JACQUELINE COHEN
JENS LUDWIG

6 *Policing Crime Guns*

B etween 1985 and 1991 the homicide rate in the United States increased by nearly 25 percent, from 7.9 to 9.8 per 100,000 residents. Almost all of this increase was accounted for by additional gun homicides committed against and by young males.[1] Although the homicide rate has declined substantially during the 1990s, homicide in the United States is still dominated by young people and firearms and remains much more frequent than in other developed nations.[2] Policymakers who are concerned about America's problem with lethal violence must ask: how can we prevent young men from shooting one another?

One increasingly popular answer is to increase the risks of carrying guns illegally through stepped-up police enforcement. Under these "directed patrol"

1. Cook and Laub (1998); Blumstein (2000).
2. Blumstein and Wallman (2000).

programs, high-crime areas are targeted for additional police resources that focus on illegally carried firearms. The hope is that targeted patrols will deter high-risk people from carrying or misusing guns in public places, consistent with evidence that criminals seem to be deterred by the threat of punishment in other contexts.[3] Such patrols may also reduce illegal gun carrying through an "incapacitation effect" by taking illegal guns or those who carry them off the street. The aim of successful targeted policing programs is to reduce illegal gun carrying in public places, a proximate cause of many lethal and nonlethal gun assaults, while avoiding many of the practical and political difficulties of regulating private gun ownership.[4]

Whether targeted policing against illegal guns reduces gun violence in practice remains unclear. To date the evidence in support of such efforts comes largely from the widely cited Kansas City Gun Experiment, which assigned additional police resources to more vigorously pursue illegal guns in one high-crime neighborhood of the city but not in another. While the "treatment" neighborhood experienced a 65 percent increase in the number of guns seized by the police and a 49 percent reduction in gun crimes, neither outcome measure showed much change over this period in the "control" area.[5]

Although the findings from Kansas City are suggestive, the program is not an "experiment" as scientists use the term since there is no guarantee that the two neighborhoods *would have* had similar crime rates had the policing intervention *not* been launched. Some support for this concern comes from the fact that gun crimes were more common in the "control" neighborhood for extended periods even before the new policing program was initiated.[6]

Indianapolis implemented a similar targeted policing program in 1997. One area of the city was targeted for stepped-up vehicle stops for minor violations, while in another area police focused on stopping the most suspicious people within these communities. The results are somewhat puzzling: the number of gun seizures increased by around half with vehicle stops but changed very little with person stops, yet the latter area experienced a decline in gun crimes both in absolute terms and in comparison to other parts of the city.[7] Because of this discrepancy in impact, there is growing interest in the distinction between interventions targeted at "people" rather than simply "places," although it is also possible that the results in the targeted area are spurious. Despite the popular

---

3. Nagin (1998); Levitt (2001).

4. James Q. Wilson. "Just Take Away Their Guns," *New York Times Magazine,* March 20, 1994, sec 6, p. 47; Sherman (2000).

5. Sherman and Rogan (1995); Sherman, Shaw and Rogan (1995).

6. Sherman, Shaw and Rogan (1995).

7. McGarrell and others (2001).

support for these police patrols and their increased use by urban police depart-
ments, reliable evidence on their effectiveness remains limited at best.[8]

In this chapter we present new evidence on the effects of police programs
against illegal gun carrying that draws on data from Pittsburgh, Pennsylvania.
As with ail nonexperimental policy evaluations, identifying causal program effects
is difficult. However, several features of Pittsburgh's 1998 policing program
offer a unique opportunity to isolate the impact of the police patrols from the
effects of other confounding factors that cause crime rates to vary across areas
and over time.

The Pittsburgh police department stepped up police patrols in some areas
but not othets of the city, which enables us to compare trends in crime rates be-
tween the treatment and control areas before and after the police patrols were
launched. Yet this type of across-area over-time comparisons is not without lim-
itations: a standard concern is that crime rates in the treatment communities
may simply follow a different trajectory over time from those of the control areas,
and so differences in trends once the patrols are implemented in the target neigh-
borhoods may not reflect the effects of the new policing intervention.

The unique aspect of Pittsburgh's program is that the police patrols were
launched on some days of the week (Wednesday through Saturday, hereafter
"on days") but not others (Sunday to Tuesday, "off days") within the treatment
(or target) communities. We compare trends between the treatment and control
neighborhoods in the periods before and after the policing program is launched,
focusing on gun misuse during the on days. If the policing program has an effect,
we would expect a greater decline during the on days of the week in the treat-
ment than control communities, and this decline should be larger than the dif-
ference in trends across neighborhoods observed during the off days. To the ex-
tent that unmeasured confounding variables cause the treatment neighborhoods
to have different trends from the control areas during every day of the week, this
approach controls for these omitted factors by comparing across-area trends dur-
ing on versus off days. This strategy thus isolates the causal effects of those fac-
tors unique to the target neighborhoods following the launch of the police patrols
during the days when these patrols are active—factors such as the police patrols
themselves.[9]

---

8. Survey data indicate that these targeted policing programs enjoyed widespread support from
both black and white residents in Indianapolis and Kansas City. Shaw (1995); Chermak, McGarrell,
and Weiss (2001); McGarrell, Chernak, and Weiss (1999).

Previous studies have also generated suggestive findings that community policing and focused
problem-solving interventions may reduce gun crime. Dunworth (2000); Kennedy and others (2001),
although these studies are susceptible to the same confourding problems as those with the Kansas City
evaluation.

9. Gruber (1994); Joyce and Kaestner (1996); Ludwig (1998).

In our judgment the Pittsburgh program provides at least suggestive evidence that targeted patrols against illegally carried guns may reduce gun crime. Our analysis suggests the policing program may have reduced shots fired by perhaps as much as 34 percent, and hospital-treated assault gunshot injuries declined by 71 percent during on days in program-treated areas. These reductions are likely to occur because of deterred gun carrying or criminal behavior rather than inca-pacitation since the number of actual arrests and guns confiscated as a result of the patrols was fairly modest. There is also no evidence of spillover or displace-ment that affect gun crime levels in untreated areas or during off days.

Given the high costs that gun violence imposes on society—about $1 mil-lion per gunshot injury[10]—the fairly modest program cost of under $35,000 in overtime expenditures is small in comparison to the potential benefits to Pittsburgh residents, which may be as large as $25 million. Perhaps more im-portant, the policing program in Pittsburgh was implemented in a way that was sensitive to concerns about individual liberty and police-community relations. No citizen complaints were filed against the police department as a result of the new program.

Policing in Pittsburgh

Pittsburgh shares many characteristics with other American cities that make the policing program evaluated in this chapter of national interest. While public at-tention has focused on the dramatic increase in gun violence starting in the mid-1980s in large cities such as New York, Los Angeles, and Chicago, similar in-creases were observed in the early 1990s in more modesdy sized cities such as Pittsburgh.[11] The increase was driven largely by gun homicides committed against and by young African American males.[12] Given the substantial residen-tial segregation by race in Pittsburgh and most other American cities, the con-centration of criminal gun violence among young black males leads to geographic concentration of gun homicides as well.[13]

This concentration suggests the opportunity for an intervention that narrowly targets resources on an identifiable law enforcement target—illegal gun carrying by youth in high-risk neighborhoods. Like other "directed patrol" efforts, Pitts-burgh's *firearm suppression patrol (FSP)* program assigned more police resources

---

10. Cook and Ludwig (2000); Ludwig and Cook (2001).
11. Blumstein (2000).
12. Cook and Laub (1998); Blumstein (2000).
13. Glacser and Vigdor (2001).

to selected high-crime areas. These patrols were relieved from responding to citizen requests for service (911 calls) in order to work pro-actively to search for illegally carried guns.[14]

Police contacts were initiated mainly through traffic stops and "stop-and-talk" activities with pedestrians in public areas. Carrying open alcohol containers in public and traffic violations were frequent reasons for initiating contact. When warranted for reasons of officer safety (usually because of suspicious actions or demeanor), these stops sometimes moved to the types of pat-downs on the outside of clothing to check for weapons that are allowed under the Supreme Court's 1968 decision in *Terry v. Ohio.*[15] When there was reasonable suspicion of criminal activity, the contact might escalate to more intrusive searches inside pockets, under coats, and in waistbands as part of an arrest.

The Pittsburgh policing initiative was constructed with concerns about individual rights and police-community relations in mind. Not long before the intervention was fielded, the Pittsburgh police department entered into a consent decree with the Department of Justice in response to complaints about abuse of force, most notably several police shootings of civilian residents. As part of the consent decree, the police department issued new regulations governing police contacts with citizens. Included were explicit guidelines on when officers could engage in *"Terry"* pat-down safety frisks and specific reporting requirements of the circumstances that precipitated more intrusive searches of persons or vehicles and seizures of guns or other property. Notably, officers had to articulate the basis for their suspicion about criminal activity by the person(s) being searched. Participating officers were specially selected by police command staff based on their demonstrated capacities in pursuing a proactive style of law enforcement tempered by a professional attitude and demeanor in citizen encounters.

The goal of the Pittsburgh policing program was to focus resources on those neighborhoods most in need. In Pittsburgh, as in many other cities, youth homicides (victims *12 to 24* years old) and citizen reports to police about shots fired are highly concentrated in a relatively small number of neighborhoods; however, unlike other cities, these high-crime neighborhoods are not contiguous in Pittsburgh. The city of Pittsburgh includes three distinct areas separated by rivers. The highest-crime neighborhoods are in police administrative zone 1 on the

14. The Pittsburgh initiative was pan ofa larger effort fianded by the Alfred P. Sloan Foundation that involved collaboration between university research teams and police in the cities of Pittsburgh, Pa., and Rochester, N.Y. Police in each city were encouraged to design their own strategies for addressing street-level gun violence. The results reported relate to the effort by police in Pittsburgh.

15. *Terry v. Ohio* (392 U.S. 1, 1968).

north side of the rivers and zone 5 in the eastern end of the city between the rivers. These two zones are the target areas for the FSPs.

Each of the target zones is a fairly large geographic area filled with neighborhoods that have different populations and problems with gun violence. The two zones each include about thirty-five census tracts and fifteen neighborhoods, spread out over nearly ten square miles each and home to 55,000 and 80,000 residents, respectively. In around a third of the tracts in zone 1 and a fifth of those in zone 5, fewer than 5 percent of the residents are African American. However, each zone also contains seven tracts in which more than half of all residents are black. Amid the high-crime neighborhoods in the target zones are many census tracts that experienced no youth homicides at all in the recent peak year, 1993.

Under the FSP program one additional patrol team was assigned to both zones 1 and 5, consisting of four officers and a sergeant (all in uniform) traveling in three vehicles—usually two marked patrol cars and one unmarked car. The teams in each zone worked four-hour shifts from 8 p.m. to midnight twice weekly for fourteen weeks from July 19 to October *24, 1998*. These patrols were focused on the high-crime evenings of Wednesday through Saturday nights. Specific patrol days were designated to ensure a mix of different days covered in each zone. The most common pattern (found in half the weeks) was alternating days, either Wednesday and Friday or Thursday and Saturday, in individual zones. During this period, fifty-one special patrol details were fielded across the two zones involving nearly 1,000 officer-hours (including the sergeants' time).

With the assistance of maps and reports of recent shots-fired activity, patrol teams identified and targeted "high-risk places at high-risk times," looking for opportunities to initiate citizen contacts for the purpose of soliciting information and investigating suspicious activity associated with illegal carrying and use of guns.[16] The earlier Indianapolis program pursued a place-based strategy in one part of the city, which focused on maximizing the number of traffic stops, and a person-based strategy in another area, which focused on stopping only the most suspicious people within the target areas. Pittsburgh's program falls somewhere between these two models—the program used pedestrian and traffic stops and included some focus on suspicious people, but stops were not limited to this group-

Implementation of this strategy in Pittsburgh differed slightly between the two treatment areas. In zone 5 the three police vehicles typically traveled together as a unit, while in zone 1 the vehicles patrolled individually. Despite the greater dispersion of police patrols in zone 1, the number of police contacts recorded was

---

16. Sherman (2001).

greater in zone 5—perhaps because of this area's higher initial crime level. Compared with zone 1, table 6-1 shows that zone 5 experienced around twice as many vehicle stops (27 versus 12), person contacts (118 versus 57), arrests (12 versus 6), and confiscated guns (5 versus 2).

Given the size of these target zones the "dosage" of the intervention may seem low in absolute terms—just four-hour patrol details twice weekly in each targeted patrol zone, covering under 5 percent of all available hours weekly. However, the patrols covered more than 15 percent of high-risk times from 7 p.m. to 1 a.m. daily and 30 percent of high-risk times on high-risk days (Wednesday to Saturday) weekly. Moreover, the three-vehicle, five-officer teams represented a large increment to customary patrol resources in the target police zones. Police vehicles increased by 20 percent and patrol officers by 25 percent in target zone 5, the city's highest crime zone. The increases were even larger in the other target (zone 1), with a 35 percent increase in vehicles and a 50 percent increase in officers.

Data

Our goal is to focus on outcome measures that capture illegal gun carrying and criminal misuse in Pittsburgh. Gun homicides are an obvious choice, although such events are too rare to be useful given our research design. More frequent events such as gun robberies or assaults provide another possibility, although previous research suggests that standard police incident reports are often unreliable about whether a gun was involved in these types of criminal events.[17] As a result our emphasis is on measures of citizen reports to the police of shots fired and on gunshot injuries treated in hospital emergency departments.

Data on shots fired come from Pittsburgh's 911 Emergency Operations Center and include information about the date, time, and address of the reported incident. These data allow us to identify whether the events occurred in the treatment or control zones during the periods that the policing program was in effect, which we define as 8 p.m. to midnight on those days the firearm suppression patrols were deployed in target areas (Wednesday through Saturday evenings).

Because discharging a firearm within the city limits of Pittsburgh is against the law, our measure of shots fired captures an event that is itself technically a "gun crime." But more important, we expect shots fired to be strongly related to the prevalence with which guns are carried in public spaces by high-risk people who are willing to use them, even if only to show off.

17. McGarrell and others (2001).

Table 6-1. *Enforcement Activities during Firearm Suppression Patrols*

Activities	Zone 1	Zone 5	Total
Person contacts	57	118	175
With pat down/search	13	21	34
No pat down	*44*	97	141
Vehicle stops	12	27	39
Stolen vehicle recovered	1	3	4
Citations	4	21	25
Vehicle	3	11	14
Open container/alcohol	1	7	8
Disorderly/noise /nuisance	0	3	3
Warnings /other[a]	17	37	54
Search/seize	2	5	7
Gun and, no other contraband	1	1	2
Gun and other contraband	I	0	1
Other contraband alone	0	4	4
Nothing found	11	17	28
Arrests	6	12	18
With gun	1	2	3
No gun	5	10	15
Guns confiscated	2	5	7
General activities with no contacts			
Pursuit—no contact[b]	17	24	41
Patrol—no contact[c]	171	*46*	217
Assist nonfirearm suppression unit	32	30	62
Nonfirearm suppression activities[d]	3	3	6
911 calls (total)	55	50	105
Person shot	0	2	2
Shots fired	13	21	34
Person(s) with gun[c]	16	8	24
Stolen gun	0	0	0

a. "Other" includes warnings without citations and requests that individuals "move along."

b. Combination "viewed only" and "pursuit—no contact" used for actors who fled or dispersed on police arrival.

c. "Patrol—no contact" also includes a few instances of stationary surveillance of an area and tactical foot patrols.

d. "Non-FSU activities" include supervision of officers and investigation of possible stolen car parked on street.

c. "Person(s) with gun" includes armed robbery calls.

One complication is that because shots fired can often be detected over a wide area, duplicate calls for the same incident may occur. We attempt to address this problem by eliminating duplicate calls that shared the same event number in the Emergency Operations Center system or calls that reported shots fired within five minutes and 2,000 feet of one another. We also eliminated reports that lacked information about the exact location of the event, since without this address information we would be unable to identify duplicate calls for the same event. Taken together these criteria eliminate 27 percent of the 9,884 original shots-fired reports in our sample.[18] Of course our procedure will not eliminate duplicate calls that are reported by residents who live more than 2,000 feet apart.

Another complication is the difficulty in pinpointing the exact location of shots-fired incidents, especially in urban neighborhoods where gunfire is common enough to make residents alert for similar sounds. Although we do not have any direct measure of this problem, gun suppression officers responding to shots-fired calls were unable to verify an actual incident in three out of four calls. In these cases witnesses or callers could not be located at or near the scene, and police could not locate physical evidence (such as shell casings or bullet holes) of shots being fired. The lack of verification does not mean that a gun was not fired, since witnesses or callers may not bother to meet with the police if they believe the clanger has passed or if the police report to the scene with some delay. Finding shell casings on the street may be difficult if the event occurs at night and witnesses are lacking or uncertain about the exact location of the shooting, or even impossible when a revolver is used (since these guns do not automatically expel the shell casing after firing).

While there necessarily remains some uncertainty about the shots-fired data, we nevertheless believe that this measure is associated with gun carrying and misuse. Support comes from the feet that the frequency of shots-fired calls and youth homicides are highly correlated in both the cross section across Pittsburgh census tracts and over time for the city of Pittsburgh as a whole.

Gunshot injuries treated in hospitals serve as a complementary outcome measure, one that is not subject to the same potential reporting problems as shots fired and which provides a more direct indicator of gun violence. Data came

18. Of the 9,884 original shots-fired calls, 238 are eliminated because they are exact duplicate records (that is, they have the same event number, address, and time), 986 are eliminated because we do not have exact address information on the event, and 1,419 are eliminated because these reports are within five minutes in time and 2,000 feet in distance from another call, leaving us with a total of 7,241 shots-fired calls in our final analysis. There do not seem to be any substantial differences across Pittsburgh police zones in the fi action of the original shots-fired calls that are eliminated undet these criteria.

from the injury surveillance system developed by the Allegheny County Health Department in cooperation with the Harvard Injury Control Research Center, which collects and analyzes data on gunshot injuries treated in four trauma centers at area hospitals. Together, data from these trauma centers capture more dian 90 percent of gunshot injuries treated in hospitals in the Pittsburgh area.[19]

These data include information about the demographic attributes of the victims, the nature of the injury, and the circumstances of the event (assault, self-inflicted, or accidental). We focus on gunshot injuries from assaults since these events should be most sensitive to the policing intervention, although we also explore the sensitivity of our findings to different subsets of gunshot injuries. For privacy reasons the data do not include individual identifying information such as the victim's exact street address. However, we do have the victim's zip code of residence, which allows us to locate victims among the large police zones used in this project. Of the 1,125 gunshot injury reports in our sample, we lose only thirty-five cases because of missing information on the victim's zip code.

How well does zip code information on the victim's address capture criminal events that occur within the same police beat? Analysis of the 328 homicides that occurred in Pittsburgh from 1990 to 1995 suggests that the use of residence zip code data performs fairly well: in 81 percent of cases the victim lives within the same police zone in which the murder occurred. The offender lives within the same police zone in which the murder takes place in 69 percent of homicides.

Methods

In the absence of random assignment of Pittsburgh's policing program across neighborhoods, the challenge is to isolate the causal effects of the intervention from those of other factors that drive variation in crime rates across communities over time. As with any nonexperimental study there necessarily remains some question about whether this analysis has successfully identified the program's effects. Nevertheless, some unique features of Pittsburgh's policing program, including the fact that the gun patrols were implemented on some days of the week but not others within targeted neighborhoods, help us eliminate various competing explanations for the crime changes observed within the target areas.

Our research design, as well as some commonly used alternatives, can be illustrated by using table 6-2. Policymakers and reporters often judge the success

19. For more information about these data, see www.hsph.harvard.edu/hicrc/nviss. In principle, perpetrators who are shot by victims or police during the commission of a crime may avoid medical treatment for fear of being arrested, but we suspect that such cases are rare in practice. See Azrael and others, chapter 10, in this volume.

Table 6-2. *Research Design for Pittsburgh Policing Evaluation*

	Preperiod (6 weeks before)	Postperiod (14 weeks during)	Estimated differences
Part A: Comparing overall averages			
Target zones	A	B	(B – A)
Control zones	C	D	(D – C)
DD			(B – A) – (D – C)
Part B: Exploiting within-week variation in patrol activity			
Target zones			
Wednesday–Saturday	E	F	(F – E)
Sunday–Tuesday	G	H	(H – G)
DD_T			(F – E) – (H – G)
Control zones			
Wednesday–Saturday	I	J	(J – I)
Sunday–Tuesday	K	L	(L – K)
DD_C			(J – I) – (L – K)
$DDD = DD_T - DD_C$			[(F – E) – (H – G)] – [(J – I) – (L – K)]

Note: Estimates of program effects rely on various differences that compare outcomes in different subsets of the data. DD is a difference-in-differences, and DDD is a difference-in-difference-in-djfferences.

of programs such as the Pittsburgh FSPs by examining whether crime rates or other outcomes decline within the jurisdiction once the program is put into place. In terms of table 6-2, this type of estimate would come from comparing the average number of gunshot injuries or shots fired per day in the treatment zones (1 and 5) during the fourteen weeks of the program from July 19 to October 24, 1998, represented by the letter B in the top part of table 6-2, with what is observed during the six-week preprogram period from June 7 to July 18, given by letter A in table 6-2.[20] The obvious problem with this before-after approach is that crime rates change over time in a generally cyclical fashion at the local, state, and national levels—often dramatically—for reasons that remain poorly understood.[21]

An alternative approach with its own limitations comes from comparing the average number of gunshot injuries or shots fired in the treatment zones during the postprogram period (letter B in table 6-2) with the control zones over the same time frame (letter D in table 6-2). The primary limitation with this cross-sectional comparison is that the treatment zones have persistently higher crime

20. We initially define a six-week period as "pre-program" to keep this time frame within the high-crime summer months, although we also explore the sensitivity of our findings to different definitions.
21. Blumstein and Wallman (2000).

rates than the control zones, even before the Pittsburgh FSP program is enacted. Criminologists often try to address this problem by using multivariate regression to control for observable differences across areas in population and other local-area characteristics. But if, as seems likely, the zones systematically differ in ways that we cannot readily measure, we will inappropriately attribute differences in crime among them to the effects of the policing program rather than to the unmeasured factors.

To address these "omitted-variables" problems, a third alternative is to focus on comparing how outcomes change in the treatment and control areas from the six-week "preprogram" period (June 7 to July 18,1998) versus the fourteen-week postprogram period. This so-called difference-in-differences (DD) estimate comes from comparing the change (or difference) in outcomes in the treatment zones (given by B -A in the top panel of table 6-2) with the change in the control zones (D - C). The estimate of program impact in this case [DD = (B - A) minus (D - C] is unbiased if the unmeasured differences between the treatment and control zones in Pittsburgh remain fixed over the sample period.

However, if there are unmeasured variables that change over time in ways that would cause the treatment and control areas to experience different *trends* in shots fired or gunshot injuries, the DD estimate will yield biased estimates for the program's impact. For example, large and small cities follow different crime trajectories over time in the United States.[22] Attempts to evaluate the effects of big city interventions by comparing their crime trends with those in small cities are likely to confound the effects of the programs of interest with those of whatever other factors lead to divergent crime experiences over time by city size.

One important check on the reliability of the DD estimation approach is to examine whether treatment and control areas follow similar trends *before* the intervention is enacted.[23] In any case, while the DD estimate improves on both the standard before-after and cross-section research designs just discussed, the approach remains vulnerable to bias introduced by time-varying omitted variables that affect crime trends as well as levels in the treatment and control neighborhoods.

In our evaluation we attempt to account for unmeasured, time-varying variables that may cause treatment and control areas to have different trends by exploiting the fact that the patrols are implemented on only selected days of the week (Wednesday through Saturday evenings). As a result observations for the off days (Sunday through Tuesdays) in the treatment areas can serve as an addi-

22. Biumstein (2000).
23. Bassi (1984); Heckman and How (1989); Smiih and Todd (forthcoming).

tional "control group" for measuring the program impact on gun crime in the on days. Put differendy, to the extent that unmeasured variables cause the treatment and control areas to have different crime trends throughout the week, comparing on days with off days within the treatment areas should control for these confounding trends and help isolate the effects of the new police patrols.

This *"difference-in-difference-in-differences"* (DDD) estimate can be described more formally with the notation outlined in the bottom panel of table 6-2.[24] The difference (F — E) represents how shots fired or gunshot injuries change from the pre- to postprogram period on Wednesdays through Saturdays in the treatment zones of Pittsburgh. As already noted, our focus on changes in this analysis helps overcome the fact that some neighborhoods have persistendy higher crime rates year after year compared with other areas. To account for the possibility that factors specific to the treatment zones may drive crime changes over time, we compare changes in the treatment areas during the on days (F - E) with changes on the off days (H - G), or $[DD_T = (F - E) - (H - G)]$.

Of course the high-crime evenings of Wednesday through Saturday may simply follow different crime trends than the lower-crime evenings of Sunday through Tuesday throughout the city of Pittsburgh as a whole. To account for this possibility, we compare the relative change over time in the treatment areas for Wednesday through Saturday versus Sunday through Tuesday $[DD_T = (F — E) — (H — G)]$, with the within-week change over the same period that is observed in the control areas $[DD_C = (J - I) - (L - K)]$. The DDD estimate in this case is given by $[(F - E) - (H - G)] - [(J - I) - (L - K)]$. This research design helps isolate the effects of those factors specific to the treatment zones during the on days of the postprogram period—factors such as the new FSPs introduced by the Pittsburgh police.

Our estimate for the effects of the Pittsburgh policing program can be derived from a simple regression framework as shown in equation 1. Let Y_{it} represent the number of either shots fired or gunshot injuries on day *(t)* within neighborhood *(i)* in Pittsburgh. The explanatory variables in the regression consist of a series of simple dichotomous indicators where *Treat$_i$* is equal to one if the neighborhood is in the treatment zones (1 and 5) and equal to zero otherwise, *Post,* is equal to one if the day falls within the fourteen-week period that the policing program is in effect and equal to zero for the six-week preprogram period, and *Day,* equals one if the observation is for a Wednesday, Thursday, Friday, or Saturday, the on days of the week when the police patrols may be operating.

24. This research design has also been applied ro study the effects of Medicaid benefits for maternity and pediatric care on abortion *(Joyce* and Kaestner, 1996), mandated maternity benefits on child bearing (Gruber, 1994), permissive gun-carrying laws (Ludwig, 1998), and the Richmond, Virginia, Project Exile program. See Raphael and Ludwig, chapter 7 in this volume.

$$Y_{it} = b_0 + b_1 Treat_i + b_2 Post_t + b_3 Day_t + b_4 (Treat_i) \times (Post_t)$$
$$(1) \qquad + b_5 (Day_t) \times (Post_t) + b_6 (Treat_i) \times (Day_t)$$
$$+ b_7 (Treat_i) \times (Day_t) \times (Post_t) + e_{it}$$

The DDD estimate in this case is given by the coefficient *by.* In our analysis we present robust standard errors that are adjusted to account for heteroskedasticity as well as nonindependence of observations drawn from the same neighborhood.[25] Only omitted covariates that vary on a daily basis have potential for affecting model estimation. More enduring factors that are typically thought to affect crime, such as demographic and economic variables, will not influence program impacts estimated on a daily basis. Furthermore, the differencing strategy of the DDD research design should account for most of the potentially confounding factors that vary over time across neighborhoods.[26]

Results

Our central finding is that the Pittsburgh FSPs appear to substantially reduce citizen reports of shots fired and gunshot injuries in the target neighborhoods. While we find some evidence of a "phantom effect" for shots fired in 1997, the year before the police patrols are put into place, the findings for gunshot injuries generally hold up to a variety of specification checks.

Shots Fired

Table 6-3 shows our key findings for the average number of shots-fired reports per day in Pittsburgh neighborhoods. In the top panel of table 6-3 is the widely used difference-in-difference estimate, which compares the pre- to postprogram change in shots fired per day averaged across all days of the week in the treatment and control areas. These calculations show that the number of shots-fired reports declined in the treatment zones of the city by -.066 during the four-hour

---

25. Estimation uses STATA (version 7) software to perform OLS regressions with the "robust cluster" options to allow for an arbitrary variance-covariance error structure.

26. Consistent with this expectation, the results are unchanged when the estimating equation includes fixed effects for each of die police zon^s individually, measures of time and time squared to capture citywide crime trends, or indicators for weeks when school was in session or weekend nights (Friday and Saturday).

Table 6-3. *Impact Estimates for Shots Fired (daily averages per police zone)*

	Preperiod (6 weeks before)	Postperiod (14 weeks during)	Estimated differences
Part A: Standard difference-in-difference estimate			
Treatment zones	.750	.684	-.066
Control zones	.274	.327	.053
DD			-.119 (.152)
Part B: Exploiting within-week variation in patrol activity			
Treatment zones			
Wednesday–Saturday	.979	.670	-.310
Sunday–Tuesday	.444	.702	.258
DD_T			-.567 (.088)**
Control zones			
Wednesday–Saturday	.323	.281	-.042
Sunday–Tuesday	.208	.387	.179
DD_C			-.220 (.120)*
DDD			-.347 (.133)**

Note: Results come from estimating daily average shots fired during the four hours from 8 p.m. to midnight in each of Pittsburgh's six police zones. Standard errors in parentheses are adjusted to account for heteroskedasticity in the error variance across different zones, and nonindependence of observations drawn from die same police zone. Asterisks identify statistically significant reductions in one-tail z tests.

  * Significant at 5 percent level.

  ** Significant at 1 percent **level.**

treatment window from 8 p.m. to midnight (hereafter "per day" for these shots-fired results), while this figure increased in the control areas by +.053. The DD difference in simple trends is thus equal to -.119, which implies a greater decline in the treatment than control neighborhoods, but one that is not statistically significant.

More compelling evidence of a program impact arises when we exploit variation across days of the week in the timing of the patrols, as seen in the bottom panel of table 6-3. Once the policing program was implemented, the number of shots fired declined by —.310 per day during the on days in the treatment zones. However, the number of shots fired increased by +.258 in the same treatment neighborhoods during the off days. The gap between the on and off days is a decline of D D $_T$— —.567 shots-fired calls. In the control neighborhoods that did not receive the policing program, the gap in shots fired between the high- and low-crime days of the week declined more modestly, DD_c = -.220. The difference-in-difference-in-differences estimate is thus DDD = -.567 - (-.220) = -.347, statistically significant at the 1 percent level using a one-tailed test for a reduction in shots fired. The estimate implies that the policing program

Table 6-4. *Robustness Checks fir Estimated Program Effects on Shots Fired (daily averages per zone)*

	DDD (difference-in-difference-in-differences)			
	Target zone 1		Target zone 5	
Estimating data	Beta	(se)	Beta	(se)
1998 program estimates				
6 weeks preprogram versus 14 weeks during program	−0.435**	(0.117)	−0.260**	(0.117)
14 weeks preprogram versus 14 weeks during program	−0.481**	(0.083)	−0.165**	(0.083)
6 weeks preprogram versus first 7 weeks during program	−0.329*	(0.213)	−0.172	(0.213)
6 weeks preprogram versus second 7 weeks during program	−0.540**	(0.080)	−0.348**	(0.080)
5 weeks preprogram versus 14 weeks during program	−0.378**	(0.090)	−0.034	(0.090)
Target zones versus control zones 3, 4, and 6	−0.504**	(0.132)	−0.329**	(0.132)
Nonprogram years				
1997 data, 6 weeks preprogram versus 14 weeks during program	−0.030	(0.102)	−1.117**	(0.102)
1997 data, 14 weeks preprogram versus 14 weeks during program	−0.266**	(0.107)	−0.915**	(0.107)
1999 data, 6 weeks preprogram versus 14 weeks during program	0.098	(0.180)	0.061	(0.180)
1999 data, 14 weeks preprogram versus 14 weeks during program	0.104	(0.103)	0.027	(0.103)

Note: Unless noted otherwise, all contrasts are between each target zone (1 or 5) and all control zones (2, 3, 4, and 6) in the six-week preprogram and fourteen-week postprogram periods. Estimates come from comparing changes over time in the daily average number of shots fired during the four hours from 8 p.m. to midnight in treatment versus control police zones (table 6-2). Standard errors in parentheses are adjusted to account for heteroskedasticity in the error variance across different zones and non-independence of observations drawn from the same police zone. Asterisks identify statistically significant reductions in one-tail z tests.

* Significant at 10 percent level.
** Significant at 5 percent level.

1997 and 1999, we expect no statistically significant differences to arise with the DDD approach. Of course the 1999 test is perhaps not as clean as that for 1997, given the possibility of "residual deterrence" of criminals even after the patrols have ended, but in any case evidence of a null effect in 1999 would provide useful evidence on the validity of our research design.

As seen in the bottom panel of table *6-4,* while we find no statistically significant evidence of a phantom program "effect" in 1999 for either of the two treatment zones (1 and 5), we do find signs of a phantom effect in 1997, the year before the police patrols were launched. When we use a six-week preprogram period for 1997, the contrast between treatment zone 5 and the controls is statistically significant. Alternative definitions, ranging from five- to fourteen-week preprogram periods, yield statistically significant phantom effects in both zones 1 and 5- These findings make us cautious about interpreting the differences in shots fired between the treatment and control zones in 1998 as signs of the patrol program's effects. However, the results for gunshot injuries presented in the next section are more robust to our various specification checks.

Gunshot Injuries

The results for assault-related gunshot injuries show an even more pronounced program impact during the program year (1998) and little consistent evidence of statistically significant program effects during the years before and after the program is in effect. Gun assault injuries declined significandy in both of the target zones by roughly the same proportional amount, although the decline is only statistically significant over a variety of conditions in the higher-crime zone 5.

The top panel of table 6-5 shows that the standard difference-in-difference estimate that relies on average gunshot injuries per day (averaged across all days of the week) suggests that the program has reduced such injuries by —.073 per day. The bottom panel of table 6-5 shows that refining the simple difference-in-difference calculation to focus on the days on which the patrols were in operation (Wednesday through Saturday) increases the magnitude of program impact to (—.161 — .007) =—.168 fewer assault-related gunshot injuries on patrol days in the target zones.[28] The simple pre- and post-trends are much more similar during the off days (.103 and .050). While the difference in trends during the on days equals —.168, the difference in trends for the off days is only about one-third as large and represents an increase of .053 = (.103 - .050). Exploiting this within-week variation in patrol implementation more formally yields a DDD estimate equal to -.222 (p<.10), a reduction in assault gunshot injuries

28. Daily rates refer to the average number of gunshot injuries during a **full** twenty-four-hour day.

Table 6-5. *Impact Estimates for Assault Gunshot Injuries*
(daily averages per police zone)

	Preperiod (6 weeks before)	*Postperiod* (14 weeks during)	*Estimated differences*
Part A: Standard difference-in-difference estimate			
Treatment zones	.155	.107	−.048
Control zones	.054	.079	.026
DD			−.073 (.022)**
Part B: Exploiting within-week variation in patrol activity			
Treatment zones			
Wednesday–Saturday	.250	.089	−.161
Sunday–Tuesday	.028	.131	.103
DD_T			−.264 (.208)*
Control zones			
Wednesday–Saturday	.073	.080	.007
Sunday–Tuesday	.028	.077	.050
DD_C			−.042 (.042)
$DDD = DD_T - DD_C$			−.222 (.165)*

Note: Results come from estimating daily average gunshot injuries from assaults in each of Pittsburgh's six police zones. Standard errors in parentheses are adjusted to account for heteroskedasticity in the error variance across different zones and non in dependence of observations drawn from the same police zone. Asterisks identify statistically significant reductions in one-taii z tests.

* Significant at 10 percent level.
** Significant at 5 percent level.

of 71 percent f .222/(.222 + .089)) from the expected level on patrol days in the target zones.

Table *6-6* shows that the results are generally not sensitive to how we define the pre- or postprogram periods. The estimated program impact on assault-related gunshot injuries in zone 5 is always statistically significant and of about the same magnitude when we partition the postprogram period in half, and whether we define the preprogram period using the five weeks before the patrols go into effect, the fourteen weeks before the patrols are launched, or any interval in between. The estimated impact in zone 1 is also statistically significant but only for longer preprogram periods (eleven weeks or more). Statistically significant proportional changes in assault gunshot injuries are approximately similar in magnitude across the two zones, ranging from 59 to 77 percent in zone 1 compared with 60 to 72 percent in zone 5.[29]

29. In shorter preprogram per.ods the proportional reductions are in the same range in zone 5 but are smaller in zone 1.

Table 6-6. *Robustness Checks for Estimated Program Effects on Assault Gunshot Injuries*

| | DDD (difference-in-difference-in-differences) | | | |
| | Target zone 1 | | Target zone 5 | |
Estimating data	Beta	(se)	Beta	(se)
1998 program estimates				
6 weeks preprogram versus 14 weeks during program	−0.015	(0.041)	−0.428**	(0.041)
14 weeks preprogram versus 14 weeks during program	−0.058**	(0.026)	−0.243**	(0.026)
All gunshot injuries (all causes)	−0.058	(0.052)	−0.542**	(0.052)
All gunshot injuries (youth only)	−0.061	(0.077)	−0.319**	(0.077)
Accidental gunshot injuries	−0.047**	(0.028)	−0.039*	(0.028)
6 weeks preprogram versus first 7 weeks postprogram	−0.036	(0.029)	−0.378**	(0.029)
6 weeks preprogram versus second 7 weeks postprogram	0.006	(0.065)	−0.479**	(0.065)
5 weeks preprogram versus 14 weeks postprogram	0.042	(0.032)	−0.348**	(0.032)
Target zone versus control zones 3, 4, and 6	−0.021	(0.055)	−0.434**	(0.055)
Nonprogram years				
1997 data, 6 weeks preprogram versus 14 weeks postprogram	−0.050	(0.051)	−0.052	(0.051)
1997 data, 14 weeks preprogram versus 14 weeks postprogram	0.028	(0.051)	−0.037	(0.051)
1999 data, 6 weeks preprogram versus 14 weeks postprogram	0.029	(0.016)	−0.118**	(0.016)
1999 data, 14 weeks preprogram versus 14 weeks postprogram	0.031	(0.010)	0.007	(0.010)

Note: Unless otherwise noted, all contrasts are between each target zone (1 or 5) and all control zones (2, 3, 4, and 6) in the six-week preprogram and fourteen-week postprogram periods. Estimates come from comparing changes over time in the daily average number of gunshot injuries in treatment versus control police zones (table 6-2). Standard errors in parentheses are adjusted to account for heteroskedasticity in the error variance across different zones and non independence of observations drawn from the same police zone. Asterisks identify statistically significant reductions in one-tail z tests.

 * Significant at 10 percent level.
 ** Significant at 5 percent level.

The pattern of large and significant effects in zone 5 and insignificant effects in zone 1 persists when we vary the subset of gunshot injuries. Zone 5 experiences significant declines in accidental gunshot injuries, the combination of all gunshot injuries, and gunshot injuries involving youthful victims (under age 25). In zone 1, the effects are smaller in magnitude and only reach statistical significance for accidental gunshot injuries.[30]

The results are similar when we use a generalized least squares (GLS) approach that corrects for serial correlation in the error structure of our regression model, or when we use models that weight each observation by the zone's population.[31] We also obtain similar findings when we reduce some of the day-to-day variation in gunshot injuries by using a panel data set that averages gunshot injuries for each zone within a week over the on days and over the off days. In this case each zone contributes only two observations per week to the panel data (one each for the on and offsets of days), rather than seven daily observations per zone per week as in our baseline estimates.

There is also suggestive evidence that our estimates for gunshot injuries are not driven by unmeasured confounding factors, shown in table 6-6. We replicate the DDD calculation using the same calendar days, neighborhoods, and days of the week but now using assault gunshot-injury data for 1997, when the program was not in effect. As with our analysis of the program year of 1998, we examine the sensitivity of our estimates to how we define the preprogram period, ranging from five to fourteen weeks. For the twenty total DDD estimates that we calculate for the preprogram year of 1997 (each of the two treatment zones individually against the control areas, using ten possible definitions of the preprogram period for each comparison), two are statistically significant at the 10 percent level. We obtain a similar fraction of statistically significant comparisons for 1999. While each point estimate is of course not a truly independent trial, we are not alarmed that two of the twenty comparisons are significant at the 10 percent level in 1997 and 1999. By way of comparison, during the program year of 1998 fully fourteen of the twenty possible comparisons are statistically significant at the 5 percent level.

---

30. The declines in accidental gunshot injuries are not likely to result from the policing program since there is evidence of even larger declines in both target zones during 1997, and in zone 1 during 1999, which further motivates our focus on the assault-related gunshot injuries that better reflect the type of violent behavior targeted by the police patrols.

31. We estimate GLS models that assume that the error terms for periods *(t - I)* and *(t)* are correlated but error terms more than one period apart are not (a first-order auto regressive process). The results are the same whether we assume that the serial correlation between error terms is the same or differs across police zones. In both cases the point estimate for the effect of the policing intervention increases in relation to the standard error and improves the significance level from .09 to .01.

Moreover, in 1998, after the fourteen-week patrol period is over, we observe a statistically significant increase in assault gunshot injuries in zone 5 equal to around one-quarter of the decline observed during the program period. This bounce back is consistent with what we might expect from the cessation of a program that has a causal effect on gun carrying or crime and generates some residual deterrence. The bounce back is observed soon after the patrol program terminates (within two weeks), which suggests that at least some subset of people at high risk of carrying and misusing guns is reasonably well informed about the local police environment. In any case, there is no similar bounce back over the same calendar period in the treatment areas in 1997 and 1999.

Discussion

Our estimates suggest that Pittsburgh's targeted policing program against illegal gun carrying may have reduced shots fired by 34 percent and gunshot injuries by as much as 71 percent in the targeted areas. Our evaluation pays careful attention to the problem of unmeasured, time-varying factors that may introduce bias into estimates of the policing program's impact. Although we cannot definitively rule out the possibility of omitted-variables bias with our analysis, the fact that the combined pattern of changes in gunshot injuries during the program period in 1998 is qualitatively different from what is observed in 1997 or 1999 is at least consistent with the idea that the FSPs had some effect on illegal gun carrying or misuse.

The relatively modest number of guns confiscated and arrests made as a result of Pittsburgh's directed patrols (table 6-1) suggest that incapacitation of illegal guns or their owners is unlikely to drive the large reductions that accompanied the policing program. Moreover, if the patrols reduce shots fired or assault gunshot injuries primarily through an incapacitation effect, by taking illegal guns and those who carry them off the street, we might expect the effects to carry over to the days of the week when the patrols were not in effect. This type of carryover does not appear to occur. Instead, the similarity in off-day trends between the treatment and control zones provides at least suggestive evidence of a deterrent effect that is specific to the treatment zones during the on days when the program is in effect and only reduces illegal gun carrying and misuse during those times. The absence of significant trends upward in the control areas during the program period also seems to rule out spatial displacement, where gun-carrying offenders shift their activities from the treatment to control neighborhoods.

Finally, one important policy question is whether the targeting of the police resources on illegal gun carrying in high-crime areas is an important part of the

program's effects, or whether simply adding more resources to routine patrol activity would achieve similar ends. Unfortunately, the Pittsburgh data do not enable us to definitively rule out either possibility. Previous research by economist Steven Levitt suggests that a 10 percent increase in the number of police reduces violent crime by around 10 percent (that is, an elasticity of violent crime with respect to police officers of around —1.0), even if the additional police resources are deployed in standard ways.[32] Our findings suggest an elasticity of gunshot injuries with respect to additional targeted police officers on the order of-1.4.[33] However, given the standard errors around both sets of estimates we cannot confidendy conclude that targeting the additional police resources in Pittsburgh to focus on illegal guns enhanced the overall impact of these expenditures on criminal behavior.

However, because the Pittsburgh program targets police resources on the most costly violent crimes—those that involve firearms—the targeting seems to enhance the cost effectiveness of the additional police resources.[34] Although blanket increases in police resources may yield from $ 1 to $5 in benefits to society for each extra dollar that is spent, the more targeted FSPs in Pittsburgh may generate even more substantial net benefits.[35] We estimate that the costs of the additional patrols in Pittsburgh are modest—something less than $35,000 in overtime expenditures during the fourteen-week program period. In contrast the costs of gun violence to society are substantial, about $1 million per gunshot injury.[36] If the estimates presented in this chapter are correct, then the investment of $35,000 or so in targeted antigun police patrols may yield benefits of as much as $25 million.[37] Equally important, Pittsburgh's experience suggests that targeted patrols against illegal guns can be implemented in a way that addresses community concerns about intrusive policing.

32. Levitt (1997).

33. This comes from comparing the 71 percent decline in assault-related gunshot injuries to the 50 percent increase in the number of police officers allocated to zone 5 during the "on days" of the Pittsburgh firearm suppression patrols.

34. Levitt (1997, p. 285) reports an average cost per violent crime of about $70,000 {in 1998 dollars), while the costs per assault-related gunshot injury are about $ 1 million. Cook and Ludwig (2000); Ludwigand Cook (2001). Of course this comparison is not perfect since the class of violent crimes includes those that do not involve injury to the victim, but the average cost per violent crime seems to be driven in large part by homicides that mostly involve firearms,

35. Levitt (1997).

36. Cook and Ludwig (2000); Ludwig and Cook (2001).

37. The estimates in table 6-4 imply a reduction in the average daily number of gunshot injuries per zone equal to 0.222, which, when multiplied by four patrol days a week over fourteen program weeks, implies a total decline of around twenty-five gunshot injuries.

COMMENT BY
Lawrence W. Sherman

The chapter by Jacqueline Cohen and Jens Ludwig is a valuable addition to an important literature. With this analysis, we now have eight published tests of a major hypothesis: that intensive police efforts to discourage illegal gun carrying in public places can reduce gun injury or death. All eight of those tests produce findings consistent with that hypothesis. Most impressive is the medical data on gunshot injury that Cohen and Ludwig contribute. With their data, all eight tests show that police efforts reduce either homicides or gunshot wounds. Not each result is statistically significant, but the likelihood of achieving all eight results in the same direction by chance alone grows lower with this study's contribution.

   The focus on medically measured gunshot injury is a crucial decision for interpreting the sum total of available published research. The measures of gun crime in the literature have been mixed, given the failure of the FBI's Uniform Crime Reporting System to establish a separate category for crimes committed with guns, let alone for injuries occurring in crimes. As criminologists well know, FBI rules require counting all four of the following events as an "armed robbery":

   — Someone pulls out a knife and threatens to stab a victim if he does not yield money;

   — Someone points a gun and threatens to shoot a victim if he does not yield money;

   — Someone shoots at, but does not hit, a victim, who then yields money; and

   — Someone shoots at and wounds a victim and then takes the victim's money. Only when a victim dies as a result of a gunshot during a robbery do we obtain a dear measure of gun violence from the FBI's crime reporting system, followed by most urban police agencies in the United States. The robbery then becomes a murder but can be identified as a robbery murder by type of gun in the FBI's supplementary homicide reports.

   The occurrence of death in a gun crime is the tip of an iceberg, as rare as 1 death per 250 commercial robberies. Most gun crimes do not cause death and probably do not even cause an injury. Police data on the percentage of incidents in which police fire at suspects and hit their human targets range from 43 percent in New York City in 1970-91 to 46 percent in Kansas City in 1972-91 to a possible high of 58 percent in Los Angeles in 1980-91.[38] Thus about half of incidents of shooting by police regularly trained in firearms cause no injury or death.

---

   38. Geller and Scott (1992, pp. 516, 502, 519).

Of the shooting victims hit by shots fired by criminals in 4,177 Baltimore incidents in the mid-1990s, the monthly percentage who died from their injuries ranged from 11 percent to 20 percent, with an overall mean of 16 percent.[39] Applying that ratio to the roughly 50 percent of gunshot incidents with no wounds yields an estimate of some 8 percent of incidents in which shots are fired resulting in a death by gunshot wound, or 92 percent without medically treatable injury. The estimates of percentages of incidents in which guns are used without being discharged are controversial, but for commercial robberies alone they outnumber the cases in which guns are shot by 19 to 1 (95 percent), and for commercial robbery by 4 to 1 (22 percent).[40] In the case of commercial robbery, then, the tip of the iceberg of gun death is roughly 4 in 1,000 gun crimes.

In their attempts to measure gun crime more reliably, experimenters in police strategies against gun crimes have been driven to reading thousands of police incident reports to detect the use of a gun as described in the narrative of the incident.[41] A check on the reliability of Indianapolis police checking a box indicating "gun use" on crime incident reports found that 35 percent of crimes in which guns were used had no check on the relevant box in the report. James Shaw found drive-by gun crimes in Kansas City coded as "criminal damage to property" or "vandalism" when bullets fired had only hit buildings or cars.[42] These heroic efforts, however, give us little confidence that the measures of gun crimes are reliable over time or even well correlated with the more measurable tips of the iceberg. In the Indianapolis experiments, for example, homicides dropped in the east patrol area target beats from 4 to zero, while recorded aggravated assaults with guns rose from 19 to 30 and total gun crimes rose from 42 to 57. Since police disliked the mayor at that time, and since police controlled the decisions to report crimes, it is hard to put greater credence on the gun crime count than on the homicide count—the latter being less stable but less vulnerable to reporting discretion.

Cohen and Ludwig neatly sidestep these problems by turning to a measurement system that is devoid of police reporting biases. Hospital records on gunshot wounds by residence of the victim is an excellent approximation for the level of gun violence in the target beats. It raises from six to eight the number of published tests, to my knowledge, of the hypothesis that murder or wounding declines after the application of increased policing of gun carrying in a defined geographic area. Seven of those tests are based on a comparison to a control area,

---

39. Long-Onnen (2000, table 4).
40. Sec Cook (1983, p. 73).
41. Sherman and Rogan (1995); McGarrell and others (2001).
42. McGarrell and others (2001, p. 126); Sherman, Shaw, and Rogan (1995).

while one is an analysis of the entire United States over time without a control.[43] Those tests are as follows:

1992: Kansas City, Mo. Homicides and drive-by shootings (causing woundings) were significandy reduced in an on-off-on comparison of extra gun patrols in one area to no extra patrols in another area.

1993-94: Cali, Colombia. Homicides were 14 percent lower during eighty-nine police intervention days than nonintervention days, with bigger effects on paydays.

1995-97: Bogota, Colombia. Homicides were 13 percent lower on 67 intervention days than on nonintervention days, with bigger payday effects.

1997: Indianapolis east. Homicides declined from 4 to 0 in this beat while they rose 53 percent citywide and were unchanged in a comparison beat.

1997: Indianapolis north. Homicides declined from 7 to 1 in this beat while they rose 53 percent citywide and were unchanged in a comparison beat.

1998: Pittsburgh zone 1. Gunshot injuries declined significantly in this zone on intervention days, relative to both control zones and nonintervention days within target zones, for an overall reduction of 71 percent fewer injuries treated in target zones on intervention days.

1998: Pittsburgh zone 5- Same as in zone 1 above.

1984-98: USA. Homicides rose from by 20 percent 1984 to 1993 as the ratio of homicides to weapons arrests remained unchanged; once that ratio rose by 50 percent and remained much higher each year, the homicide rate began a steady drop to over 20 percent below the 1984 figure.[44]

Explaining Consistent Results

These results are encouraging but not clearly explained. The question of causal mechanism is hard to answer by the seizure of more guns: three of the eight (Pittsburgh 1 and 5 and Indianapolis north) did not report any increase in guns seized, and one (Bogota) had no data on this point; the other four report increases in guns seized. Whether the police focused on high-risk places or people is somewhat easier to resolve; only the Indianapolis north experiment reportedly focused on people, rather than on anyone looking suspicious in gun crime hot spots. But the experiments all lack specificity on this point, and it would require new research with systematic observation of police conduct to resolve it.

43. Sherman (2000).

44. For 1992, Sherman and Rogan (1995); for 1993-94 and 1995-97, Villaveces and others (2000); for 1997, East, and 1997, North, McGarrell and others (2001); and for 1998, Pittsburgh, Zones 1 and 5, Cohen and Ludwig, chap. 6, in this volume; for)984-98, U.S.A., Sherman (2000).

The most likely causal mechanism seems to be communication of a threat that police are looking for guns in certain places at certain times. That hypothesis is consistent with four of the eight tests (Pittsburgh and Colombia), which employed interventions that were limited to a few days at a time and then were turned off. How the threat was communicated, how widely it was received, and whether it would decay over time remain important and unanswered questions.

What Is An Experiment?

A final comment on the Cohen and Ludwig chapter is not a trivial point. They distinguish repeatedly between experimental and nonexperimental research. They imply that none of the eight tests we have, including their own, are "truly" experimental. This suggests that only a randomized experiment can be "true." As an ardent advocate of using randomized designs whenever possible, I appreciate the spirit in which their remarks are offered. But it does little good to reject the basic definition of an experiment as research on intentional changes in variables. All of these tests happened because someone decided to send police out to "get the guns off the street." That is a fundamentally different kind of test from a correlational analysis of "natural" differences in policing over time or across places.[45] We must recognize the Cohen and Ludwig chapter for being just as "experimental" as Boyle's tests or Newton's and not disparage the value of their careful methods. Our causal inference from experimental research is always limited, no matter how well bias is controlled. These tests collectively report effects whose sizes are large enough and consistent enough to draw substantial inferences even without random assignment.

COMMENT BY

Jeffrey Fagan

As crime rates in U.S. cities fell through much of the 1990s, "Do police matter" became a hotly contested question. On one side of this debate are researchers who suggest that police strategies directly and exclusively contributed to crime declines. These scholars claim that "but for" new police strategies, crime rates would have fallen neither as sharply nor as persistently as recent trends show.[46] On the other side are researchers who claim that crime has declined steeply

45. Cox (1958).
46. See, for example, Kelling and Sousa (2001); Silverman (1999); Heymann (2000).

across cities that applied quite different policing strategies, inviting explanations beyond policing to account for the steep declines in crime. Researchers looking across cities describe policing as one of many interacting social and economic forces that together produced downward pressures on crime.[47] Unfortunately, there is little hard evidence on either side of these claims.

Despite controversies in between-city comparisons, several studies *within* cities have shown that police practices can make a difference through small-scale, carefully targeted interventions. Many of these innovations have targeted guns, some with impressive results. These studies suggest that directed patrol practices and other selective and targeted strategies have produced declines in gun crimes that exceed the general regression in most cities. In Boston, police and other law enforcement agencies parsimoniously selected individuals for police surveillance and interdiction. In Indianapolis and Kansas City, police used directed patrols and proactive stops of citizens to reduce gun carrying and gun crimes in certain areas where gun violence rates were highest. Police in Chicago and San Diego engaged citizens in ongoing analysis of crime data to identify situations and locations with the highest crime rates and strategically focus interventions in those areas. Police in Jersey City targeted violence reduction strategies at specific problems within public housing projects to reduce violence.[48] The Cohen and Ludwig analysis of the *Firearms Suppression Program* (FSP) in Pittsburgh adds to this growing dossier of empirical evidence on the salutary effects of directed patrol and tightly focused proactive policing. The Pittsburgh Police Department launched an intervention in two of its six police districts that targeted guns via "gun suppression patrols." Cohen and Ludwig estimate the effects of these patrols on gun violence using two measures: gunshot injuries and gun shots fired. In the absence of an experimental design, Cohen and Ludwig use a "difference in difference in differences" (DDD) design to address three research challenges: heterogeneity of both crime problems and social structural factors in the six police districts, an intermittent application of the "treatment," and a relatively small number of observational units. Their expression of the DDD approach is well described in the chapter.

The structure of FSP creates a challenging identification problem and addressing it with this design is at once the strength and weakness of the chapter. The DDD model fits well with the quasiexperimental conditions. The third "D" is a time parameter to estimate the effects of the intermittent intervention by isolating the nonexperimental time intervals in the treatment sites. A fourth

---

47. Eck and Maguire (2000); Fagan, Zimring and Kim (1998); Bowling (1999); LaFree (1998).
48. Sherman (2000); Kennedy (1997); McGarrell and others (2001); Sherman and Rogan (1995); Skogan and Hartnett (1997); Greene (1999); Braga and others (1999).

"D" in this paper includes additional time periods to test for possible regression effects that predate the program. The method has been used productively by Jens Ludwig in his study of concealed firearm laws and in several studies of social welfare policy initiatives.[49] To avoid measurement problems associated with official statistics, Cohen and Ludwig use two independent measures of firearm activity in each of the two police zones where FSP was implemented. First, they use firearm injury surveillance data to estimate gunshot injuries. These public health records offer the comparative advantage of avoiding police "filtering" of reports. Second, they balance surveillance data with police reports of gun shots fired, carefully cleaned to reduce repeats and false reports. However, two facts suggest the possibility of reliability problems in shots-fired data: the elimination of 20 percent of these calls because of duplications and the inability of officers to validate actual incidents of shots fired in three of four calls. Nevertheless, the use of two alternate measures and data sources is first-rate measurement and design strategy.

The results were positive though internally contradictory and inconsistent. A significant and sizable reduction in shots fired was observed in both target districts. A significant and sizable reduction in gunshot injuries was observed in zone 1 but not in zone 5. For injuries, there were significant differences when zone 1 was included with zone 5, but there were no significant differences when zone 1 was analyzed separately. Moreover, there were no changes in gunshot injuries attributed to accidents, suggesting that perhaps the overall level of firearm possession remained the same but gun owners in zone 5 perhaps were less likely to carry their guns. Because of measurement problems in estimating gunshots fired, I put more stock in the injury data and hence the second analysis.

The accuracy of these estimates depends on assumptions in the modeling strategy that are provocative. First$_t$ the DDD model eschews statistical controls to estimate differences between observational units. This strategy fits well with state-level data to examine the effects of policy, where endogeneity of policy and structural variables may be less prominent in the adoption of policies to be tested. Besides, in the FSP study, measures for small areas such as census tracts were not available, a practical reason to use DDD models. Indeed, one of the strengths of DDD designs is their ability to remove endogeneity or heterogeneity as a confounding effect. DDD solutions assume that the between-unit differences in exogenous factors are accounted for in the initial differences among police zones in gun violence, and that changes over time in these differences are attributable solely to changes in the application of the intervention. The effects of the covariates can be "removed" because they are constant across time periods.

49. Ludwig (1998).

In other words, the factors that produce the initial differences—whether crime problems or social structure—are assumed irrelevant, because they are assumed to be invariant over the relatively short period m the study and hence remain stable among zones.

This is a Bayesian gamble, and the downside risk is not small. The covariates that are set aside in a DDD design are likely predictors ofboth crime problems *and the* decision to target the neighborhood.[50] Accordingly, there may be alternate contemporaneous effects that are not accounted for in the design. Put another way, crime problems and social structure may be endogenous to the selection of the areas to implement the intervention, a problem not easily addressed by the DDD design and its inherent treatment of covariates. Consider one example: since the two target zones were-selected because of their high crime rates, general police practices may differ in these two places. If such contextual factors interact with the interventions, then they cannot be removed. This is a classic threat in quasi-experiments that may well confound the design.[51]

Nevertheless, there are some strategies to address these concerns. One method to account for differences in the intensity of policing among different areas is to use an instrument such as traffic stops or misdemeanor arrest rates. Cohen successfully used such a strategy in a study of city-level differences in deterrence. Since gun violence often reflects the dynamics of drug markets, a proxy for drug markets in local areas—for example, drug arrests, overdoses, or seizures—can be another hedge against selection.[52] Finally, as a specificity analysis, it would be conceptually helpful to include an instrument to account for other violent crime, such as (nongun) robbery, to sort out the unique effects of the gun suppression strategy on gun crimes versus general levels of violence versus crime generally. An instrument of robbery would help referee between competing measures and conflicting results (injuries versus gunshots) and resolve general crime reduction versus gun-specific effects.

Second, the DDD model itself is not without its controversies. There are potential biases in estimating the *standard error* around the estimates of treatment effects.[53] The estimation of standard errors in DDD models opens the door to potentially serious serial correlation problems, owing to the lengthy time periods (fourteen time points or more in this study), the general serial correlation in the dependent variables, and the relatively narrow range of the treatment variable. Marianne Bertrand and colleagues used "placebo laws" to test for these effects and

50. Fagan and Davies (2000); Sampson and Raudenbush (1999).
51. Cookand Campbell (1979).
52. Sampson and Cohen (1988); MacCounand Reuter (2001).
53. Bertrand, Duflo and Mullainathan (2002).

showed a tendency toward overrejection of the null hypothesis across a range of studies using DD methods.[54] Using the third "D"—a diagnostic for preexisting effects, similar to Cohen and Ludwig—only slightly lowers the rejection rate. The problem persists, and the standard errors are underestimated. What Cohen and Ludwig could do, and I hope will in subsequent work, is to allow for a different covariance matrix that explicidy addresses the problem of repeated measurements over time. Comparison of these results with the DDD estimates in Cohen and Ludwig's chapter will provide a valuable benchmark for interpretation of DDD models and their applications to other experiments.

Finally, there are normative and constitutional questions about stops of citizens based on broad subjective notions of "suspicion" that animate police stops of citizens. There are many ways to police guns, some that promote strong citizen-police cooperation, and those that exact a high cost in citizen trust and cooperation with police.[55] The FSP was sparing and judicious in the extent of citizen contacts, compared with the directed patrol experiments in Kansas City or Indianapolis, or New York's policy of aggressive and widespread stops and frisks that most adversely affects predominantly African American communities. In cities that have adopted aggressive street-level enforcement, there have been explicit trade-offs of citizen rights and protections to obtain reductions in crime, reductions whose causal links to police efforts are contested. This ambivalence is strongest in the neighborhoods most affected by violence, the most heavily patrolled, and where social and economic disadvantage is most concentrated. Pittsburgh's recent history of police-citizen conflict suggests that the legitimacy of polking is a ractor that should be a focal consideration in the assessment of programs like FSP.[56] If small gains are produced at high costs in due process and fair (procedural) treatment, citizens might withdraw their cooperation with police in the creation of security. If aggressive patrols have high social and legal costs, in the end they will be self-limiting.

References

Bassi, Laurie J. 1984. "Estimating the Effect of Training Programs with Non-Random Selection." *Review ofEconomics andStatistics 66* (1): 36-43.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2002. *How Much Should We Trust Differences-In-Differences Estimates? Working* Paper 8841. Cambridge, Mass.: National Bureau of Economic Research.

54. Bertrand, Duflo, and Mullainathan (2002).
55. Fagan (2002); Tyler and Huo (2003).
56. Fagan and Davies (2000); Harcourt (2001); Weitzer (2000); Livingston (1999).

Blumstein, Alfred. 2000. "Disaggregating the Violence Trends." In *The Crime Drop in America,* edited by Alfred Blumstein and Joel Wallman, 13-44. Cambridge University Press.

Blumstein, Alfred, and Joel Wallman. 2000. *The Crime Drop in America.* Cambridge University Press.

Bowling, Ben. 1999. "The Rise and Fall of New York Murder." *British Journal of Criminology 39 (4): 531-54.*

Braga, Anthony, and others. 1999. "Problem-Oriented Policing in Violent Crime Places: A Randomized Controlled Experiment." *Criminology* 37 (3): 541-80.

Chermak, Steven, Edmund F. McGarrell, and Alexander Weiss. 2001. "Citizen Perceptions of Aggressive Traffic Enforcement Strategies." *Justice Quarterly* 18: 365-91.

Cook, Philip J. 1983. "The Influence of Gun Availability on Violent Crime Patterns." In *Crimeandjustice: An AnnualReview of Research,* vol. 4, edited by Michael Tonry and Norval Morris, 49-90. University of Chicago Press.

Cook, Philip J., and John H. Laub. 1998. "The Unprecedented Epidemic in Youth Violence." In *Youth Violence, Crime and Justice: A Review of Research, Volume 24,* edited by Michael Tonry and Mark H. Moore, 27-64. University of Chicago Press.

Cook, Philip J., and Jens Ludwig. 2000. *Gun Violence: The Real Costs.* Oxford University Press.

Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Houghton-Miflin.

Cox, David. 1958. (reprinted 1992). *Planning of Experiments.* Wiley.

Dunworth, Terence. 2000. *National Evaluation of the Youth Firearms Violence Initiative.* NCJ 184482. National Institute of Justice, U.S. Department of Justice.

Eck, John, and Edward Maguire. 2000. "Have Changes in Policing Reduced Violent Crime?" In *The Crime Drop in America,* edited by Alfred Blumstein and Joel Wallman, 207-65. Cambridge University Press.

Fagan, Jeffrey. 2002. "Policing Guns and Youth Violence." *Future of Children* 12 (2): 133-52.

Fagan, Jeffrey, and Garth Davies. 2000. "Street Stops and Broken Windows: Terry, Race, and Disorder in New York City." *Fordham Urban Law Journal* 28 (2): 457-504.

Fagan, J., F. E. Zimring, and J. Kim. 1998. "Declining Homicide in New York: A Tale of TwoTrends." *Journal ofCriminal Law and Criminology* 88 (4): 1277-1324.

Geller, William, and Michael Scott. 1992. *Deadly Force: What We Know.* Washington: Police Executive Research Forum.

Glaeser, Edward L, and Jacob L. Vigdor. 2001. "Racial Segregation in the 2000 Census: Promising News." Brookings Institution Center on Urban and Metropolitan Policy.

Greene, Judith. 1999. "Zero Tolerance: A Case Study of Police Policies and Practices in New York City." *Crime and Delinquency 45* (2): 171-87.

Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *American Economic Review* 84 (3): 622-41.

Harcourt, Bernard. 2001. *Illusion of Order: the False Promise of Broken Windows Policing.* Harvard University Press.

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." *Journal oj'the American Statistical Association* 84 (408): 862-80.

Heymann, Philip. 2000. "The New Policing." *Fordham Urban Law Journal2H* (2): 407-56.

Joyce, Ted, and Robert Kaestner. 1996. "The Effect of Expansions in Medicaid Income Eligibility on Abortion." *Demography* 33 (2): 181 -92.

Kelling, George, and William Souza, Jr. 2001. "Do Police Matter? An Analysis of the Impact of New York City's Police Reforms." New York: The Manhattan Institute (December).

Kennedy, David M. 1997. "Guns and Violence: Pulling Levers: Chronic Offenders, High-Crime Settings, and a Theory of Prevention." *Valparaiso Law Review* 31 (2): 449-80.

Kennedy, David, Anthony A. Braga, Anne M. Pichl, and Elin J. Waring. 2001. *Reducing Gun Violence: The Boston Gun Projects Operation Ceasefire.* NCJ 18874. National Institute of Justice, U.S. Department of Justice.

LaFree, Gary. 1998. *Losing Legitimacy: Street Crime and the Decline of Social Institutions in America.* Westview Press.

Levitt, Steven D. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *American Economic Review* 87 (3): 270-90.

―――. 2001. "Deterrence." In *Crime: Public Policies for Crime Control,* edited by James Q. Wilson and Joan Petersilia, 435-50. Oakland, Calif.: Institute for Contemporary Studies.

Livingston, Dcbra. 1999. "Police Reform and the Department of Justice: An Essay on Accountability." *Buffalo Criminal Law Review 2* (2): 815-57.

Long-Onnen, Jamie Rene. 2000. "Measures of Lethality and Intent in The Geographic Concentration of Gun Homicides: An Exploratory Analysis." Ph.D. dissertation, University of Maryland at College Park,

Ludwig, Jens. 1998. "Concealed Gun-Carrying Laws and Violent Crime: Evidence from State Panel Data." *International Review of "Law and Economics* 18: 239-54.

Ludwig, Jens, and Philip J. Cook. 2001. "The Benefits of Reducing Gun Violence: Evidence from Contingent-Valuation SurveyData." *Journalof Risk and Uncertainty* 22 (3): 207-26.

MacCoun, Robert, and Peter Reuter. 2001. *Drug War Heresies.* Cambridge University Press.

McGarrell, Edmund F., Steven Chermak, and Alexander Weiss. 1999. Reducing Firearms Violence through Directed Police Patrol: Final Report on the Evaluation of the Indianapolis Police Department's Directed Patrol Project. Report submitted to the National Institute of Justice, U.S. Department of Justice.

McGarrell, Edmund F., Steven Chermak, Alexander Weiss, and Jeremy Wilson. 2001. "Reducing Firearms Violence through Directed Police Patrol." *Criminology and Public Policy* 1 (1): 119-48.

Nagin, Daniel. 1998. "Criminal Deterrence Research: A Review of the Evidence and a Research Agenda for die Outset of die 21st Century." In *Crime and Justice: A Review of Research, Volume 23,* edited by Michael Tonry, 1—42. University of Chicago Press.

Sampson, Robert J., and Jacqueline Cohen. 1988. "Deterrent Effects of the Police on Crime: A Replication and Theoretical Extension." *Law and Society Review* 22 (1): 163—89.

Sampson, Robert J., and Stephen W. Raudenbush. 1999. "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods." *American Journal of Sociology* 105 (3): 603-51.

Shaw, James W. 1995. "Community Policing against Guns: Public Opinion of the Kansas City Gun Experiment. *"Justice Quarterly* 12 (4): 695-710.

Sherman, Lawrence W. 2000. "Gun Carrying and Homicide Prevention." *Journal of the American Medical Association 2Kb* (9): 1193-95.

―――.2001. "Reducing Gun Violence: What Works, What Doesn't, What's Promising." *Criminal Justice* 1 (1): 11-25.

Sherman, Lawrence W., and Dennis P. Rogan. 1995. "Effects of Handgun Seizures on Gun Violence: 'Hot Spots' Patrol in Kansas City." *Justice Quarterly* 12 (4): 673-93.

Sherman, Lawrence W., James W. Shaw, and Dennis P. Rogan. 1995. *The Kansas City Gun Experiment.* NCJ 150855. National Institute of Justice. U.S. Department of Justice.

Silverman, Hi. 1999. *The NYPD Battles Crime.* Northeastern University Press.

Skogan Wesley, and Susan M. Hartnett. 1997. *Community Policing, Chicago Style.* Westview Press.

Smith, Jeffrey, and Petra Todd. Forthcoming. "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics.*

Tyler, Tom R., and Yuen J. Huo. 2003. *Trust and the Rule of Law.* New York: Russell Sage Foundation Press.

Villaveces, Andres, Peter Cummings, Victoria E. Espitia, Thomas D. Koepsell, Barbara McKnight, and Arthur L. Kellerman. 2000. "Effect of a Ban on Carrying Firearms on Homicide Rates in 2 Colombian Cities." *Journal of the American Medical Association* 283: 1205-09.

Weitzer, Ronald. 2000. "Racialized Policing: Residents' Perceptions in Three Neighborhoods." *Law and Society Review* 34 (1): 129-55-