced in the target area to a greater extent than across the control areas, then the researcher continues to calculate a buffer displacement measure. The ratio of these two measures creates a range of outcomes for the buffer area; more crime than was reduced in the target area can be displaced, some of the crime can be displaced, there can be a diffusion of some benefits to the buffer area, or the buffer area can sometimes even outperform the crime reduction in the target area. These outcomes are estimated from the equation;

\[
WDQ = \frac{B_{t1}/C_{t1} - B_{t0}/C_{t0}}{A_{t1}/C_{t1} - A_{t0}/C_{t0}} = \frac{\text{(Buffer displacement measure)}}{\text{(Success measure)}}
\]  

(1)

where \(A\) represents crime in the target areas before the operation started \((t_0)\) and during the operational period \((t_1)\), \(B\) represents crime in the buffer areas, and \(C\) is crime in the control zones.

Determination of the appropriate buffer area around target sites was made by field researchers, within guidelines, based on their knowledge of the experiment sites and where offenders were likely to be displaced. The guidelines for buffer areas were;

- Buffers can generally extend no more than two of the longer side of a Philadelphia (rectangular) street block,
- Buffers cannot overlap with other target, buffer or comparison areas, and
• Buffers should not extend across clear urban barriers (such as railway lines) unless there are easy access routes (such as bridges or pedestrian tunnels).

Ten target areas were so close to another target area that the buffer areas for these areas were combined into five buffer areas, each of which contained two target areas, resulting in a total of 60 target areas and 55 buffer areas. The displacement or diffusion of benefits effect was calculated for the outcome measure of violent crime, and a total net effect measure (TNE) reports the overall program outcome after inclusion of any buffer area effects (Clarke and Eck, 2005; Guerette, 2009). The total net effect of the operation can be calculated in relatively simple terms by examining the ratio of the crime reduction in the target areas after factoring in the general change in the control areas and then taking into consideration any displacement or diffusion to the buffer area. Adapting the equation from Guerette (2009: 41)\(^{16}\) to the terminology employed at (1) above, the TNE is calculated:

\[
TNE = [A_{10}(C_{11} / C_{10}) - A_{11}] + [B_{10}(C_{11} / C_{10}) - B_{11}]
\]

The equation is comprised of two parts, a target net effect \([A_{10}(C_{11} / C_{10}) - A_{11}]\) and a buffer net effect \([B_{10}(C_{11} / C_{10}) - B_{11}]\). The citywide success measure of -0.277 reiterates the reduction in violent crime and confirms the value in proceeding with a buffer displacement measure. Combining the buffer displacement measure (0.11) with the success measure results in a weighted displacement quotient of -0.41. Referring this value to table 1 of Bowers and Johnson (2003: 286) suggests there was displacement of violent crime during the experiment, but that the displacement was less than the direct benefits achieved in the target areas.

The total net effect of the operation can be calculated by examining the ratio of the crime reduction in the target area after taking into consideration any displacement or diffusion to the buffer area and factoring in the general change in the control areas, as shown in equation (2). It can be seen that the algorithm is the combined ratio of the net change in the target areas (relative to changes in the control areas) added to any net change in the buffer areas. Replacing the equation with values shows a net total effect of 53.11, effectively a total program effect of 53 prevented violent crimes.

\[
TNE = [359(327/296) - 306] + [320(327/296) - 392] = 90.60 + (-37.49) = 53.11
\]

This crime reduction of 53 violent crimes is comprised of a reduction of 90 crimes in the target area, offset by a 37 offense increase occurring in the displacement areas immediately surrounding target areas\(^{17}\).

---

\(^{16}\) Rob Guerette acknowledges the origins of this in the work of Bowers and Johnson (2003), and extended by Eck and Johnson (see Clarke and Eck, 2005; Guerette and Bowers, 2009).

\(^{17}\) This approach, based on the work of Bowers and Johnson (2003) and Guerette (2009), does not measure displacement around control areas. We recognize that a couple of place-based experiments have measured displacement around control sites in order to account for any movement of crime that might be unrelated to the intervention. By taking the approach of Bowers and Johnson (2003) we ignore any potential displacement out of control areas on the grounds that areas with no treatment would not expect to have any displacement. As a result of adopting the Bowers and Johnson approach, we recognize that our displacement findings may overestimate the degree of displacement and are inherently conservative with regard to the benefits of the intervention.
Limitations

We are cautious in saying that the crime reduction outcome in the foot patrol areas was entirely the result of foot patrol officers. Like many places, the Philadelphia Police Department does not employ crime analysts, and centrally-generated spatial crime intelligence disseminated to district patrol officers and supervisors is fairly sparse (see Ratcliffe, 2008). While knowledge of foot patrol locations was not formally disseminated beyond the necessary districts, neither were the sites exactly a secret, and in the absence of little other guiding information it is possible that officers not involved in the experiment were called in to periodically assist foot officers, or used the known foot beat areas as indicative of crime hotspots to which they should also pay attention. Table 4 would suggest, however, that this was not a significant issue.

We should also caution that in terms of violent crime count, the numbers examined in this paper are quite small. Although the aggregate crime counts are over 300 for the target areas, the effect becomes diluted when distributed across all target areas. At an individual foot patrol area level, the effect represents a net gain of less than two violent crimes per foot patrol area, and this drops to less than one when the total net effect of changes in the displacement area is factored. This is to be expected given that violent crime levels are often lower than the public imagine, and especially given the constrained spatial units employed by the experiment. This is at least partially responsible for the low observed power found in this experiment. For this reason we are reluctant to report results for individual police districts or foot patrol areas where one or two violent offenses either way could have an impact on an area’s individual effectiveness.

We were unable to support sufficient field research time to generate robust measures of patrol time within each foot patrol area due to financial limitations. Graduate students observed foot patrol officers in each the 60 treatment areas for approximately 2 hours for each day shift and 2 hours for each night shift, totaling 240 hours observation time. This is an insufficient observational period from which to extrapolate and develop an estimate of the total time spent by officers in their beats. During observations, officers likely focused on showing researchers around their assigned beats. Even if officers went beyond beat boundaries in the presence of observers there was no way of measuring how long they stayed and worked in such areas during the course of an (unobserved) shift of 8 hours. Foot patrol officers did conduct a few official activities in control areas (see footnote 15) but this accounted for less than four percent of all incidents within the control areas. It is anticipated that analysis of both field notes as well as post-experiment interviews with foot patrol officers may in the future enable a more nuanced understanding of officer staffing and officer compliance with patrol boundaries.

This issue of potential crime displacement or diffusion of benefits is therefore addressed in our research. As stated earlier, there are two methodological schools of thought, and we adopted the Bowers and Johnson approach. As explained in footnote 17, we recognize that our findings may overestimate the degree of displacement and thus are conservative with regard to the benefits of the intervention.

We also report descriptive output statistics on the differences between treatment and control areas with regard to a number of official indicators of proactive police activity. The data reported are official data only, the limitations of which are well known. As Durlauf and Nagin (2011) point out, measures of apprehension risk based on official records of crime or enforcement are incomplete because they are unable to incorporate the risk of apprehension for opportunities overlooked by offenders because the risk was too high. While these data suggest a component of the violent crime reduction may have its origins in proactive policing, disentangling specific deterrence effects of mere presence versus officers’ proactive activity was beyond the reach of this paper. It is acknowledged that we are not able to parse the observed crime reduction into an officer presence component and an activity component, thereby
limiting our study to a partial test of deterrence. Articulating the dimensions of this distinction would be an excellent avenue for future research.

Discussion

We found that violent crime hotspots that were recipients of foot patrol officers for up to 90 hours per week had a reduction in violence of 90 offenses (with a net effect of 53 offenses once displacement is considered), outperforming equivalent control areas by 23 percent; however, the benefits were only achieved in areas with a threshold level of pre-intervention violence. When that threshold was achieved (in our study an average of 6 violent crimes in the three months pre-intervention), these target areas in the top 40% on pre-treatment violent crime counts had significantly lower levels of violent crime during the operational period, even after accounting for natural regression to the mean.

Our findings therefore raise the possibility that the Newark Foot Patrol Experiment and subsequent follow-on studies are not necessarily the last word on foot patrol effectiveness. In theoretical terms, our study suggests that the foot patrols operated as a ‘certainty-communicating device’\textsuperscript{18} within the micro-spatial contexts of the hotspot areas. As our analysis focused on outdoor crimes, the data suggests that there were more behaviors which the police had the capacity to influence in the target areas with high thresholds of pre-treatment violence. As Stinchcombe pointed out long ago (1963), police activity is structured by the location of crimes in terms of whether they occur in public space or within the ‘institutions of privacy’. In dense urban settings with high levels of outdoor criminal behavior, more police-initiated activity in the form of enforcement and order maintenance is likely to occur. From this perspective, spatially focused foot patrol may communicate an increased level of certainty that crimes will be detected, disrupted, and/or punished. This perceived risk of detection might be especially high for individuals ‘on the run’, such as those with arrest warrants who may seek to minimize the chances of police encounters in public spaces (see Goffman, 2009). Overall, this theoretical explanation is consistent with the conclusion of deterrence researchers that certainty of apprehension plays a stronger role than severity of punishment as a mechanism of general deterrence (Nagin and Pogarsky, 2001; Nagin, 2010; Durlauf and Nagin, 2011).

The overall crime reduction in foot patrol areas is not trivial and the reduction represents a net outcome of 53 fewer crime victims in a city wrestling, like many American cities, with the individual and public health impact of violence. If, as we suggest, that deterrence is highly localized, one possible explanation for the difference in crime outcome from Newark to Philadelphia may be an issue of spatial dosage. The Newark experiment began with existing foot beats, some of which were commercial corridors up to sixteen blocks in length. The chances that patrolling officers would soon return to an intersection once perambulated would be slim. Benefiting from the application of GIS, in collaboration with senior commanders at the Philadelphia Police Department, we designed foot patrol areas that averaged just 1.3 miles in total street length. It is likely that if foot patrols are only effective due to a certain spatial concentration, then larger foot patrol areas become ineffective. When the local police department in Flint (MI) expanded their foot patrol areas against the wishes of the research team (in one case up to 20 times the original area) crime reduction effectiveness decreased substantially (Trojanowicz, 1986). If dosage, either in terms of spatial foot beat extent or the number of officers assigned to a given area, is fundamental to the effectiveness of foot patrols as a violent crime reduction tool, our research represents an important first step rather than the final word. We say a first step because we acknowledge that, due to the speed with which the operation was conceived and implemented, there

\textsuperscript{18} We are grateful to one of our anonymous reviewers for this terminology.
was no time to find the funds necessary to enable a robust measure of dosage. This limitation should be considered by researchers looking to replicate this study.

A second potential distinction between Newark and Philadelphia relates to an operational difference. In Newark, a number of foot beats had existed for at least five years, and part of the experimental design used random selection to either retain or drop these beats. If some of the Newark beats had been patrolled for a number of years, there is the possibility that patrolling officers had become jaded or tired of the assignment, resulting in crackdown decay (Sherman, 1990). Equally possible, offenders had learned the rhythm of the foot patrols and adjusted to the conditions, finding new opportunities to commit crime in the target area in different ways. With regard to the Newark experiment, Pate (1986) also raises the issue of internal validity; because of the mechanism of the way pre-experiment foot beats had been selected there was a potential for the selective assignment of new beats to non-equivalent groups. These distinctions between the foot patrol experiments in Newark and Philadelphia reinforce the assertion of Durlauf and Nagin (2011) that ‘police-related deterrent effects are heterogeneous; they depend on how the police are used and the circumstances in which they are used.’

The change score analysis provides an overall assessment of the outcome, but the linear regression incorporating the interaction term of treatment status with violence pre-intervention may provide the most significant finding from both an operational and theoretical perspective. The lack of statistical significance for hotspots with a lower level of pre-intervention violence suggests that foot patrols are not a silver bullet to the problem of violence. Only when a pre-intervention violence count of six crimes (the 60\(^{th}\) percentile in table 3) was achieved did the intervention become successful. The broader implication is that foot patrols may only be able to deter violent crime once a threshold of violence exists. In the future, police organizations may benefit from a more situational approach that is tailored to neighborhood characteristics and crime levels (Nolan, Conti, and McDevitt, 2004). For instance, police departments may want to target their foot patrol resources in only the highest crime places in order to improve overall security and maximize the chance of success, while other solutions, such as the targeted application of a problem-oriented policing approach may be more suitable to neighborhoods with a lower threshold of violence and greater community capacity.

A situational approach may help address concerns of some researchers of hotspots policing that their findings be interpreted as providing a carte blanche for a more aggressive policing stance (for example Sampson and Cohen, 1988; Sherman, 1986). We definitely concur. The data shown in table 5 indicate a substantial jump in proactive activity for foot patrol officers in the highest quintile crime areas of the experiment; however, we are reluctant to suggest that proactive policing alone resulted in the crime reduction found in this experiment. Being unable to gauge the level of informal community contacts during the foot patrols we cannot state categorically that these formal activities alone were able to communicate the increased certainty of police intervention, which is essential to deterrence. Mere presence, or (unmeasured) community interactions may have contributed equally to the crime reductions seen in the foot patrol areas. Proactive police work resulting in more traffic tickets, more pedestrian field interviews and more arrests can run the risk of alienating the local community. Furthermore, increased police activity could potentially increase other public health risks. For example, increased enforcement of drug-related behavior may deter drug users from seeking services at a syringe exchange program (Davis et al., 2005). Thus, while our results suggest that foot patrols were effective in the higher crime hotspots, this may be too high a price for community-police relations in some areas and certainly more work needs to be conducted to examine potentially harmful outcomes of focused police efforts (Weisburd and Braga, 2006; Durlauf and Nagin, 2011; National Research Council, 2004).

At least in Philadelphia, both anecdotal feedback from police commanders and documented field observations indicated that there was no noticeable public backlash in response to additional police.
activity in the target areas. Rather, community figures in many areas complained when the summer foot patrol experiment finished and officers were reassigned. This should not read as a mandate to promote complacency in community relations. Our study was a largely pro bono venture to assist our local police department and the limited funding we garnered in a short time was not sufficient to provide the resources to fully assess the community impact of the intervention. It is to be hoped that any replication in other jurisdictions will be able to examine the impact of foot patrols for a longer time period, as well as examine broader impacts on community relations and public health.

Additional potential negative consequences relate to the increase in arrests and other enforcement. Many cities are facing over-crowded jails and prisons, and criminal justice systems straining under the weight of too few resources to address too many needs. Given that target area arrests increased 13 percent relative to the control areas, significant consequences on the criminal justice system in terms of increased criminal processing time or increases in the number of fugitives may ensue (Goldkamp and Vilicică, 2008). Furthermore, we did discover some displacement of violent crime (see footnote 17), and this is obviously of concern to residents of areas surrounding police intervention sites. That the operation in question was an overall success and knowing that any displaced crime was of a lower volume than the crime prevented through the Philadelphia foot patrol experiment would be little comfort to a crime victim in a surrounding area. This leaves police commanders with somewhat of a conundrum. They could plan enforcement operations for neighborhood-wide areas that demonstrate action to a wider community but potentially be unsuccessful at measurably reducing crime, or focus scarce resources in a small area and show effectiveness but have to accept the possibility of some collateral damage to nearby areas. It is to be hoped that if this experiment is repeated, either in Philadelphia or elsewhere, that the displacement observed here was anomalous and future outcomes demonstrate the more common diffusion of benefits seen in many other studies (Hesseling, 1994; Ratcliffe and Makkai, 2004).

Finally, patrolling officers did little to address the underlying causes or social determinants of violence (WHO, 2002). Environmental criminology theory stresses the importance of the situational and contextual moment of a crime event, and any deterrent capabilities of the police were likely place-based, but transitory. In a recent randomized experiment, the tactic of saturation patrol in police cars was found to under-perform problem-oriented policing interventions (Taylor, Koper and Woods, 2010). It may be that not only are vehicle-bound patrol officers unable to impact crime levels significantly, but also that foot patrol officers develop greater situational knowledge. A useful future direction with any foot patrol studies would be to develop in officers an appreciation for the merits of a problem-solving/problem-oriented policing approach that could leverage their local knowledge developed over months of foot patrol into a long-term problem reduction strategy.

**Conclusion**

This research has been a response to an identified need to discover which specific hotspots strategies work best in particular types of situations (National Research Council, 2004; Weisburd and Braga, 2006). Foot patrols have until now been written off as unsuccessful in combating crime, and especially violent crime, a view largely emanating from the Newark foot patrol experiment of over 25 years ago. We estimated that police foot patrols prevented 90 crimes in violent crime hotspots, though there was apparent displacement of 37 of these crimes to nearby areas; thus, the net crime prevention effect from the foot patrol experiment was 53 crimes prevented. This crime reduction was most likely achieved through a combination of community contacts and interaction, alongside more proactive enforcement
and field investigations. This additional level of police activity may seem overly aggressive in the eyes of some members of the community, but with others there may be considerable relief that the police are having a more active presence in their neighborhoods. Community surveys or some other form of societal litmus test could help police find a state of equilibrium with effective and proactive enforcement on the one hand and community approval, or at least reluctant tolerance, on the other. If these findings can be replicated, and a suitable balance can be struck between police intervention and community perception, it may be that police are able to reap the crime reduction and public health outcomes of the focused foot patrol intervention examined here, while retaining the community support reported many years ago in Newark.

References


If we examine the target area in isolation, ignoring any displacement or changes in the control areas, each violent crime reduction was associated with an additional 89 pedestrian stops, 8 vehicle stops, 4 arrests, 35 disturbance and 4 disorder incidents, and an additional narcotics incident; however, no causal connection is directly claimed, and this list only includes official activities and does not include any potential crime prevention from informal, more community oriented work.


Acknowledgements

The authors would like to thank Philadelphia Police Commissioner Charles Ramsey and the executive team at the Philadelphia Police Department for their collaborative approach to research and their support of this project, including Deputy Commissioner Richard Ross, Deputy Commissioner Tommy Wright, Deputy Commissioner Kevin Bethel, Chief Administrative Officer Nola Joyce, and Director of Strategic Communications Karima Zedan. We would also like to show our appreciation to the District Commanders and patrol officers for their hospitality. The authors would like to thank Evan Sorg, Lallen Johnson, and Cory Haberman for their assistance with fieldwork, and John Goldkamp, Ralph B. Taylor and the anonymous reviewers for their insightful comments on various drafts of this paper. Aspects of this project were supported by the Public Health Law Research Program, Robert Wood Johnson Foundation and the Temple University College of Liberal Arts Research Award (CLARA) Program. Points of view or opinions expressed herein are those of the authors and do not necessarily represent the official position of the Philadelphia Police Department or the City of Philadelphia.


For more details, press releases and a short three-page research brief, visit the project website at www.temple.edu/cj/FootPatrolProject