POLICY ESSAY

COMMUNITY-DRIVEN VIOLENCE REDUCTION PROGRAMS

Too big to fail

The science and politics of violence prevention

Andrew V. Papachristos
Harvard University
The University of Massachusetts—Amherst

Criminologists and politicians walk to the beat of different drummers. The 4-year rhythm of political terms and the 24-hour buzz of the postmodern news cycle disrupt the slow and steady cadence of academic research. Criminologists strive for analytic rigor, sound research design, and objectivity, especially when trying to understand causal effects such as those demanded in most evaluation research. Politicians, however, pledge their allegiance not to the scientific method but to their constituents. Problems need fixing, lives need saving, and most nonacademics need solutions at a pace quicker than the processes of peer review. As a result, whereas the criminologist waits to make claims about causality and program efficacy until field experiments and mathematical models are complete, politicians and other denizens of the "real world" often rely on back-of-the-envelope calculations or simple cross-tabulations made on spreadsheets to discern whether a violence prevention program "works." Linking specific programs to decreases in crime becomes more of an art than a science.

Such divergent worldviews between science and politics have profound implications on violence prevention efforts. "Successful" violence prevention programs typically can secure better funding and resources, not to mention the attention of community leaders, politicians, and the press. Yet who determines whether a violence prevention effort is a success? The academic with his or her regression tables and field experiments? Program administrators and front-line workers with their on-the-ground knowledge and experience? Or the politicians addressing the problems of constituents? And what impact does a programmatic "failure" have on subsequent prevention efforts?

I would like to thank Christopher Wildeman and Anthony Braga for comments on earlier drafts of this essay.
Direct correspondence to Andrew V. Papachristos, Department of Sociology, University of Massachusetts, Amherst. (e-mail: andrewp@soc.umass.edu).

DOI:10.1111/j.1745-9133.2011.00774.x

I maintain that in the absence of consistent evidence and, more importantly, acknowledgment of *scientific* evidence, politics and political rhetoric will lead the charge in determining programmatic success. Many programs—such as the Chicago Project for Violence Prevention (hereafter, simply Chicago CeaseFire)—are often anointed as "successful" before any rigorous scientific evaluation has been commissioned, let alone completed.[1] In the case of Chicago CeaseFire, external evaluation of the program was not completed until nearly a decade after its inception (Skogan et al., 2009). By that time, the program was already dubbed a success by program staff, the media, and politicians and had even expanded within Chicago and into other cities. Although the formal evaluation of Chicago CeaseFire was mainly positive, results from other CeaseFire-like replications in Baltimore, MD, Newark, NJ, and now Pittsburgh, PA, are less promising. Even with such modest and, at times, conflicting scientific evidence, CeaseFire is still packaged as a blueprint for national and even international violence prevention efforts.

The advancement of programmatic and political agendas in the face of limited scientific evidence fundamentally alters the ecology of the violence prevention world. As the title of this essay suggests, I argue that promoting programs as "models" without sufficient evidence can create a situation where violence prevention initiatives become "too big to fail." Massive amounts of political and economic capital have been diverted to programs such as Chicago CeaseFire, perhaps at the expense of smaller and equally successful programs. When such programs work—and at least some evidence suggests that CeaseFire has positive effects—they advance both science and practice, and hopefully, they are associated with significant reductions in violence. But, when they fail—or rather when negative results may develop—no one pays attention. More often than not, negative or null results of juggernaut programs like Chicago CeaseFire do not do what they should do: raise even a modest alarm about the direction of violence prevention efforts. Because programs such as these have been preordained to succeed, failure would spell the end of political and programmatic careers, not to mention it would represent a waste of taxpayer and foundation dollars. More than that, the failure of such programs would leave a massive void in our approach to violence prevention, in large part because we have put all of our eggs into one basket. When programs are too big to fail, good alternatives are simply too difficult to locate or else too obscure to sell to political audiences—without the mantel of "success," people will be less likely to invest. Programs such as these must succeed. And they do, regardless of what good science has to say about it.

The evaluation of Pittsburgh's One Vision One Life program by Wilson and Chermak (2011, this issue) provides an excellent example of a thoughtful scientific analysis of a

1. For the sake of disclosure, I have worked on the evaluations of both Project Safe Neighborhoods (PSN) and CeaseFire in Chicago (see Papachristos, Meares and Fagan 2007; Skogan, Hartnett, Bump and Dubois, 2009).

program with massive political momentum. The One Vision program is partially modeled after Chicago CeaseFire (Skogan et al., 2009). Strikingly, Wilson and Chermak's findings run *counter* to the Chicago CeaseFire evaluation as well as to the political assessments of CeaseFire-like programs more generally: Not only do the authors fail to find evidence of a violence reduction effect, but also their analysis finds that One Vision actually is associated with *increases* in violence. Wilson and Chermak provide several important discussion points about how their findings build on and call into question CeaseFire-style programs, all with an eye toward improving future strategies. Sadly, although the analysis is thorough and convincing, I suspect that it will not even cause a hiccup in the political machinery promoting CeaseFire-like programs.

In the remainder of this essay, I extrapolate a bit more on the science and politics of Chicago CeaseFire. Then, I focus on the Chicago evaluation results as well as on the results from replications in Baltimore, Newark, and Pittsburgh. I conclude by discussing the implications such a debate has for the idea of "evidence-based practice."

The Science and Politics of Chicago CeaseFire

Wilson and Chermak (2011) analyze a replication of one of the biggest and most politically vibrant violence programs today—a version of the Chicago CeaseFire program.[2] At its core, the Chicago CeaseFire initiative is based on an old approach commonly referred to as "street work": the use of outreach workers to work directly with gangs, gang members, and troubled youth to provide direct services and mediate disputes before they become violent.[3] Like previous street work efforts, CeaseFire relies not necessarily on professionally trained social workers but on street-oriented individuals (in particular, ex-gang members and ex-offenders) who have local knowledge of the neighborhood and gangs targeted for intervention. To be sure, CeaseFire has made significant advancements in the street work approach. First, CeaseFire reframed violence prevention in a public health framework; specifically, the program believes that changing attitudes toward gun violence requires changing norms and behaviors in the same way other public health efforts have tried to alter behaviors like cigarette smoking, drunk driving, and risky sexual activity.[4] Second, CeaseFire has expanded previous street work models by differentiating more fully the roles and functions of its staff. For example, the current CeaseFire model incorporates "violence interrupters," whose job it is to mediate gang/neighborhood disputes, and more traditional

2. It is important to note that the Chicago CeaseFire program is operationally and organizational distinct from the Operation Ceasefire that was part of the Boston "pulling levers" strategies, although several significant similarities exist, including the use of street workers.

3. The origins of street work programs such as these can be found as early as World War II when sociologists Clifford Shaw and Henry McKay initiated the Chicago Area Project. See Tita and Papachristos (2010) for a recent review of street work programs (also see Klein, 1971; Spergel, 1968).

4. The "crime as a public health problem" is also a much older idea that gets revisited in the CeaseFire model (see Hemenway, 2006, for a review).

caseworkers, who are engaged in direct service provision. Third, CeaseFire's street work efforts are integrated into larger community-level activism such as media campaigns, rallies, protests, town hall meetings, and so on. In a sense, CeaseFire has become a *social movement*, not simply an intervention program.

CeaseFire started in 1999 in several of Chicago's highest crime neighborhoods. Since its inception, researchers on the CeaseFire staff have monitored neighborhood-level crime indices, boasting of dramatic programmatic effects as early as the first year after implementation (CeaseFire, 2011). CeaseFire expanded rapidly based, in part, on such internal assessments and, in part, on political posturing and rhetoric: By 2005, the program expanded into approximately 12 operational sites in Chicago (Skogan et al., 2009: 2–19). The program became so politically popular that in June 2005, First Lady Laura Bush visited with street workers and lauded CeaseFire as a model of how to help at-risk youth. CeaseFire's activities and success has been spotlighted in dozens of media outlets—including an article by Alex Kotlowitz (2008) in the *New York Times Magazine*—and recently, it was the subject of a documentary movie (also by Kotlowitz) that premiered at the 2011 Sundance Film Festival.

External evaluation of CeaseFire did not commence until 2005 and was not complete until 2008—nearly a decade after the program began. Like many other programs before it, the evaluation of CeaseFire and the fact that the program's initial design did not include any formal experimental design components forced the National Institute of Justice (NIJ)–funded research team to try and reconstruct "comparison" neighborhoods to match with "treatment" neighborhoods by statistically matching on similar sociodemographic and crime characteristics.[5] Therefore, any true experimental design was stymied by the politics of funding, state legislation, and the nonrandom assignment of treatment clients and gangs—a common casualty, but not necessarily a fatal one, in the evaluation of criminal justice and violence prevention programs (Skogan et al., 2009: 2–20). The results of this external evaluation were by and large *positive* and have since been published in a NIJ report (Skogan et al., 2009). However, to the best of my knowledge, the Chicago CeaseFire findings have not been subject to peer-review publication (outside of the NIJ report) nor have they been replicated successfully.[6]

5. The lack of research design and external evaluation in the original CeaseFire model is surprising given the founder's medical and public health background as well as the housing of the program in a school of public health—a field that stresses the need for randomized trials and experiments in determining program efficacy.

6. An early analysis of PSN in Chicago conducted by my colleagues and myself (Papachristos et al., 2007) included a brief analysis of the presence of CeaseFire in several neighborhoods; using a quasi-experimental design and propensity score matching to create statistical counterfactual groups, we found no discernable effects of CeaseFire above and beyond PSN. Although our analysis was *not* intended to evaluate CeaseFire specifically, it does suggest that perhaps some findings of the NIJ report might be sensitive to methodology, timing, and the selection of comparison groups.

Is CeaseFire a "Success"?

Skogan et al.'s (2009) evaluation of Chicago CeaseFire represents a detailed and massive evaluation of both outcomes and outputs. In fact, one of the greatest contributions of this evaluation is the documentation and assessment of CeaseFire activities, clients, and logic models. The results from a longitudinal analysis of crime trends in CeaseFire neighborhoods and comparison areas find significant drops in shootings with decreases ranging from 16% to 34%. Furthermore, Skogan et al. (2009) also (a) considered the concurrent occurrence of the PSN program that was present in some (but not all) of the CeaseFire neighborhoods, (b) discussed the limitations of the models and data, and (c) described the difficulties of politics and funding in determining the selection of treatment neighborhoods. Overall, Skogan et al. worked with the available data, and their evaluation represents a good faith effort at assessing programmatic effects.

The media (not the researchers) called these Chicago results "beyond dispute" ("Seeking safe passage," 2009), thus providing the impetus for expansion and replication. However, a closer examination of the 238-page report and 212-page technical appendix may perhaps leave some room to question the political designation of these findings as "indisputable." An interesting aspect of the report by Skogan et al. (2009) is the inclusion of two independent analyses by Richard Block (technical appendix, pp. B1–B34), who analyzed the changes in shooting "hot-spots" in CeaseFire and treatment districts, and myself, who analyzed changes in network patterns of gang homicide (Skogan et al., 2009: technical appendix, pp. C1–C35). Both of these efforts were conducted independently from the main evaluation, and each of us used our own data. In his analysis, Block reviewed seven CeaseFire program neighborhoods and found positive results in three CeaseFire areas, one "probably" positive result, and three inconclusive or null results. Strikingly, my findings mirrored Block's almost exactly. Of the eight CeaseFire locations I analyzed, I found positive programmatic effects in two target areas (which were identical to two of Block's positive CeaseFire areas) and potentially a third positive result. The remaining five areas exhibited either no effect or the results were inconclusive.

Considering these full results, one might conclude more accurately that the Chicago program was successful in some of the program's targeted areas. In fact, such an interpretation would be more consistent with the findings of other CeaseFire replications, including those in Baltimore, Newark, and Pittsburgh.

In 2005—even before the NIJ-funded evaluation was underway on Chicago CeaseFire—a replication was already beginning in Newark (see Boyle, Lanterman, Pascarella and Cheng, 2010). Whereas the Chicago evaluation of CeaseFire's influence on violence relied on data provided mainly by the police department, the Newark study analyzed changes in gunshot wounds reported at trauma centers. The analyses reported by Boyle et al. found no evidence of a statistically significant decrease in gunshot injuries in the CeaseFire areas compared with similar neighborhoods.

With the assistance of a $1.6 million grant from the U.S. Department of Justice and technical assistance from the Chicago CeaseFire team, another CeaseFire replication was implemented in Baltimore in 2007—once again, before the Chicago external evaluation was complete (Webster, Vernick and Mendel, 2009). The results from the Baltimore CeaseFire model also produce mixed findings, but the overall findings lend little support of a convincing programmatic effect. On the positive side, one intervention neighborhood experienced a significant decrease in homicides postintervention. However, this finding is somewhat offset by the fact that two target neighborhoods experienced *increases* in nonfatal gunshot injuries (Webster et al., 2009, especially Tables 6 and 7).[7]

Taken in this context, the results presented by Wilson and Chermak (2011) are perhaps not surprising either analytically or politically. Analytically, the results coincide with those of Newark and Baltimore but diverge from the way the Chicago study has been portrayed. Politically, Wilson and Chermak are taking on a behemoth that was self-christened as a success and validated by politicians and the media well before the positive results of external evaluators. More than that, the Pittsburgh study illustrates how, once again, cities were willing to partake in a program that was deemed a "success" without really considering the scientific evidence.

Is CeaseFire Too Big To Fail? The Implications for Evidence-based Practice

Thus far we have observed that (a) Chicago CeaseFire was anointed a success well before external evaluation; (b) the NIJ evaluation provides some support for programmatic success, but a more detailed examination of the technical reports might suggest that such a positive assessment is not infallible; (c) subsequent replications of the CeaseFire model have by and large produced negative results; and (d) cities continue to look toward CeaseFire despite the aforementioned results.

The continued proliferation of CeaseFire-like programs in the face of largely negative replication results may seem scientifically odd. How many other public health initiatives or clinical trials would continue to receive funding in the face of similar results? Politically, however, this is an indication of the "too big to fail" phenomena: Programs draw a lot of attention and resources without consistent empirical evidence, and consequentially, stakeholders have a lot riding on their perceived success. CeaseFire programs like the ones in Baltimore, Newark, and Pittsburgh invested in a "successful" model and, therefore, potentially had much more to lose in the face of failure than if they had invested in a smaller program without the mantel of success. Negative results are dismissed all together or else flatly ignored by the media and, no doubt, by program administrators. When the First Lady and Alex Kotlowitz say a program is a success, who is a criminologist to say otherwise?

7. This interim report showed some significant changes in attitudes toward gun violence, although these did not seem to translate into changes in gun use or aggregate patterns of gun violence in intervention areas.

To be sure, public safety cannot always wait for academic findings, and often, practitioners will accept "best guesses" as evidence of programmatic success. If held to their word, politicians must be better equipped to engage and respect scientific findings. For our part, we criminologists must be able to work with practitioners and politicians in an intelligible and jargon-free manner outside of our thought experiments and counterfactuals (see Skogan, 2010). Regardless, as scientific evidence begins to mount, a reassessment of our violence prevention strategies is clearly in order. In the case of CeaseFire, the evidence demonstrates some positive results in Chicago followed by replications with null effects.[8]

How *should* we interpret these results? And, what affect should they have on subsequent proliferation of the CeaseFire model? "Too big to fail" would suggest that we cannot do anything—that stopping the momentum of massive programs would derail violence prevention efforts and leave a void in programmatic efforts. A more balanced approach suggests that both practitioners and academics would reassess our efforts to improve the focus and results of violence prevention strategies.

The good news is that the violence prevention world is perhaps ready to make serious progress on both scientific and political fronts. Politics, especially since the Obama administration, has experienced considerable advancements in the application of "evidence-based practices"—a preferential use of programs that are demonstrated to have positive effects on reducing violence (Robinson, 2010). The mounting evidence on CeaseFire-like programs provides a moment for us to understand *how* the CeaseFire model holds under the criteria of evidence-based practice. For instance, compare the idea and use of the idea of "evidence-based practice" as applied to Chicago CeaseFire with the evidence-based ethos employed by place-based policing strategies (Braga, Papachristos and Hureau, 2011). Place-based policing strategies have been and continue to be subject to rigorous evaluation. More importantly, the results of such evaluation are systemized and subjected to additional scrutiny by the scholars and policing experiments. For example, a recently updated Campbell Collaboration Systematic Review and meta-analysis of 18 randomized controlled trials and quasi-experiments reported significant crime control gains associated with hot spots policing programs (Braga et al., 2011). Negative results are not discarded but are gleaned for how to proceed in future policing interventions.

If we truly wish to subject the CeaseFire model of violence prevention to the rigors of evidence-based practice, then, like the approach taken in place-based policing, we must learn from the positive *and* negative evaluations alike. One area that comes to mind for expanding the use of evidence-based practice as applied to CeaseFire-like programs is how to measure what actually is going on. To date, *all* the evaluations examined only the changes in aggregate rates of crime or violence—that is, crime either increased or decreased after

8. The Minneapolis Domestic Violence Experiment follows a similar trajectory in which initial success was followed by failures in replication. See Sherman (1992) for a summary of this policy issue.

the CeaseFire intervention. We have very little statistical assessment of *group effects* directly linking reductions in mediated conflicts to declines in violence or of *individual effects* linking the participation of CeaseFire clients to increased participation in prosocial activities and to decreased criminal activities.

A particularly important aspect in need of serious consideration is the way we evaluate street work. The press and program administrators most often attribute such positive effects of CeaseFire to the work of outreach workers or "violence interrupters" (indeed, the documentary on Chicago CeaseFire spotlights these street workers). Yet, *none* of the evaluations of CeaseFire-like programs have measured the "treatment" given by street workers—we simply have not devised a way to measure what it is street workers do, how often they do it, and how such efforts are linked to changes in violence.[9] If street work is a model of violence prevention that we should continue to pursue—and I think it *is*—then we need to understand its effects more clearly. Once we can pinpoint the effects of some of CeaseFire's moving parts, we can then potentially assess the model more effectively as well as assess which aspects of the model might be breaking down in replication. In sum, we should seize the moment to make our programs better, not simply bigger. Evidence-based practices can help us in this quest but only if we have clear expectations of what constitutes evidence and what constitutes success.

References

Boyle, Douglas J., Jennifer L. Lanterman, Joseph E. Pascarella, and Chia Cheng. 2010. *Impact of Newark's Operation CeaseFire*. Newark: Violence Institute of New Jersey, University of Medicine and Dentistry of New Jersey.

Braga, Anthony A., Andrew V. Papachristos, and David Hureau. 2011. *Effects of Hot Spots Policing on Crime* (revised systematic review report submitted to Campbell Collaboration Crime and Justice Group; manuscript available upon request from authors).

CeaseFire. 2011. Organizational website. Retrieved July 9, 2011 from ceasefirechicago.org/results.shtml.

Hemenway, David. 2006. *Private Guns, Public Health*. Ann Arbor: The University of Michigan Press.

Kennedy, David. M. 1997. Pulling levers: Chronic offenders, high-crime settings, and a theory of prevention. *Valparaiso University Law Review*, 31: 449–484.

Klein, Malcolm. 1971. *Street Gangs and Street Workers*. Englewood Cliffs, NJ: Prentice Hall.

9. For example, in my own report in the CeaseFire evaluation, I tracked changes in the exchange of homicides between gangs. However, because of the lack of data on actual interventions and mediations of street workers, I could not link street worker efforts explicitly to the actions of specific gangs or specific violent exchanges (see Skogan et al., 2009, technical Appendix, pp. C-1–C-35). That is, although I reported that retaliatory shootings decreased in a particular area or among a particular gang, I could not identify whether those were the actual gangs that were clients of outreach workers.

Kotlowitz, Alex. 2008. Blocking the transmission of violence. *New York Times Magazine*. May 4.

Papachristos, Andrew V., Tracey L. Meares, and Jeffrey A. Fagan. 2007. Attention felons: Evaluating Project Safe Neighborhoods in Chicago. *Journal of Empirical Legal Research*, 4: 223–272.

Robinson, Laurie. 2010. The federal role in promoting evidence based practice. *The Academy of Experimental Criminology Newsletter*, April.

Seeking Safe Passage: Right under our noses? 2009. *Chicago Tribune*. October 18.

Sherman, Larry. 1992. *Policing Domestic Violence*. New York: The Free Press.

Skogan, Wesley G., Susan M. Hartnett, Natalie Bump, and Jill Dubois. 2009. *Evaluation of CeaseFire-Chicago*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Skogan, Wesley. 2010. The challenge of timeliness and utility in research and evaluation. In (John Koflas, Natalie Kroovand Hipple, and Edmund McGarrell, eds.), *The New Criminal Justice: American Communities and the Changing World of Crime Control*. New York: Routledge.

Spergel, Irving A. 1968. *Street Gang Work: Theory and Practice*. Reading, MA: Addison-Wesley.

Tita, George and Andrew V. Papachristos 2010. The evolution of gang policy: Balancing intervention and suppression. In (Robert J. Chaskin, ed.), *Youth Gangs and Community Intervention: Research, Practice, and Evidence*. New York: Columbia University Press.

Webster, Daniel W., Jon S. Vernick, and Jennifer Mendel. 2009. *Interim Evaluation of Baltimore's Safe Streets Programs*. Baltimore, MD: John Hopkins Bloomberg School of Public Health, Center for the Prevention of Youth Violence.

Wilson, Jeremy M. and Steven Chermak. 2011. Community-driven violence reduction programs: Examining Pittsburgh's One Vision One Life. *Criminology & Public Policy*. This issue.

Andrew V. Papachristos is currently a Robert Wood Johnson Health and Society Scholar at Harvard University and an assistant professor of sociology at the University of Massachusetts, Amherst. His research examines neighborhood social organization, street gangs, interpersonal violence, and social networks.