

Addressing Methodological Challenges in Evaluating Costs and Benefits of Policing Practices

Lieutenant Stuart Greer¹

January 24, 2017

Introduction

Policing is a decidedly complex task. No two interactions are the same and every contact between a police officer and citizen has implications for our society at large. Although there are expanding bodies of work and ongoing research into many aspects of policing, evaluating the benefits and costs of specific strategies, interventions and technologies remains an area that is largely untouched. In an effort to address this issue, this paper will attempt to provide a practitioner's perspective to addressing some of the methodological challenges that surround conducting such analysis. This perspective comes with recognized limited understanding of existing methodological frameworks but hopes to add value by providing a working police officer's viewpoint. The context for this paper was provided by the Policing Project at NYU School of Law and surrounds the proactive enforcement efforts of the hypothetical Ridgemont Police Department's Proactive Patrol Unit.

How can you quantify the benefit of recovering an illegal gun or warrant arrest?

One of the most challenging aspects of trying to quantify something is ensuring that there is a logical, transparent and consistent method of determining value. In this case, having two vastly different crime types as examples just highlights the difficulty that will be presented when trying to assess values across a broad spectrum of crime types and policing practices. There are clearly differences between seizing a weapon and making a warrant arrest, differences which continue to compound as you add filters to each category individually. A quick example would be the difference in value of a warrant arrest made for an outstanding violent offender who has yet to be processed and a person who has a bench warrant issued for missing a court date or payment. It's clear that not all crimes are created equal therefore any attempt to quantify efforts to address those crimes must take that into account. Such an understanding has led some criminologists to suggest a move towards a crime-harm index that can apply a weighted scale to crimes that more closely measures the amount of harm inflicted upon a community relative to all other crimes (Sherman et al, 2016). It is my suggestion that a similar approach be used to quantify the enforcement efforts of our hypothetical police officers.

In a recent publication surrounding the operationalization of the Cambridge Crime-Harm Index, Professor Lawrence Sherman (2016) suggested that evaluation methods

¹ Morristown Bureau of Police

used to measure the harm from crime “cannot be too complex to understand, too changeable to provide comparisons over time, nor too expensive to be widely used.” In order to achieve such a goal, the Cambridge model utilized the maximum penalties associated with a first-time offender convicted of the crime to assess the harm values. Such an approach could be used in our example, using the penalties for a given crime type and then assigning a monetary value to that number. In the case of the recovered weapon for example, the maximum penalty could be used as the baseline with the value being assigned based upon the judicial costs associated with the process and subsequent prison sentence.

A second approach to quantifying the benefit of a proactive arrest or seizure would be to evaluate that action against the possible harm that it could have caused a victim in terms of lost wages, medical bills, psychological counseling or other issues. The value of those categories could be assessed, either directly using available cost data or through studies of average recovery times following certain crime types. That data could then be used to generate a benefit value based upon the recovery of the weapon or the removal of an offender who had a history of committing certain types of crimes.

How can you assess whether proactive policing has any effect on crime rates without conducting a randomized-control trial?

While recognized by many as the gold standard amongst research designs, the dynamic nature of policing tends to pose operational and ethical challenges to conducting randomized-control trials. While the RCT has long been recognized for its ability to more accurately detect predicted effects, the time and money required to conduct such analyses pose challenges to the rapidly changing world of today’s police commanders. While there are now some more compressed RCTs being conducted in criminal justice studies in order to provide information on a timely basis for decision-making (BetaGov), those remain outside the scope of this analysis and may not be available to all agencies.

The first challenge associated with measuring effects from proactive policing would be to recognize the difficulty in defining exactly what was meant by that term. In one jurisdiction, such as our hypothetical example, proactive policing might entail being dropped off in a high-crime area with a wide remit and little direction. Such an example would be dramatically different from a more targeted, data-driven approach where clear parameters were used to identify the area to be addressed, the people responsible for driving crime and the tactics or strategies that would be deployed to address those issues. This is an area that would require clear definition and design if any assessment was to be possible.

Once this definition issue was addressed, I would suggest that it is possible to measure whether or not proactive policing is having an effect on crime rates however, the lack of a randomized control group would certainly result in less strength for any perceived results. Recognizing those limitations however, I would suggest that there is strength in numbers. By creating a standardized method for identification, targeting and tactics,

proactive policing units could be deployed across multiple jurisdictions or areas within a city and the crime rates could be monitored. The lack of rigor that would result from not using randomization or control groups could be offset by having access to large datasets as long as that level of standardization was maintained. As much as possible, all other variables would have to be controlled but by monitoring activity in those areas where the proactive policing was being deployed, comparisons could be drawn between the other treatment areas as well as those not assigned to the patrol. This would include the number of proactive stops, searches, seizures and arrests but should go further to assess less concrete data. In an expansion of the harm-focused concept, Dr. Jerry Ratcliffe (2014) has suggested that the perception of crime is critical to developing effective strategies as there is frequently a disconnect between the priorities of policing commanders and those in the neighborhood they police. As a result, accurately accounting for and measuring those perceptions of crimes would be an important second metric used to assess the value of the proactive policing efforts.

How would you monetize the cost of a consent search?

A consent search is something by definition which is an allowed behavior. Typically, it requires that the consenting party make an informed decision to waive a right, free of duress and with an understanding that the consent can be withdrawn. As such, it is important that a distinction be made between a consensual search and one which is more intrusive, based upon the standards set forth in *Terry v. Ohio* (1968). Such distinction must be reflected in the quantification of an encounter, even though some of the data used to evaluate those stops would be similar.

In order to assess a cost to a consent search, we would first need to measure the length of time required for the search to take place. Research from multiple disciplines has shown that people tend to apply greater weight to the time and ease of interactions than to the reason or significance of that encounter (Mazzerole et al, 2015) so a sliding scale could be utilized to apply more weight as the length of the stop increased.

The second metric that would be required to monetize a consent search would be the hit rate, or the rate at which contraband items were discovered. In terms of cost effectiveness, a consent search may be considered a valuable tactic but if there is a consistently low rate of seizures based upon those searches, it would be possible that more harm was being conducted than good. A low hit rate could also have ramifications for the manner in which officers were being trained, evaluated or deployed and would suggest that such stops were a poor use of limited resources.

I would suggest that a move to evaluate the effectiveness of consent searches could draw some parallels from the Australian move to random breath testing as a way to combat drunk driving. In that case, the length required for a stop was fairly insignificant, less than 2 minutes in most cases, and as a result, the tactic was met with limited resistance. Studies show that the hit rate, or the level at which drunk drivers were detected, varied by region and were highly dependent upon the manner in which the random testing was deployed. I would suggest that similar discrepancies would be found with any attempt to monetize consent searches.

One final metric that could be considered as part of the process of quantifying these stops would be to include the method used to quantify the recovered weapon or warrant arrest as part of the overall evaluation of the benefit of the search. If searches were conducted within a reasonable amount of time, contraband was located with relatively high levels of success and the items seized were determined to be of high value with respect to crime prevention or apprehension, those factors would be combined to determine a quantification score. The challenge again would be to clearly define and standardize what was required for a search to be considered consensual and what tactics would be allowed during such a search.

How would you monetize the dignitary cost of a stop to somebody who believes that they have been unfairly targeted because of race?

The social contract theory suggests that for a citizenry to willingly concede the right to be stopped and seized to government officials, they must believe that those extraordinary powers will be used objectively, fairly and transparently. Any action that a citizen feels supersedes the bounds of that contract will result in decreased perceptions of the legitimacy of those government actors. To assess a value to the perception that one has been targeted improperly, my immediate thought was to evaluate the length of the stop and then attempt to add a quantifying metric based upon that duration. While this would have the benefit of providing some standard for measurement, there would be a concern that the time of the stop alone would not address the feelings of the parties that were detained. In order to accurately capture that information, a measurement instrument would have to be developed that could be administered to the detained parties to determine the level of dignitary loss they perceived as a result of the stop. While recognizing that this would be an imperfect instrument, it should allow for factors such as the circumstances of the stop, the demeanor of the investigating officers and the level of procedural fairness administered to influence the final score. By combining the values based upon the length of time and perception of the detained party, a standardized value could be assessed for each and ranked according to time and score to develop a final dignity-loss value.

Should arrest and incarceration costs be included as part of the cost-benefit analysis?

Policing decisions should not be viewed in isolation. Every action that is taken by a police officer will have a ripple effect within the interconnected series of systems that surround this profession. Only by assessing the complete value or cost for each decision can a society make an informed decision as to what is acceptable. The concept of “lawful but awful” policing suggests that simply because an action is allowed under existing laws, there must always be an analyst to determine whether or not it is still morally or ethically acceptable to the public we serve (Wogan, 2016). Tragic events such as the encounter that lead to the death of Eric Garner in New York remind us that this is much more than a conceptual discussion and highlight the importance of determining what should be considered when setting police tactics

and priorities. If one does not consider the full costs of an action, it is impossible to properly frame that action within the overall context of the larger system.

A simple example of this concept was provided by Malcolm Gladwell in his story about million-dollar Murray (2006). This narrative followed the actions of a single homeless alcoholic in Reno, Nevada and outlined the costs associated with providing him emergency care as a result of his inebriation. In the example, Murray was contacted several times a day by police officers, was transported to local hospitals by emergency medical services and was treated within an emergency department setting. Each one of those actions took resources away from other priorities and had a very real economic impact on the systems at play. By considering the total costs of contact, transport and treatment, rather than simply using a more narrowly focused policing response view, the actual cost of dealing with this individual was calculated and it became clear that a different approach was warranted based upon a value proposition. I would suggest that a similar approach should be taken in policing. If the combined costs of arrest, processing, transport and court were considered when establishing response protocols to complaints of people selling loose cigarettes on the street, it is doubtful that the fateful series of events leading to the death of Mr. Garner would have taken place. I would caution that this should not be a policing decision alone but would suggest that having a more robust measurement of the actual cost of dealing with varied crimes and strategy responses would allow for a citizenry to make more informed decisions about what would be tolerated by their police agencies.

How would you measure, and quantify, the effect that proactive policing has on community trust?

To measure public trust and confidence in policing, you would have to conduct frequent surveys with representative samples of people within the varied communities that were served. As proactive policing would typically be deployed in high-crime neighborhoods or areas with emerging problems, it would be essential that the people in those areas were able to voice their opinions. The measurement of trust and confidence alone would be challenging but in order to quantify the effect that proactive strategies had on that metric, the survey would have to be designed in such a way that the tactics or methods being used by those units were identified and included as questions on the instrument. Specifically measuring components of the proactive strategies would allow police commanders to determine what tactical or procedural aspects were associated with increasing levels of trust amongst those they served. As increasing perceptions of trust have been strongly linked to improved relations and cooperation with police (Tyler, 1988) this is a metric that could have very real implications for the design and modification of policing strategies. If police commanders could identify those tactics that increased feelings of trust amongst the community and had been proven effective through analysis of hit rates and crime statistics, the benefits could be seen through lower use of force rates, lower injury levels to police, higher rates of solving crimes from increased citizen cooperation and lower overall crimes rates in the community.

References

BetaGov is an organization dedicated to promoting innovation in the public sector. Access and information available at betagov.org

Gladwell, M. (2006). Million-dollar Murray. *The New Yorker*, 13, 96.

Heaton, R., & Tong, S. (2015). Evidence-based policing: from effectiveness to cost-effectiveness. *Policing*

Mazerolle, L., Bates, L., Bennett, S., White, G., Ferris, J., & Antrobus, E. (2015). Optimising the length of random breath tests: Results from the Queensland Community Engagement Trial. *Australian & New Zealand Journal of Criminology*, 48(2), 256-276.

Ratcliffe, J. H. (2014). Towards an index for harm-focused policing. *Policing*

Sherman, L., Neyroud, P. W., & Neyroud, E. (2016). The Cambridge Crime Harm Index: measuring total harm from crime based on sentencing guidelines. *Policing*, 10(3), 171- 183.

Terry v. Ohio, 392 US 1 (1968)

Tyler, T. R. (1988). What is procedural justice?: Criteria used by citizens to assess the fairness of legal procedures. *Law and Society Review*, 103-135.

Wogan, J.B. (2016). How Police Chiefs Plan to Avoid ‘Lawful but Awful’ Shootings. *Governing*