The Contagiousness of Police Violence

Thibaut Horel, Trevor Campbell, Lorenzo Masoero,
Raj Agrawal, Andrew Papachristos and Daria Roithmayr[*]

Explanations for unlawful police violence focus on individual "bad apple" officers or deviant top-down departmental culture. Recent research suggests that violence may diffuse through social networks, much like a disease spreads on networks of interaction. We investigate whether police violence is contagious—whether an officer's exposure to earlier police shootings by network neighbors increases the probability that the officer will engage in future police shootings. Drawing on data from Chicago, we construct and analyze dynamic patterns of diffusion of shooting on police professional networks. We find structural and dynamic evidence consistent with a dynamic of contagion in police-involved shootings, even after controlling for homophily. Fitting a dynamic model of contagion to police network data from Chicago, we find that a single shooting event at the beginning of the study period gives rise to 0.5 additional offspring shootings during the eight year period under study. Contagion appears to affect a significant number of shootings—it contributes to the occurrence of 141 of 488 (29%) police-involved shootings in our study. Most remarkably, within two years, exposure to a single shooting more than doubles a network neighbor's probability of a future shooting. Our findings suggest that interrupting the transmission of violence in officer networks may well be an important avenue to reduce police shootings.

Introduction

The October 2018 verdict against police officer Jason Van Dyke for killing Laquan McDonald is the first time in the last fifty years that a Chicago officer has been convicted of murder in an on-duty shooting. Nationwide, police officers shoot between 900 and 1000 civilians every year. Police use fatal force most often against civilians armed with knives or guns, but unarmed black men constitute a significant number of civilian deaths—36 in 2015 alone. Mental health also plays a significant role in fatal force: one in four civilians shot by the police are reported to be experiencing some form of mental health distress at the time of their encounter with police.[1]

---

[1] Fatal Force, Washington Post dataset for 2017 (published January 2018, https://www.washingtonpost.com/graphics/national/police-shootings-2017/) (last accessed October 4, 2018).

Police shootings that involve excessive force do great harm. They erode the legitimacy of law enforcement and create a relationship of profound mistrust between officer and civilian.[2] Like gunshot violence for civilians, police shootings tend to cluster in socially and economically struggling neighborhoods in communities of color. Social movements like Black Lives Matter have organized around the issue of police brutality, and have worked to make visible the distinctly racial profile of police shootings. Several high-profile killings of civilians have spurred nationwide protests and sparked debate over police use of force.

Data from Chicago provide a useful up-close look at the structure and dynamics of police shootings. Between 2010 and 2016, Chicago police engaged in 435 shootings. Officers killed 92 people and wounded 170. Almost 80 percent of the 262 people shot by Chicago police were African-American. Latinos accounted for 35 of the shooting victims, nearly 14 percent of the total. Only fourteen of those shot were white, less than 6 percent of the total.[3]

Sorting out which of these shootings were justified and which involved excessive force is extremely difficult. Almost all police shootings in Chicago have been adjudicated by department officials and independent review authorities as justified.[4] At the same time, a recent investigation by the Department of Justice in the wake of Laquan McDonald's killing has cast doubt on such findings. The DOJ report found fault with a number of police practices that create a high risk of excessive force. Of particular concern was the practice of shooting civilians during a foot chase when they are running away from officers and do not pose a threat to officers or the public, and shooting into cars that are driving away but are not posing an imminent threat to the public.[5]

Theoretical literature on the causes of police violence falls into two categories. The first focuses on the so-called "bad apple" micro-level theory of violence, which attributes violence to the deviant traits (like the authoritarian personalities) of individual officers. On this view, violence is generated bottom up. A second macro-level category ascribes police violence to top-down police department incentives, set up by departmental leaders and enforced by way of both formal and social approval and/or punishment.

Recent research on the contagiousness of violence suggests a third possibility: that violence is also generated at the meso-level, spreading from officer to officer much like a disease. A literature on the diffusion of violence supports this possibility as an additional source of excessive force. Research on violent uprisings and gangs suggests that under certain conditions, people who are exposed to earlier

---

[2] See Tammy R. Kochel, Assessing the initial impact of the Michael Brown shooting and police and public responses to it on St. Louis County residents views about police, 4–6 (2015), http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1001&context=ccj_reports [http://perma.cc/RK99-DVLW]. See generally Anthony Bottoms & Justice Tankebe, Beyond Procedural Justice: A Dialogic Approach to Legitimacy in Criminal Justice, 102 J. Crim. L. & Criminology 119, 132 (2012); Tom R. Tyler, Enhancing Police Legitimacy, 593 Annals Am. Acad. Pol. & Soc. Sci. 84, 85 (2004).

[3] Id.

[4] See A Historic Murder Conviction, The Chicago Tribune, October 6, 2018.

[5] U.S. Department of Justice Civil Rights Division and US Attorney's Office, Northern District of Illinois, Investigation of the Chicago Police Department, January 13, 2017 (DOJ Report).

violence will engage in violence at a later time at a greater rate than those who are not exposed.[6] In addition, other literature suggests that social learning plays an important role in police department dynamics. Studies suggest that officers frequently learn from each other key scripts about how to exert control over civilians or reduce the risk of a civilian encounter.[7]

In this paper, we investigate the contagiousness of police violence in police networks, using Chicago as a case study. More specifically, we investigate whether an officer's exposure to a shooting by a neighbor on his network increases the probability that he will engage in a shooting in a subsequent encounter with a civilian. We analyze data from civilian complaints and shootings, collected by the Chicago Police Department (CPD) and Chicago's Independent Police Review Authority (IPRA) between 2008 and November 2015.

Drawing from this data, we construct networks of officers who interact with each other, to test whether violence—here, police shooting—is contagious. We generate networks by connecting officers who interact with each other; in network parlance, we draw "edges" between the "nodes" of officers who are "tied" or "connected." Here, we construct a layered network that connects police officers who engage in shooting.

We first construct a network of CPD officers who are listed together on the same civilian complaint (the "complaint" network), to map officers who interact with each other in responding to calls with civilians. We assume that being listed on the same complaint constitutes evidence of an existing relationship among the officers listed, a relationship that might mediate the transmission of violence.[8]

We then layer onto the complaint network a separate network of police-involved shootings for which the shooting officers are connected on the complaint network. Layering shooting data onto the complaint network allows us to create this separate network consisting of shooting officers who have been listed together on a civilian complaint and are statistically likely to have had an existing relationship during the shooting. We analyze this separate shootings network. In particular, we investigate the relationship between an officer's exposure to violence by his neighbors on the shootings network and his future acts of violence.

---

[6] Henk W. Houweling and Jan G. Siccama, The Epidemiology of War, 1816-1980, 29 J. Conflict Res. 641 (1985) (violent uprisings cluster and diffuse across time and space); Ben Green, Thibaut Horel and Andrew Papachristos, Modeling Contagion Through Social Networks to Explain and Predict Gunshot Violence in Chicago, 2006 to 2014, 177 JAMA Intern Med. 326 (2017) (diffusion of gunshot violence exhibits evidence of contagion).

[7] For a full review of the literature, see Daria Roithmayr, The Dynamics of Excessive Force, UNIVERSITY OF CHICAGO LEGAL FORUM: Vol. 2016 , Article 10.

[8] Empirical studies support the idea that co-offender relationships are highly correlated with existing relatively stable relationships between co-offenders. Jean Marie McGloin and Holly Nguyen, the Importance of Studying Co-Offending Networks for Criminological Theory and Policy," in PROC. THIRD ANNUAL ILLICIT NETWORKS WORKSHOP, Montreal, Quebec, October 2011; Jean Marie McGloin, Christopher J. Sullivan, Alex R. Piquero, and Sarah Bacon, Investigating the stability of co-offending and co-offenders among a sample of youthful offenders," 46 Criminology 1 (2008). See also Albert J. Reiss, Co-offending and criminal careers, in CRIME AND JUSTICE: A REVIEW OF RESEARCH (1988); Albert J. Reiss and David P. Farrington, Advancing knowledge about co-offending: Results from a prospective longitudinal survey of London males, 82 J. of Crim. Law and Criminology 23 (1991).

Using the shootings network as our subject of study, we use three approaches to investigate the possibility of contagion in police shootings. First, we investigate the structure of the shootings network, comparing the connectedness of officers who have engaged in shooting to the connectedness of randomly chosen officers on the underlying complaint network. A relatively greater number of connections among shooters suggests that these connections in the complaint network serve as potential pathways of transmission for contagious violence. We find that officers who are engaged in shooting are far more connected to other shooting officers than randomly chosen officers on the underlying network. Shooting officers have on average 20 connections, while non-shooting officers have only 12. Thus, shooters appear to have more pathways through which to transmit violence than a random officer on the network.

Second, we compare the inter-event periods between shootings by officers who are connected on the shooting events network (on the one hand) with shootings by randomly chosen officers (on the other hand.) Shorter periods between shootings among connected officers suggest that officers who are exposed to earlier shootings become more likely in a given period of time to engage in a later shooting. We find that the inter-event period elapsed between shootings by connected officers is much shorter—739 days—than the period for randomly chosen officers—994 days. This suggests that shootings are more likely in a given time period for officers who have been "exposed" to shootings on their network. There are alternative explanations for this finding, which we discuss below.

Finally, we fit a dynamic point-process model of contagion to observed data. We estimate three relevant parameters that describe the contagion relationship: (i) the background rate of shooting independent of contagion, (ii) the increase in probability of a future shooting from exposure to a prior shooting, and (iii) the rate at which this spike in probability decays over time. Interpreting those parameters, we are able to estimate that owing to social contagion, each shooting gives rise to another 0.496 "offspring" shootings during the eight-year period under study. In addition, we estimate that the number of shootings affected by contagion is 141 out of 488 shootings during the period under study, the remainder being background shootings. Most remarkably, we estimate that exposure to a single shooting doubles the probability of a future shooting within two years.

As with all network analysis, to make any claim about contagion, we must separate the effects of social contagion and homophily. The tendency of similar individuals to associate together could potentially explain some of our results. We control for homophily by using randomizing techniques that focus on the way in which contagious shootings depend on shooting time patterns and sequences. Our results suggest that neither homophily nor other omitted variables are likely to be the source of evidence consistent with contagion.

The rest of this paper is organized as follows. Part I describes the existing literature on police violence and the literature on contagiousness. Part II describes our data. Part III describes our conceptual framework for investigating contagion. Part IV describes our results. Part V discusses the results and considers avenues for further research. We then conclude.

I.   Existing literature

      a.   Causes of police violence

Existing scholarship offers three types of causes for excessive force: (i) individual police officers, so-called bad apples—are responsible for a significant fraction of police violence; (ii) the top-down culture of particular police departments, shaped by leaders and organizational incentives, encourages violence; (iii) racial and gender dynamics, which prejudice officers against black men, explain police brutality. We consider each in turn.

The majority of literature on police violence traces excessive force to so-called bad apples who possess deviant "authoritarian" personality traits or attitudes. This claim finds mixed support in the literature at best. Although some survey research correlates a permissive attitude towards police force with a likelihood of using improper force,[9] other studies have found no consistent link between attitudes or personality traits and actual office behavior.[10] Most contemporary scholars now dismiss the bad apples hypothesis as unsupported, though a recent literature continues to focus on the small number of officers responsible for a disproportionate number of complaints or shootings.

Indirect evidence also suggests that malignant organizational culture in particular police departments may contribute to high rates of excessive force. This evidence primarily comes from studies on the effect of changing organizational rules on levels of excessive force.[11] At the same time, these studies are hard pressed to link organizational structure or culture at the top with police behavior on the ground level, and to adequately distinguish between organizational and occupational culture, which describes the profession of policing more generally.

As an alternative argument from culture, mixed evidence exists for the argument that race and gender stereotypes contribute to high levels of excessive force against black male civilians. Although scholars find uncontroversial the claim that the use of deadly force occurs disproportionately against civilian men who are black or Latino, and that implicit bias contributes to the use of excessive force, ascertaining the precise cause of this disparity has proved far more difficult.[12]

---

[9] See Brad W. Smith & Malcolm D. Holmes, Community Accountability, Minority Threat and Police Brutality: An Examination of Civil Rights Criminal Complaints, 41 Criminology 1035, 1048-57 (2003) (hereinafter "Community Accountability").

[10] See Robert E. Worden, The Causes of Police Brutality: Theory and Evidence on Police Use of Force, in JUSTICE FOR ALL: UNDERSTANDING AND CONTROLLING POLICE ABUSE OF FORCE 37-46 (William A. Geller & Hans Toch eds., 1996).

[11] See James J. Fyfe, Administrative Intervention on Police Shooting Discretion: An Empirical Examination, 7 J. Crim. Just. 309, 313-22 (1979); Warren Christopher, REPORT OF THE INDEP. COMM'N ON THE L.A. POLICE DEP'T (1991) (hereinafter "Christopher Commission").

[12] See Malcolm D. Holmes & Brad W. Smith, Intergroup Dynamics of Extra-Legal Police Aggression: An Integrated Theory of Race and Place, 17 AGGRESSION AND VIOLENT BEHAVIOR. 344, 345 (2012); Robert J. Kane, The Social Ecology of Police Misconduct, 40 Criminology 867, 887-88 (2002); Michael R. Smith & Geoffrey P. Alpert, Explaining Police Bias: A Theory of Social Conditioning and Illusory Correlation, 34 Crim. Just & Behavior., 1262, 1269-74 (2007); Eric A. Stewart et al., Neighborhood Racial Context and Perceptions of Police-Based Racial Discrimination Among Black Youth, 47 Criminology 847, 855, 870-75 (2009).

b. Contagion as a supplementing potential cause

We develop here the hypothesis that contagion, in which the use of excessive force spreads from officer to officer, contributes to excessive force. The bad apples hypothesis focuses on individual traits at the micro-level, finding the source of violence at the level of individual. Relatedly, the bad culture hypothesis focuses on top-down, macro-level patterns. The contagion hypothesis suggests that violence may also emerge at the meso-level, in the social learning that takes place when police officers interact with each other.

Recent scholarship on network diffusion has described the way in which information and behavior spreads along networks.[13] Scholarship over the last several decades has established that criminal violence spreads among peers, often members of a group. For example, work in criminology and sociology documents that at the population level, criminal violence (terrorism, mass killings, gang and drug violence) displays the cluster patterns common to diffusion.[14] This paper builds in particular on methodology developed in earlier work by Green, Horel and Papachristos on the contagious dynamics of gang violence.[15]

Beyond this empirical work on population-level diffusion, a number of studies suggest that violence might spread because criminals learn particular scripts to follow during violent crimes. These scripts involve a particular sequence of actions that criminals execute much like a script in a play. So for example, criminals might follow a script that involves overwhelming a robbery victim with a sudden show of force (as opposed to an initial non-threatening engagement) to destabilize the victim, leaving the victim vulnerable to robbery. Empirical evidence suggests that these scripts cluster regionally along criminal networks, and are transmitted via social learning both in and out of prison.[16] Scholars suggest that a range of contextual facts determine the contagiousness of these scripts of violence: most importantly, social learning occurs when engaging in the script is positively reinforced, and when individuals identify with each other.[17]

---

[13] For an overview of network research applied to sociology, criminology, public health, psychology and a wide range of other disciplinary social science investigations, see Mark Newman, NETWORKS: AN INTRODUCTION (2010).

[14] See Towers S, Gomez-Lievano A, Khan M, Mubayi A, Castillo-Chavez C, Contagion in Mass Killings and School Shootings. PLoS ONE 10(7) (2015): e0117259; M.B. Short, et. al, Gang Rivalry Dynamics via Coupled Point Process Networks, 19 Discrete and Continuous Dynamical Systems Series (2014); Charles Loeffler and Seth Flaxman, Is Gun Violence Contagious? S. J Quant Criminol (2017); Gentry White, Fabrizio Ruggeri, and Michael Porter, Modelling the Spread of Terrorism via Diffusion and Contagion, arXiv:1612.02527. See also Jeffrey Fagan, Deanna Wilkinson, Garth Davies, The Social Contagion of Violence in THE CAMBRIDGE HANDBOOK OF VIOLENT BEHAVIOR 688 (Daniel Flannery et al. eds 2007).

[15] See Ben Green, Thibaut Horel and Andrew Papachristos, Modeling Contagion Through Social Networks to Explain and Predict Gunshot Violence in Chicago, 2006 to 2014, JAMA Intern Med. 2017;177(3):326–333.

[16] L. Rowell Huesmann & Lucyna Kirwil, Why Observing Violence Increases the Risk of Violent Behavior by the Observer, in THE CAMBRIDGE HANDBOOK 545, 564-65 (Daniel Flannery et al. eds., 2007).

[17] Id. For a fuller discussion of scripts and police violence, see Daria Roithmayr, The Dynamics of Excessive Force, 2016 Chicago Legal Forum 233 (2016).

Like criminals, police officers appear likely to learn scripts of violence from each other. In a study of policing of protest movements, Della Porta and Tarrow charted the contagious spread of a particular force tactic deployed by police in their battle against an increasingly organized protest movement. When confronted by innovative protestors who blockaded officers in the street, police deployed a particular script of violence: they mobilized special coercive units in advance, and then used massive force to temporarily incapacitate protestors after which they pushed protestors outside the boundaries of legal protest. Scholars traced the diffusion of this particular script of violence from unit to unit and from city to city, as officers from other cities were imported to participate in joint intervention.[18]

This literature supports the possibility that police shootings might spread through the social learning of scripts of violence, which get reinforced as officers observe that shootings reduce the violence of a potentially risk encounter.[19] Recent work in sociology documents the way in which "the danger imperative," a collective preoccupation with violence and ensuring officer safety, dominates police decision-making.[20] Officers who are highly attuned to the threat of violence from a population saturated with guns adopt a range of strategies, including excessive force, to reduce this perceived threat of violence. Officers impart information about these models and strategies to each other, both explicitly and observationally, as a way of mediating risk or danger.[21]

This project investigates whether exposure to earlier shootings increases the likelihood of later shootings by police officers. It is the first project to our knowledge to empirically investigate the contagiousness of police violence.[22] Importantly, this paper is a preliminary step, seeking only to ascertain whether there is evidence consistent with an increase in probability of later shooting from exposure to earlier shooting. Our dataset contains insufficient information to test whether police shootings spreads from officer to officer through the mechanism of social learning.

More specifically, our empirical analysis does not allow us to identify the mechanism through which police violence might spread. Other mechanisms may be at work. For example, officers may infer from exposure to a shooting that the risk of a civilian encounter is higher than they had assessed, and this perception of increased risk could increase the likelihood of a future shooting. As discussed

---

[18] See Donatella Della Porta & Sydney Tarrow, Interactive Diffusion: The Co-Evolution of Police and Protest Behavior with an Application to Transnational Contention, at 10-34 (presented at European Consortium for Political Research 2011); John D. McCarthy et al., The Diffusion and Adoption of Public Order Management Systems, in SOCIAL MOVEMENTS IN A GLOBALIZING WORLD 71, 71-93 (Donatella Della Porta et al. eds., 1999).

[19] Thirty percent of all police shootings between 2010 and 2016 involved foot chases, in which the civilian is running away from the officers. Tribune.

[20] Michael Sierra Arevalo, American Policing and the Danger Imperative (working paper).

[21] Id. at 6. See also Jerome H. Skolnick and James J. Fyfe, ABOVE THE LAW: POLICE AND THE EXCESSIVE USE OF FORCE (1994).

[22] Earlier work by one of the authors theorizes more fully the arguments presented here on the social learning of scripts among police officers.

in the conclusion, further research may help to isolate the precise mechanism to explain the contagious spread of violence among officers.

II.   Data

A. Description and Summary Statistics

In this project, we investigated contagion by analyzing the structure and dynamics of two connected networks. First, we mapped a network of police officers who are listed together on civilian complaints against police officers ("the complaint network.") This complaint network connects officers who respond to calls together and for whom evidence suggests they engage in misconduct together. Being listed on a complaint together also suggests a pre-existing relationship through which scripts of violence could be transmitted. Generating networks using co-offending as a proxy for relationships enables scholars to map patterns of diffusion for gun violence and a number of other types of violence.[23]

Second, we analyzed a smaller network of police-involved shootings, in which officers involved in the shootings are connected on the complaint network described above. This "shootings network" connects officers who respond to calls together and who engage in shootings. We focus our investigation on this shootings network, but on occasion compare this network to the larger, underlying complaint network.

We constructed the complaint network from complaints of misconduct filed by civilians against CPD police officers between January of 2008 and November of 2015 (N = 19,843). This network contained a smaller number of unique officers because many officers received more than one complaint (O = 9737).[24] In the Chicago Police Department, officers work in 25 police districts and in a number of special units, both within the district (e.g., tactical response teams) and across districts (e.g., the gang unit).[25] While individual officers move across districts over the course of their careers, officers are paired with partners and assigned to beats based entirely on their district of service; only on rare occasions, such as multi-jurisdiction cases or specialized units, do officers work directly with officers from different districts. Officers are assigned to districts at the beginning of their appointments. Assignments are not random; owing to the union agreement with the Fraternal Order of Police, Chicago officers are able to some degree to bid for vacancies based on seniority.[26]

---

[23] See Green, Horel and Papachristos, supra note x; D. L. Haynie, Friendship networks and delinquency: the relative nature of peer delinquency, 18 J Quant Criminol. 99 (2002); See generally, Michael Warr, COMPANIONS IN CRIME: THE SOCIAL ASPECTS OF CRIMINAL CONDUCT. New York, NY: Cambridge University Press; 2002.

[24] For more details on the complaint network, including an extended analysis of the distribution of officers and complaints, and the factors predicting the existence of co-offenders, see George Wood, Sinclair Ewing-Nelson, Jennifer Wu, Daria Roithmayr and Andrew Papachristos, The Structure of Police Misconduct (working paper.)

[25] See Appendix for a map of police districts and beats.

[26] Section 23.8, Agreement Between the City of Chicago Department of Police and the Fraternal Order of Police Chicago Lodge No. 7. (2012).

The data used in this dataset were extracted from a larger dataset of complaints covering a larger number of years and types of complaints, obtained by collaborative effort of the Citizens Police Data Project at the Invisible Institute and the University of Chicago Law School's Mandel Legal Aid Clinic (N = 107,276). These institutions obtained the data from the CPD pursuant to a Freedom of Information Act request and subsequent litigation, and have publicly released the data.[27]

We study civilian-facing complaints, on the assumption that officers who are connected because they interact in connection with a civilian encounter are significantly more likely to spread police violence against civilians through those connections. Complaints by civilians made up 60.3% of the complaint dataset and 39.7% of complaints were generated within the department.

Civilian complaints constitute an important though limited window into police conduct. Under a long-standing CPD policy in effect during the relevant period, civilians with grievances against police officers were given the opportunity to file administrative complaints. During the relevant period, filing a complaint required civilians to swear out an affidavit attesting to the truth of their allegations. Filing also required affidavits to be signed in person, at a limited number of locations.[28] After a filing was complete, the complaint was investigated by an independent authority or other agency. For those investigations that were completed, allegations were either sustained, un-sustained or the officer was exonerated.[29]

For the period under study, the Independent Police Review Authority (IPRA) handled initial responsibility for processing and investigating civilian complaints. IPRA was a then-existing separate oversight agency whose leadership and funding came from outside the department.[30] Less than 3% of complaints were sustained, and officers were almost never disciplined.

Over 55% of civilian complaints in our dataset name two or more officers, indicating that the majority of alleged police misconduct occurs in a group setting. Civilian complaints in our dataset

---

[27] Kalven v. Chicago, 7 N.E.3d 741 (2014). Data is available at https://invisible.institute/police-data (last accessed on October 12, 2018). For detailed information on these data, see Bocar A. Ba, Going the Extra Mile: The Cost of Complaint Filing, Accountability, and Law Enforcement Outcomes in Chicago (working paper 2016), and Kyle Rozema and Max Schanzenbach, Good Cop, Bad Cop: Using Civilian Allegations to Predict Police Misconduct, American Economic Journal, Economic Policy (2018).

[28] Under Illinois law, investigations into allegations made by civilians against the police can proceed only if there is a sworn affidavit. The law further directs that "[a]ny complaint, having been supported by a sworn affidavit, and having been found, in total or in part, to contain knowingly false material information, shall be presented to the appropriate State's Attorney for a determination of prosecution." 50 ILCS 725/3.8. For a discussion of the limitations posed by requiring physical presence and limiting the number of locations, see Bocar Ba, Going the Extra Mile: The Cost of Complaint Filing, Accountability, and Law Enforcement Outcomes in Chicago (working paper).

[29] DOJ Report, supra note x.

[30] In 2007, the Chicago City Council created IPRA to replace the CPD's Office of Professional Affairs. IPRA investigates certain types of complaints against CPD members, reviews legal settlements involving police misconduct, and makes policy recommendations designed to increase the efficiency of CPD. In November of 2015, the City of Chicago replaced IPRA with COPA, owing to findings by the City Council that IPRA followed deficient practices of oversight and data collection. DOJ Report supra note x.

included a range of types of allegations, including: First Amendment violations, wrongful arrest, illegal search and seizure and excessive force. To guard against endogeneity problems, we excluded all allegations (__%) connected in any way to the discharge of a firearm.

Officers listed in the complaints were identified using personnel records on all Chicago police officers that had been separately obtained by filing a Freedom of Information Act (FOIA) request with the City of Chicago Department of Human Resources. The bi-annual personnel data include each officer's name, position title (rank), and original hire date. These records were available from 2002 to 2014. We used the officers' badge number and the original hire date to link records with the complaint and shooting datasets, and to identify unique officers.

We constructed the shootings network from detailed incident-level data collected by IPRA on police-involved shootings occurring between 2008 and November 2015. Chicago city regulations required IPRA to investigate "all cases in which a department member discharges his or her firearm, stun gun, or Taser in a manner which potentially could strike an individual, even if no allegation of misconduct is made."[31] CPD General Orders required officers to notify IPRA each time a CPD member discharged a firearm. Once notified, IPRA created an electronic record of its investigation by assigning each incident a log number in CPD and IPRA's electronic case management system (CLEAR).[32]

City regulations required IPRA to issue quarterly reports on weapons-discharge investigations. In its reports, IPRA included five categories of weapons-discharge notifications: "hit shooting" of firearm, "non-hit shooting" of firearm, shooting of firearm at an animal, shootings with taser, and oleoresin capsicum (OC). We concentrate analysis on the first category of data consisting of police-involved "hit shootings" (meaning shootings that hit a civilian) with a firearm, from January 2008 to November 2015.[33]

Reports provided detailed incident-level information from these categories.[34] Key variables in this dataset included: the date and location of the shooting, identifying information associated with the officers who had discharged their gun, and information about the organizational assignment of the relevant officer. Individual officers involved in police shootings were linked through a unique identifying number to those officers in civilian complaints.

We excluded both same-day same-event shootings, involving connected officers who were involved in a police shooting at the same event, and same-day different event shootings. Excluding the latter (for now) biases our investigation in a conservative direction against a finding; same-day different event shootings are likely to reflect contagion, particularly in light of the short period of time between shootings.

---

[31] MCC § 2-57-040.

[32] DOJ Report, supra note x.

[33] "Miss" shootings are less reliably reported; only 18 are included in the dataset, but are excluded from our analysis.
[34] MCC § 2-57-040.

In the shootings network, nodes in the graph represent individual shootings, meaning a discharge of a firearm by an individual officer. Each node in the graph pairs an individual shooting with the date of the shooting and the identity of the shooting officer (n= 488 shootings/officers). Because a number of officers committed more than one shooting, the shootings network contained a smaller number of unique officers (o = 418 unique officers).[35]

Table 1 summarizes the complaint and shootings data:

	Complaints	Shootings
Number	N= 19,843	n=488
Unique officers	O=9737	o=418

Table 1 Complaints and Shootings Data

These data were not without significant limitations. A full 50% of allegations in our original dataset were not investigated because civilians did not complete the required affidavit.[36] Although complaints with affidavits might be more reliable than those without, in general, civilian complaints are unreliable as a source of information about the events and associated individuals reported, with both under- and over-inclusion posing a significant limitation. A large portion of potential police misconduct goes unreported by civilians, frequently owing to administrative requirements but also owing to a range of factors associated with a strong distrust of authority, a lack of faith in accountability, and a reluctance to risk retaliation.[37]

Likewise, civilians are not reliable reporters of the events in question; reports suffer from civilian biases against the police, faulty memory and other common sources of mistake or misreporting.[38] We make no claims about civilian complaints other than that the officers that are listed as subjects of the complaints interacted with each other and with civilians in the encounter at the reported place on the reported date. Given data limitations, the complaint co-offender network cannot be said to accurately represent the social connections among officers because we capture some subset but not the full set of social connections among officers. Nor can we determine whether this network is representative of other police departments outside of Chicago.

In addition, the IPRA dataset reporting police-involved shootings has several important limitations. Most notably, subsequent analysis by the City of Chicago's Office of Inspector General (OIG) found that IPRA did not follow best practices in reporting use of force. Among other difficulties, IPRA relied on CPD notifications, and did not independently verify that the Department had provided all of the required weapons-discharge notifications. At the same time, the data appears

---

[35] We discuss below the distribution of shootings in the larger shootings network.

[36] Research suggests that incomplete allegations that are not processed for lack of a complainant affidavit are as informative regarding officer future behavior as the allegations that are fully investigated. Bocar Ba, Rozema and Schanzenbach.

[37] Kim M. Lersch, "Police misconduct and malpractice: a critical analysis of citizens' complaints," 21 Int'l. J. of Police Strategies & Management 80 (2016).

[38] Id.

relatively reliable. The OIG report compared the IPRA reported shootings to those reported in CPD internal "tactical response reports." The comparison found that IPRA's shooting data from September 2007 to September 2014, which reported 488 "hit shootings," had overreported by only four shootings compared to the CPD tactical response reports, an error rate we deemed to be within an acceptable range.

Finally, the subset of officers in the complaint network are not necessarily representative of the larger departmental population. If officers in the complaint network differ from those not in the network, our results might not represent actual contagion for the department. Representativeness poses less of a concern because all 418 shooting officers are present in the complaint network, and more generally, 70-75% of officers appear at least once in our complaint network.[39] Thus, the analysis captures the full population of shooters and a significant fraction of non-shooters.

More importantly, any selection towards more connected officers' biases against a finding. As the discussion below elucidates, the "connectivity" analysis relies on a difference between the greater connectedness of shooters (officers have on average 20 connections) and smaller connectedness of non-shooters (on average 12 connections). If non-shooters are more connected than other officers not in the network, then the difference between shooters and non-shooters should be smaller, not larger. Because we cite to a large difference in connectivity to show pathways of potential transmission, any relationship between complaints and connectivity actually biases the inquiry against finding an effect.

B. Data processing

In this section, we describe how we used complaint and shootings data to construct the two networks under study: (i) a co-offending police complaint network containing 9737 unique officers and (ii) a shootings network containing 488 shootings with 418 unique officers.

1. Missing Data

We first cleaned the civilian complaints to fill in missing information from other records. The number of complaints collected during the period of study totaled 107,276. A significant number of complaints were missing any officer identities, race, gender, dates of appointments and other information, and another group of complaints were missing affidavits from the complainants. Our dataset included 19,843 complaints involving 9737 unique officers that had complete information.

We obtained personnel records on all Chicago police officers for the relevant time period January of 2008 to November 2015 by filing a Freedom of Information Act (FOIA) request with City of Chicago Department of Human Resources. The bi-annual personnel data include each officer's name, position title (rank), and original hire date. We use the officers first and last name and the original date of appointment to link records with the other datasets.

We linked information across datasets (the civilian complaint dataset and the shootings dataset) by constructing a unique identifier for each and every officer on the roster. In the event that officers

---

[39] The complaint network contains 9737 unique officers. For the relevant dates, on average, the personnel rosters from the CPD show that there were between 12,000 and 13,000 officers in the department at any one time.

shared names, we verified the officer's identity through badge numbers and/or date of appointment where available on either the relevant complaints or the shootings data from IPRA.

2. Forming the Complaint Network

Drawing from the civilian complaint data, we constructed a complaint network that connects officers whose relationship is reflected in the fact that they appear on the same civilian complaint. Previous research suggests that co-offending represents strong and enduring relationships between individuals.[40] We therefore treated co-offending as evidence of an existing relationship between the two individuals involved, rather than as a point-in-time estimate or marker of when that relationship formed. We generated a static snapshot of the network: an edge indicates that two individuals have co-offended together at least once at some time during the study period.
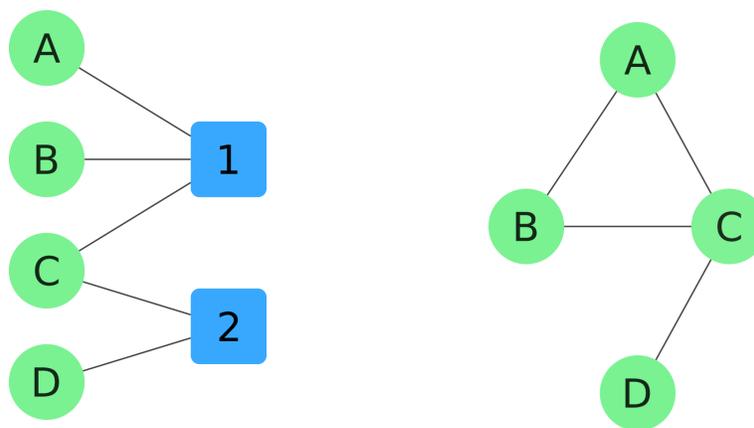

Fig. 1 Co-offender Complaint Network. On the left, the two-mode relationship between officers (letters) and complaints (numbers); on the right, the resulting one-mode officer relationship.

After identifying unique officers and civilian complaints that contain allegations against these officers, we created a two mode (person-by-event) matrix that links officers to a specific complaint. Fig. 1 on the left depicts the relationship of officers to complaints.

We then converted this matrix to a one-mode, person-by-person matrix by assuming that officers who respond to calls together and engage in conduct that triggers a civilian complaint are associated with each other and interact in ways that facilitate the spread of shooting violence. We draw an edge between two officers who are listed on the same complaint. Fig. 1 on the right depicts this person-by-person relationship.[41]

This network connects every pair of officers who were listed together on the same civilian complaint during the period of study. The complaint network contains 9737 unique officer nodes and 47,037 edges. Importantly, we did not weight the edges of our complaint network to reflect the number of times that a pair of officers are listed together on the same civilian complaint. The vast majority

---

[40] Michael Warr, COMPANIONS IN CRIME: THE SOCIAL ASPECTS OF CRIMINAL CONDUCT (2002); .

[41] See Stephen P. Borgatti and Martin G. Everett, Network Analysis of Two-Mode Data, 19 Social Networks 243 (1997) (describing the construction of a network of co-offenders linked through a common document).

(83%) of edges have a weight of 1, meaning that for the vast majority of officer pairs, these officers appeared together on a complaint only once. The average officer received 1.72 complaints. Approximately 15% of officers have 10 or more unique co-complaint recipients for civilian complaints. The top 1% of officers have received a civilian co-complaint with 26 other officers, on average.

3. Forming the shootings network

We created a second and separate shootings network by layering shooting data onto the complaint network. This multi-layered shootings network maps the relationship among shooting events committed by officers who are also listed on the same civilian complaint.

We used IPRA incident-level shooting data to layer shooting events occurring during the period under study onto the complaint network of co-listed officers described above. For each shooting event, we recorded the date and location (street address) of the shooting. We added this information to the complaint network by adding the shooting event data to the officer-nodes in the network, and in particular, to the officers who were recorded as having committed the shootings.

From this layered information, we constructed the shootings network as a network capable of being analyzed separately from the complaint network. In this network, a vertex in the shootings network represents a shooting event at a particular time and date, paired with the officer who committed the shooting. Edges connect shooting events on the underlying complaint network, for which the pair of officers who engaged in the shootings appeared together on a civilian complaint at any time during the period under study. As with the underlying complaint network, the shootings network is a static network, representing a snapshot of all connected shootings that occurred during the period of study. Fig. 2 below depicts the construction of the shootings network from the complaint network.
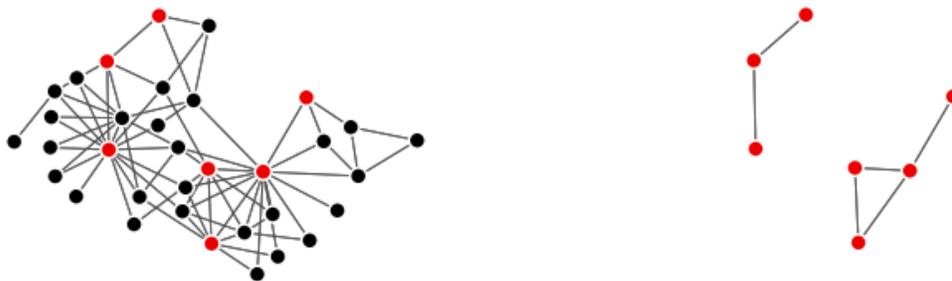

Fig. 2 Construction of the shooting network. On the left, the complaint network with shooting officers in red; on the right, the resulting shooting network.

In contrast to the complaint network, the shootings network is an event-focused network. Some officers are involved with more than one shooting. A significant fraction (86%, n=358) of officers in the complaint network were associated with only one shooting. A smaller fraction (12%, n=51) of officer nodes were associated with two shootings; 2% (n=8) were involved in three shootings, and one officer (Tracey Williams) was involved in five shootings. All of these shootings are represented in the shootings network.

4. Largest Connected Component

Mapping the network yields several smaller components and one giant connected network component. As is true for most social network analysis, we use this largest connected component (LCC) as the focus of our study, for both the complaint network and shootings network.

The largest connected component in the complaint network yields 80 % of the nodes in the network (N = 7825), and 99% of the edges (N = 46,568). As well, the LCC of the complaint network contains 418 of the 450 shooting officers in the larger network. As detailed in the Appendix, the LCC also has a similar clustering coefficient, an average path length, and its degree distribution follows a power-law distribution as is true for the larger network.

	Complaints	LCC	Shootings	LCC	Shooting officers	LCC (complaint)
Nodes (No)	9737	7825 (80%)	488	326	488	418
Edges (E)	47,037	46,568 (99%)	?	487	713	663 (93%)

Table 2 Complaints and Shootings Data

III.   Conceptual Framework for Analyzing Contagion

To determine whether police shooting is contagious, we investigate whether exposure to a police-involved shooting on an officer's network increases the likelihood that the officer will engage in a police shooting later in time. We approach this question in three ways. First, we investigate the structure and connectivity of the relevant networks to see if a greater potential for pathways of transmission exists among shooting officers.

Second, within the shootings network, we compare the periods between shootings for connected officers with those for randomly chosen officers, to determine whether the probability of shooting for connected officers will be higher in a given period of time. Third, we fit a dynamic Hawkes model of contagion to our observed data, using maximum likelihood estimation to estimate the parameters of the contagion effect. We use these parameters to generate various estimates of the magnitude of the contagion effect.

Network models of contagion have been developed to describe the diffusion of disease (or information, money, data, etc.)[42] While many models assume that a population is well mixed and will interact randomly, more recently developed models take into account that people do not randomly interact with one another. Most people have a network of coworkers, neighbors, friends and family with whom they interact regularly, and others with whom they almost never interact. The architecture of this network can shape the likelihood of interaction and the resulting dynamic patterns of contagion.

---

[42] For an overview of network research applied to sociology, criminology, public health, psychology and a wide range of other disciplinary social science investigations, see Mark Newman, NETWORKS: AN INTRODUCTION (2010).

The dynamics of contagion on a network resemble a branching process, as each infection gives rise to new "offspring" infections. In a highly connected network, these branching processes may trigger a spatial and/or temporal cascade or cluster of infections on the network, as infections trigger new infections close by in space and time. Modeling and analyzing such processes can be challenging.

Scholars from a number of disciplines have begun using point process models to describe the branching process of contagion, in which each infection triggers other infections. A point-process model counts events, or points, as they occur in space and/or on a timeline. These models also keep track of the intervals of time between events, called inter-event periods, in order to describe the distribution of points using probability distributions.

For example, Poisson processes describe random distributions of points on a timeline or in space. In a Poisson process, the probability of an event is independent of the occurrence of other events, and the occurrence of an event has no effect on the probability of a further event occurring—what some called a process without after-effects. If police-involved shootings followed a Poisson process, shootings would occur completely independently of each other, with no relationship between the occurrence (or timing) of earlier shootings and the probability (or timing) of later shootings.[43]

Contagious distributions, in contrast, show dependencies: with some probability, earlier shootings trigger later shootings and the timing of an earlier shooting affects the timing of a later shooting. As early as 1931, Pólya introduced a family of probability distributions that he and colleagues called "contagious distributions," to describe distributions in which the occurrence of an event increases the probability that a further event would occur during a given period of time. In these contagion distributions, history plays an important role, as events trigger future events.[44]

The Hawkes model, developed by Alan Hawkes in the 1920s, describes the dynamics of these contagious distributions.[45] The Hawkes process model is a self-exciting point-process model that studies infectious processes in which an earlier event has an influence on the occurrence of a later event. The term self-exciting point-process models describes a class of models in which events at time t are dependent upon events of the same type, e.g. police shootings, at an earlier time, $t - i$.

The Hawkes analysis is probabilistic: it focuses on whether exposure to an earlier event, here by a network neighbor, increases the likelihood of a later event. As applied to shootings, in the Hawkes model of contagion, the effects of multiple exposure are additive: an officer exposed to multiple earlier shootings experiences an increase in probability that is the sum of influence from all earlier shootings.

More specifically, the Hawkes model assumes that the instantaneous likelihood of an officer shooting a civilian at a particular time depends on the sum of two factors: (i) the background rate of shooting, owing to other factors that are independent of contagion; and (ii) the net increase in

[43] See Jerzy Neyman, On a new class of "contagious" distributions applicable in entomology and bacteriology, 10 Ann. Math. Statist. 35 (1939).

[44] Georges Pólya, Sur quelques points de la théorie des probabilités. Ann. Inst. Henri Poincaré **l,** 117-162 (1931). Florian Eggenberger, and Georges Pólya, Über die Statistik verketteter Vorgänge. Z. Angew. Math. Mech. 3, 279-289 (1923).

[45] Alan G. Hawkes, Spectra of some self-exciting and mutually exciting point processes, 58 Biometrika 83 (1971).

probability from exposure to earlier shootings by network neighbors (taking into account that the influence of those earlier shootings decays over time.)

$$\lambda(t) = \lambda_0(t) + \alpha \sum_{t_i < t} \phi(t_i - t)$$

instantaneous probability of shooting at t   background rate at t   historical kernel: sum of influence from earlier events

- λ_0 is the background rate. We will consider it constant at the outset, $\lambda_0(t) = \lambda_0$, but in further refinements, we will include a function that reflects the seasonal and day of week variation in shootings for police officers.

- α is the "influence" rate: this parameter measures the scaling up of probability above the baseline rate that an earlier shooting event has on the probability of a future shooting. Put differently, this parameter measures the increase in probability of a later shooting from exposure to an earlier shooting. We have the constraint $\alpha \in [0,1]$.

- ϕ is the historical kernel function that reflects the influence of earlier shooting events from network neighbors. This function should be non-negative and causal, i.e., $\phi(t) = 0$ for $t < 0$. This implies in particular that influence is only summed over the earlier data points $i \in [N]$ such that $t_i < t$. We assume that social influence decays over time exponentially. Thus, we choose an exponential kernel to describe an exponential decay of social influence over time:

$$\phi(t) = \begin{cases} \exp(-\beta t) & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where β is the decay rate.

For our purposes, the Hawkes process model is most useful because it separates the role of contagion from the other factors unrelated to contagion that affect the rate of shootings. Not all upticks in the rate of an event come from contagion. For example, an outbreak in food poisoning from contamination could explain a dramatic spike in poisoning over the ordinary base rate, which might otherwise appear to be evidence of contagion.

IV.   Analysis and Results

A. Network Descriptions and Statistics

Table 3 presents summary descriptive data collected for each of the complaint and shooting networks, and visualizations of each network.

Descriptive Feature	Complaints (SD)[46]	Shootings (SD)
Average Degree	11.64	20.56
Density	0.002	0.016
Average Path Length	5.21	8.04
Transitivity	0.42	0.39
Modularity Score	0.567	?

Table 3 Features of complaint and shooting networks by network type

The corresponding mean count of complaints per officer received is 1.72. Among officers named in at least one civilian complaint, approximately 78% are co-named alongside another officer at least once during 2008 to 2015. One in four officers (25%) did not receive a civilian complaint alongside another officer.

Table 4 below presents summary statistics on the complaint and shooting networks. The vast majority of officers in both networks are male (86% and 89%). White officers make up just over one half of each complaint and shootings network (53% and 52%), on average, with substantial variation across districts. For both networks, approximately one in five officers in the data are black and one in four officers are Hispanic, Asian, Pacific Islanders, or Native American. The mean tenure is 4 years and 3.8 years in the civilian-facing and shooting networks, respectively.

Statistic	Complaints (SD)	Shootings (SD)
Proportion male	0.86 (0.02)	0.89 (0.04)
Proportion white	0.53 (0.10)	0.52 (0.13)
Proportion black	0.20 (0.13)	0.23 (0.17)
Proportion Hispanic and other	0.27 (0.08)	0.25 (0.09)

Table 4: Summary statistics for the district co-complaint networks by network type

The distribution of complaints among officers shows a small number of officers with a high number of complaints. The majority of officers have received co-complaints with fewer than two of their peers: 58% of officers have 2 or fewer unique co-complaint recipients for civilian complaints. Approximately 15% of officers have 10 or more unique co-complaint recipients for civilian complaints. The top 1% of officers have received a civilian co-complaint with 26 other officers.

Fig. 3 depicts a visualization of the complaint network, for a single police district. We can observe a large connected component with a dense core at the center of District Five. Other work on the

---

[46] To be calculated.

structure of misconduct describes the architecture of such a network, and the various factors that influence the likelihood that two officers will form an edge or tie between them. This network resembles most social networks in its architecture in that it appears to be a network characterized by preferential attachment, in which those who have a greater number of connections acquire even more.
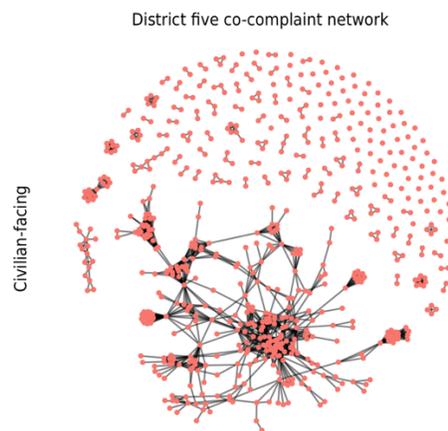

Fig. 3 Co-offending Complaint Network Visualization (District 5)

B. Evidence of Contagion

1. Connectedness and Clustering

Evaluating the connectedness and clustering of shooting officers relative to randomly chosen officers can offer some early evidence of the possibility of contagion. Evidence of relatively greater connectedness and clustering among shooting officers is consistent with a finding of contagion because such evidence demonstrates the relatively greater potential for transmission of social scripts (or other information) among shooters. To compare the connectedness of shooters to that of officers in general, we average the number of edges per node, or "degree," for shooting officers and compare this number to an average degree for randomly chosen officers in the underlying complaint network. Similarly, to compare the clustering of shooting officers to randomly chosen officers, we count edges between shooting officers and compare this number with a count of edges among an equal number of randomly selected officers (418) on the underlying complaint network.

Table 5 below shows structural evidence on the relatively greater connectedness of shooting officers that is consistent with the contagious spread of shooting violence on the officer complaint network. Comparing shooters and randomly chosen officers on the complaint network reveals that shooting officers are more connected to each other than randomly chosen officers. Shooting officers have an average degree of 20, compared to randomly chosen officers on the complaint network, who have an average degree of 12.

	Random	Involved in Shooting
Avg. Degree	12	20
Avg. Edges	112	487

Table 5 Connectedness (degree or links per node) and clustering (edges)
on complaint network by officer type

Shooting officers are also more clustered together than non-shooting officers. A full 53% of shootings and their associated officers are connected to at least two other shootings and shooting officers , compared to only 24% of non-shooting officers. Likewise, shooting officers are more tightly clustered than are non-shooting officers in the complaint network. As Table 5 shows, on the complaint network, shooting officers are connected by a total of 487 edges, where a random subset of 418 officers are connected by only 112 edges.

This evidence is consistent with a finding of contagion. More connections and greater clustering among shooting officers offer potential pathways for transmission of shooting violence. Of course, this evidence might also be consistent with homophily; particular officer traits like race, gender, district, membership in a special unit or tactical team might explain the existence of more edges among shooting officers. We rule out homophily as a potential explanation in the discussion below.

2. Periods Between Shootings

Beyond the architecture of the shooting network, dynamic evidence is also consistent with the contagion of police violence. We measure periods between shooting events to calculate inter-shooting periods, as a measure of the increased likelihood of shooting from exposure to earlier shooting events. We compare inter-shooting periods by officers who are connected on the shootings network to inter-shooting periods for shooting officers who are randomly selected from the network. Shorter periods for connected shooters suggest an increased probability of shooting in a given period of time.

Table 6 shows that inter-shooting periods are in fact shorter (739 days) on average for shootings committed by officers who are connected on the network than for randomly chosen shootings on the network (994 days). This evidence is consistent with the contagiousness of police violence: shorter shooting periods imply that in a given time period, there is a higher likelihood that an officer who is directly exposed to shootings will subsequently engage in shootings.

Group	Avg. inter-shooting period (days)[47]
Connected Shooters	739
Randomly Chosen Shooters	994

Table 6 Inter-shooting periods by group type

---

[47] SD TBD (given that quantity is important, better to include distributions with regard to the delays).

3. Homophily and Omitted Variables

Factors other than contagion could potentially explain our results in Tables 5 and 6. For example, common membership on a tactical team could explain both the relatively higher incidence of edges among shooters as well as shorter inter-shooting periods among connected shooters(because tactical officers are more likely to be connected, and shoot more frequently compared to officers who are not on these higher-risk tactical teams). Likewise, the clustering together of "bad cops"—officers with higher numbers of shootings—might generate evidence consistent with our results. Still other omitted variables might explain both periods and edges: for example, a common commanding officer among connected shooters might explain shorter inter-shooting periods for this group if the officer counsels more aggressive interaction with civilians.

To separate the effects of homophily or other omitted variables from those of contagion, we use a shuffle test, randomly shuffling the dates of officer shootings to observe the effect on the duration of inter-shooting periods. This shuffle test is based on the idea that if contagion is present, an officer's shooting will depend on the timing of other connected officer prior shootings: contagion involves sequential shooting times relatively close in time, as shootings spread from officer to officer.[48] Conversely, in the absence of contagion, the timing of officer shootings should be independent of the timing of other officer shootings. Likewise, if homophily or some other omitted variable plays a role in our results, an officer's probability of shooting should depend on whether connected officers are shooters, but should not depend on *when* connected officers shoot.

For the shuffle test, we (i) identify all shooters on the underlying complaint network; (ii) randomly permute the shooting dates by randomly reassigning shooting dates from one officer to another; and then (iii) measure the period of time between shootings among connected officers who are directly connected by an edge. All else remains the same: officer identities, officer neighbor positions and edges among officers. Thus, we draw from the exact same distribution of shooting times as in the observed data, with the exact same network architecture, but shuffling disrupts any potential dependencies in the timing of officer shootings. After date shuffling, we compare the average duration of inter-shooting delays for shuffled dates with the average associated with observed data.

Fig. 4 below illustrates the shuffle test with hypothetical dates. In this figure, the observed shooting dates show close-in-time and sequential shooting times consistent with contagion. Shuffling dates disrupts those dependencies, and increases the average inter-shooting periods, as seen on the right.

---

[48] For an extended discussion of randomization as a test for omitted variables and homophily, see Aris Anagnostopoulos, Ravi Kumar, and Mohammed Mahdian, Influence and correlation in social networks, in KDD '08: Proceeding of the 14th ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING (2008); Rodrigo Belo and Pedro Ferreira, Identifying Social Influence in Viral Products Using Randomization over a Large Mobile Network (working paper 2016).
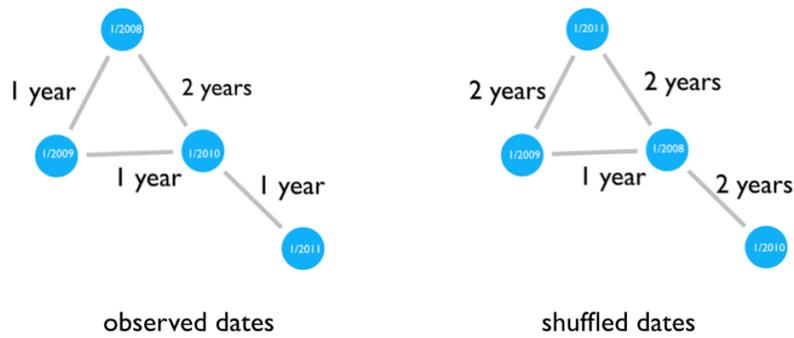
Fig. 4 Shuffle Test

To conduct the shuffle test, we ran 10,000 Monte Carlo simulations in which we randomly shuffled the dates of shooting events for all shooting officers, and then measured the average period between shootings of officers connected directly by an edge.

Our shuffling test results are reported in Table 7 below. For the general population of shooters on the complaint network, we find that shuffling dates lengthens the period of shooting delays, from 739 days for observed dates to 993 days for shuffled dates (consistent with the measurement between randomly selected shooting officers, which was 994 days). This result is consistent with the presence of contagion. As described earlier, the timing of an officer's shooting appears to depend on the timing of other officers' shootings, and shuffling disrupts that relationship, with no remaining dependency created by an omitted variable. For that reason, these results suggest that an omitted variable common to all shooters is not likely to explain the relatively shorter inter-shooting period among connected officers.

Group	Avg. inter-shooting period (days)[49]
Connected Shooters (observed)	739
Shuffled connected shooters	993

Table 7 Inter-shooting delays by group type

Using the same logic regarding the timing of shootings, we further test specifically for homophily by investigating within-group variation between connected and unconnected officers. Here we construct artificial homophily groups, and then randomly shuffle the dates of shooting events within those groups. We focus on traits or group identities that might potentially explain both greater edge density and shorter inter-shooting periods among officers connected by an edge. As before, if homophily is the sole source of our effect, then shuffling the dates should not change the duration

---

[49] TBD (given that quantity is important, better to include distributions).

of inter-shooting periods.[50] Conversely, if contagion plays a role, then shuffling dates within artificially-constructed homophily groups will change the duration of the periods between shootings.

To shuffle dates within artificial homophily groups, we (i) divide officers into various permuted groups on the basis of potentially homophilous traits or group-identities—for example, numbers of shootings or complaints (one vs. more than one), race, gender, or membership in a tactical team (TBD); (ii) randomly permute the shooting dates within the identified group by randomly reassigning shooting dates from one officer to another in the group. All else remains the same: officer identities, officer neighbor positions and the edges among officers. After date shuffling, we again measure the average duration of inter-shooting delays. Because averaging can mask the effects of particular groups, we also report measurements for individual groups (for example, a particular district) that show the smallest differences between shuffled measures and observed measures. These "smallest-difference" measures set the lower bound on homophily.

Table 8 below, which measures inter-shooting periods for shuffled homophily groups, shows that homophily does not appear to explain our results. Recall that the average inter-shooting delay for shootings connected by an edge is 739 days. The average period for a group of connected officers within groups categorized by race is 994 days; for groups categorized by gender is 991 days; for groups categorized by district is 957 days; and for groups categorized by frequency of shooting (one shooting versus more than one shooting) is 981 days.

Even the smallest reported differences leave behind a significant difference between observed connected shooters and shuffled connected shooters. Thus, homophily does not appear to explain or be associated with the shorter inter-shooting delay for shootings that are connected on the network, and contagion may play an important role. Interestingly, organizing around district appears to show that homophily with regard to district may play a very small role.

Group	Avg. inter-shooting period (days)[51]	Smallest reported difference (days)
Connected Shooters (observed)	739	--
Shuffled Race	994	? (TBD)
Shuffled Gender	991	?
Shuffled District	957	?
Shuffled # of Shootings	981	?

Table 8 Inter-shooting delays by artificial homophily group type

It is important to note the limits of this analysis. Randomization provides important information about the role that contagion plays. Further, controlling for the number of shootings controls not just for this group but for a number of other groups both tested and untested. For example, if district membership were to explain both increased edges among officers and shorter inter-shooting

---

[50] Aris Anagnostopoulos, Ravi Kumar, and Mohammed Mahdian, Influence and correlation in social networks, in KDD '08: Proceeding of the 14th ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING (2008); Rodrigo Belo and Pedro Ferreira, Identifying Social Influence in Viral Products Using Randomization over a Large Mobile Network (working paper 2016).

[51] TBD (given that quantity is important, better to include distributions).

periods, it likely would be because officers in certain districts shoot more often in a given period of time. Accordingly, controlling for number of shootings operates to control for any unobserved group membership that would explain edges and periods because the group shoots more frequently.

At the same time, it is important to note that if there are non-observed confounded variables or homophily traits correlated with the timing of the shootings but not with any of our observed variables, including the number of shootings, then our analysis cannot distinguish homophily from contagion without use of an instrumental variable (which is unlikely here, owing to the non-random assignment of officers to units and the limited availability of other data) or a randomized design. The Appendix contains the description of a fanciful hypothetical to illustrate the kinds of confounding variable not detected by our analysis.[52]

4. Maximum Likelihood Estimation of the Hawkes Model

Social influence is not the only factor that affects the probability of police shootings. The dependent branching process of observed data might also be explained by background factors that influence the rate of police-involved shootings independent of contagion, owing to a range of factors that affect both individual and groups. This background rate of shooting will inevitably affect any measurement of probability owing to contagion.[53]

To differentiate the relative contributions of background rate and social influence, we estimate background and contagion parameters by fitting a Hawkes model of contagion to the observed data. As described earlier, the Hawkes process is a "point process" that involves discrete points or events located in time and space—in our case, individual shootings. In "self-exciting" point processes, past events influence the occurrence of future events: for example, past infections trigger future infections, and earthquakes give rise to aftershocks.

In the Hawkes model, past events influence the occurrence of future events, and frequently, that influence decays over space and time. Point process models have been used to study contagious influence in disease, urban crime, high frequency trading, economics, neuronal firing and earthquakes, among other things.

We modeled the contagion of gun violence as a Hawkes process. Each network node on the shooting network represents a shooting, paired with the officer who committed the shooting. In the Hawkes process model of shootings, the probability that an officer will shoot a civilian at a particular time depends linearly on two contributing factors: (i) the background or "intrinsic" rate at which police shootings occur owing to all other factors independent of other shootings, and (ii) a "history kernel" that reflects the cumulative influence of all earlier shootings by an officer's network neighbors on the relevant network. Equation 1 below describes this relationship.

---

[52] See Appendix for a fanciful hypothetical example, the Sunday Shooter Syndrome, to illustrate such an omitted variable.

[53] We assume at the outset that the background rate of shooting is constant for all officers. Preliminary diagnostics suggest that background rates of police shooting do not vary significantly time of year and weekend.

We can define the instantaneous probability of the officer-involved shooting for each event on the network $v \in V$ using the following intensity (instantaneous probability) function derived from the Hawkes process model:

$$\lambda_v(t) = \lambda_0(t) + \alpha\beta \sum_{i:(v_i,v\in E)} e^{\beta(t_i-t)} \text{ [54]} \tag{1}$$

As explained earlier, λ_0 is the background rate, and α is the social influence rate—the scaling up in probability that comes from exposure to a shooting. β is the rate at which the social influence decays over time.[55] We can estimate the value of these three parameters—the background rate, the social influence rate from earlier shootings per shooting, and the rate at which this influence decays--by way of maximum likelihood estimation.

Hawkes processes possess flexible statistical properties that allow the analyst to incorporate auto-correlations and self-exciting features. In contrast to time-series models, Hawkes processes remain analytically tractable. We can analytically compute likelihood functions as well as conditional distributions, moments, Laplace transforms and other characteristic functions. Owing to their tractability, Hawkes processes have been applied to fit models to data in seismology, biology, criminology and finance.

We derive in the Appendix the log-likelihood function corresponding to Equation 1 for use in maximum likelihood estimation. This log-likelihood function describes the (joint) probability of our observed data as a function of data and parameters. To find the parameter values that maximize the log-likelihood of our observed data, we solve an optimization problem that maximizes the log-likelihood function. Because the function is complex, and no analytical solution is possible, we used numerical optimization using grid search, gradient descent, and a combination of the two techniques.

We first used grid search, exhaustively searching a bounded space in three dimensions (corresponding to three parameters) in intervals of .01 for parameters that would maximize the log-likelihood of our observed data. As described more extensively in the Appendix, we also engaged in gradient descent, taking the derivative of the log-likelihood function to determine the gradient of the function, and then iteratively searching the path of steepest ascent until the gradient had flattened, indicating that we had reached a maximum. We initialized this gradient descent search with 100 random starting points consisting of three randomly chosen parameter values.

Finally, we combined grid search and gradient descent. After a grid search, we then input ten of the most promising three parameter combinations as our initial point of departure for a subsequent algorithmic search. Searches were run until parameters converged. For comparison's sake, we compared our results from our chosen algorithm with results generated by multiple other search algorithms. As detailed in the Appendix, all methods and all algorithms converged to the same parameter values.

---

[54] It is important that our Hawkes process be well-defined such that $\alpha \int_0^\infty \phi(t)dt < 1$. If not, the process will diverge and there can be infinitely many events occurring at all time t. For this reason, we normalize the kernel function so that its integral is 1 and we add the constraint $\alpha \in [0,1]$.

[55] We also add to the equation a β multiplier modifying the historical kernel, for purposes of normalizing. This is for mathematical convenience only and does not change the quantitative results.

5. Interpretable Results

    A. Definitions of background and contagion shootings:

We define a "background shooting" and a "contagion shooting" using the background shooting as reference. A background shooting is a shooting for which no earlier shooting influences the probability of the shooting. A contagion shooting is a shooting for which an earlier shooting influences the probability of the shooting. Importantly, a contagion shooting is a shooting for which both the background likelihood of shooting and the influence of an earlier shooting can contribute to the probability of a shooting.[56] The Appendix contains the mathematical definitions for each type of shooting. For all calculations, we reduce the 8 years of the study period to a unitary 0 to 1 timeline, where 1 represents the end of the study period.

    B. Parameters of the Hawkes model that maximize the likelihood of observed data

We obtained the following values of the parameters at the optimum

λ_0 (background) = 3.89 X 10^{-2} α (social influence) = 1.9 X 10^{-2} β (decay rate) = 4.388

- λ_0 represents the instantaneous background rate of shootings that are primary shootings, independent of any social influence from exposure to earlier shootings. We estimate this parameter value to be 3.89 X 10^{-2}.[57] Much like multiplying instantaneous speed by time of travel can generate distance measures, we can multiply the instantaneous rate of background shootings by the period under study , to obtain the number of shootings in the study that are background shootings, meaning that they are not influenced by earlier shootings. This calculation is below.

- The parameter α, represents at a high level the cumulative contagion effect of exposure to a single earlier shooting. We estimate this parameter to be 1.9 X 10^{-2}. More formally, this rate equals the average increase in the total number of future shootings by a network neighbor given a previous shooting, and the existence of an edge between the officer and a network neighbor. Further, by incorporating the structure of our complaint network, we can also compute the threshold value of α past which the network generates an "epidemic" cascade of infections of violence. We find that this critical threshold value of α is 2.38 X 10-2, which

---

[56] Mathematically, we could also interpret our results to probabilistically assign shootings to one of two categories, either background or contagion, as these two interpretations are mathematically equivalent. Interpreting shootings as a mixture of contributions from background and contagious influence is more in keeping with the conceptualization of contagion as a force that increases the likelihood of shooting beyond an existing baseline of background probability.

[57] The relevant time scale of the study, eight years, is reduced to a unitary 0 to 1 timeline.

is greater than the value from our maximum likelihood estimate, and hence our model does not predict this cascading effect.[58]

- The parameter β represents the rate of decay for the social influence of a single shooting. More specifically, this is the rate at which the previously-described increase in probability associated with α decays over time. We estimate this parameter to be 4.388. We can calculate the half-life of social influence as follows: $1/\beta = 1/4.388 * 8$ years = 1.82 years.

We can calculate several statistics of interest from these parameters:

- First, we calculate the expected number of offspring shootings to which the average shooting "gives birth" over the period of study. To obtain this number, we assume that the first shooting happens at a random location in our network of police officers, and then keep track of the number of offspring and offspring's offspring over the period of the study. We estimate that the first shooting of the period will give birth to an additional **0.496 shootings** over the study period. The mathematical calculation is contained in the Appendix.

- Second, we calculate the expected number of shootings that are influenced by earlier shootings in any way. If N is the total number of shootings, the number of expected shootings associated with contagion is equal to N minus the sum of the probabilities that are associated with background shootings. For our data, we estimate that **141 shootings** are the product of contagion. The calculation is included in the Appendix.

- Third, we specify the way in which contagion-influenced shootings are influenced by earlier shootings. Survival functions track the way in which some object of interest (say a patient) dies before a given time Here, we describe a general "survival function" to track the way in which a network neighbor officer shoots before a given time. We can track the relationship between the level (number) of exposures and the probability that the network neighbor shoots before a given time. More specifically, we plot in Fig. 3 the probability that Officer 1 will shoot before time t after exposure to K shootings by Officer 2 (alternatively, shootings by two network neighbors in a short period of time), assuming that Officers 1 and 2 are network neighbors and connected by an edge.[59]

---

[58] This critical threshold is based on the maximum eigenvalue of the network adjacency matrix K. To calculate the threshold, the maximum eigenvalue of K must be small enough that α multiplied by this maximum eigenvalue is less than 1. The maximum eigenvalue of K is 41.9. Thus, α must be less than 1 / 41.9 = .0238.

[59] On this figure, the x-axis measures time since K shooting events by Officer i_2 at t =0. The y-axis represents the probability of shooting by Officer i_1 at time t after Officer i_2's shooting. Formally put, given two random Officers i_1, and i_2 in the network, and given that Officer i_2 has shot exactly K times at t=0, the probability that Officer i_1 shoots before a given time t (x-axis) is given by p(t,i_1,i_2, K).
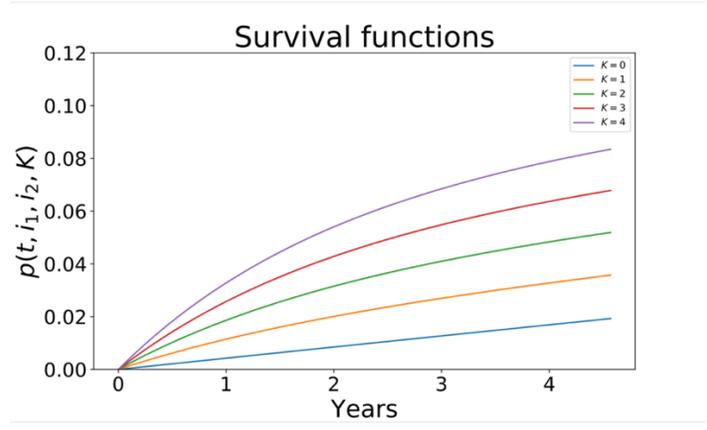
Fig. 5 Probability that Officer 1 will shoot by time t after Officer 2 shoots K times at t=0.

Fig. 5 illustrates the way in which an officer's exposure to shootings changes the probability of future shootings. Suppose we draw a vertical line at $T = 3$ years. Officer 1's **exposure to a single shooting more than doubles the probability of a future shooting** for a network neighbor within the first two years after exposure, compared to the probability of shooting with no exposure.

We also observe from Fig. 5 that being exposed to K shootings at $T = 0$ increases the probability of shooting before T years by at least a factor of K. (This heuristic appears to hold for any number T of years into the future). Put more simply, exposure to K shootings scales up the probability of a future shooting by a factor of K.

- We can also examine individual shootings to discover the relative contributions of contagion and background. For example, for Shooting #319, contagion contributes 90% of the probability of the shooting (p_contagion = 0.90.) We can identify the earlier shootings that were likely sources of contagion, in addition to identifying the timing of those shootings and their associated contagion effect.
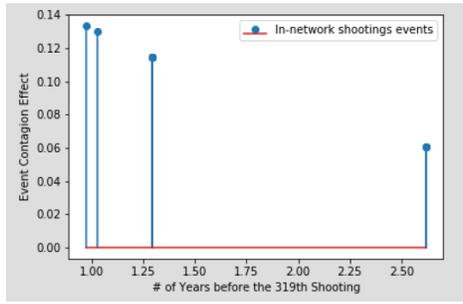

Fig. 6 Four prior neighbor shootings and associated contagion effects for Shooting #319

In this case, Fig. 6 shows that the officer responsible for Shooting #319 had been exposed to four earlier shootings, and the figure displays the relative contributions of those earlier shootings to the contribution from contagion.

In sum, fitting the Hawkes contagion model to the observed data, and the resulting statistical analysis, produces evidence that is strongly consistent with contagion. The first shooting of the period of study produced .496 additional offspring shootings over the period of study. Contagion contributed to the probability of 141 of 488 shootings in the dataset. Exposure to a single shooting doubles the probability of exposure for a network neighbor with the first two years.

Further research will re-fit the Hawkes model to the data in a Bayesian framing. Calculating posterior probabilities will enable us to determine how sensitive the log-likelihood of observed data is to parameter variation, to determine whether the Hawkes model is well-fitted to the data. Such analysis may enable us to make stronger causal claims with regard to exposure and future shootings.

V.   Discussion

Our findings advance the study of police violence by marshalling evidence that is consistent with the finding that police shooting is contagious. Understanding police shootings as contagious can potentially supplement violence prevention efforts that focus only on individual officers or departmental culture. Traditional scholarship explains police violence as either top-down, involving police department culture or leadership, or bottom-up, involving the individual traits of bad-apple officers. To supplement these micro- and macro-explanations, our investigation illuminates the dynamics of police violence at the meso-level, located in the patterns of social influence that shape behavior. Emergence of violence from the collective is much more difficult to manage from a policy perspective, because intervention points are so diffuse. Violence prevention strategies must disrupt transmission or diminish the strength of social learning to be effective, an insight we discuss more below.

Beyond this meso-level insight, these models also underscore the general importance of a dynamic approach to violence, an approach that incorporates not just the role of time but also the role of space and network structure into theory about police violence. In terms of structure, the architecture of police networks condition the likelihood of police interaction, and the corresponding flow of social influence. Officers who are close to each other on a police network are likely to interact with and influence each other. Network connections often correspond to organizational connections—officers in the same district are likely to interact—but here, some cross-district connection and special unit interactions with civilians illuminate the value of mapping the network.

With regard to the role of time, dynamic analysis of the contagion model illuminates how police shootings can become self-reinforcing over time. These results suggest that police violence can endogenously increase over time, as earlier events trigger later events in cascades of events. Officers who are exposed to multiple shootings appear to be at particularly at risk, as the effect of exposure is cumulative.

Focusing on the contagion of police-involved shooting gives rise to a whole new range of strategic interventions to reduce violence. Current research on violence prevention includes the use of

machine learning to create early intervention systems, focused mostly on individual officer traits.[60] Research on the contagiousness of violence supports the need to expand the focus of such programs to include a focus on the contagious transmission of violence from officer to officer.

Potentially effective strategies could include structuring assignment practices so as to limit repeated exposure to shootings, moving vulnerable officers from locations on networks with high exposure to locations with less exposure, or pairing officers who have engaged in shooting with more senior mentors who are trained in de-escalation. Police departments could regularly rotate officers among units and districts to mix exposure to shootings more uniformly among officers, paying attention to whether the infected officer is senior or junior (which affects infectiousness) and the number of officers with whom the officer will have contact in his new network location.[61]

Alternatively, departments might also investigate the possibility of "seeding" the contagious spread of good scripts, like de-escalation of violence. Officers with the highest degree centrality in the network—meaning the most connections to other officers—might make good candidate seeds for spreading good scripts.[62] In the same vein, departments might choose to assign officers who have engaged in shooting to be paired with senior officers who have not engaged in shooting or who have been trained in de-escalation of violence.

Future research will focus on confirming the contagion dynamic, ascertaining the mechanisms that transmit the contagion of police-involved shootings, understanding the features of a department that make it vulnerable to contagion, and mapping the dynamic patters of violence contagion at the population level. Beyond confirming the existence of contagion in other departments, ascertaining the precise mechanism of transmission is an important next step. Qualitative investigation through survey and interviews of officers would help to determine whether officers learn scripts involving shooting from each other, what those scripts look like and under what conditions such social learning happens. Interviews could also shed light on alternative, non social-learning mechanisms as well. For example, surveys might illuminate whether officers infer from exposure to a police-involved shooting that they are at greater risk from civilians, which might explain an increased likelihood of a future shooting.

Importantly, the argument that police shooting is contagious serves to supplement, rather than supplant, existing theories of police violence. At the same time, if police violence is contagious, it introduces the possibility that most departments are at risk from a few bad apples. After all, bad apples are a problem because they spoil the whole barrel: spoilage is contagious. This project suggests that policymakers in this space will need to attend to contagion in thinking seriously about violence prevention.

---

[60] Samuel Carton et. al, Identifying Police Officers at Risk of Adverse Events, PROCEEDINGS OF THE 22ND ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, KDD '16.

[61] Mixing the barrel of bad apples may reduce the number spoiled if the number of apples infected over a given period of time by an apple (and its neighbors) in its new location is fewer than the number of apples that would have been infected by the apple (and its neighbors) in the old location over a given period of time.

[62] See Bryan Wilder, Laura Onasch-Vera, Juliana Hudson, Jose Luna, Nicole Wilson, Robin Petering, Darlene Woo, Milind Tambe, Eric Rice, End-to-end Influence Maximization in the Field (working paper 2017).

Conclusion

We model the spread of police shootings as a dynamic process of contagion. Drawing from data on police networks and police shootings in Chicago, we generate evidence consistent with contagion. Officers engaged in shooting are more connected with each other (avg. degree 20) on police networks than randomly chosen officers (avg. degree 12). The period of time between shootings is significantly shorter (739 days) for shooting officers who are connected on a network than for randomly chosen shooting officers who are not necessarily connected (994 days). Data analysis allows us to estimate that each police shooting gives rise to an additional 0.5 shootings over eight years of study. We find that contagion contributes to 29% of police-involved shootings in our dataset. Most significantly, we find that exposure to an earlier shooting doubles the likelihood of a future shooting after two years. Our research constitutes compelling preliminary evidence of contagion, and suggests avenues of violence prevention that involve disrupting or combating contagion among officers.

APPENDIX

1.0 Omitted Variables—The Sunday Shooter Syndrome

A fanciful hypothetical example which helps illustrate potentially unobservable confounding is the following. Assume that there is a personality disorder, the "Sunday and Wednesday Shooter Syndrome"'(SWSS), which causes people who suffer from it to have an irrepressible desire to fire guns on Wednesdays if there was a shooting on a Sunday, while being absolutely repulsed by the idea the rest of the week. Further assume that people who suffer from SWSS tend to get along well together. If there were officers suffering from SWSS in our dataset, then the shootings from these officers would appear close together in time (since they would all be on Sundays) while also committed by officers connected by edges (since they suffer from SWSS, they are more likely to be connected to each other). This would lead us to believe that there is contagion at play, while the true explanation is homophily, in the form of the SWSS disorder. Because this homophily trait does not work through increasing the frequency of shooting, it would not be controlled for by "number of shootings" randomization in our analysis.

2.0 Fitting the Hawkes Model to the Observed Data

Formally, let G = (V, E) be the graph of officer-involved shooting events $\{(v_i, t_i)\}$ $1 \leq i \leq n$, where $v_i \in V$ (vertices) represents the identity (as specified by both location and officer involved) of the shooting event, and $t_i \in \{0, T\}$ is the time of the shooting located in the relevant period of study described by $\{0,T\}$.[63] Based on previous studies of violence in social networks, we assume that infections are able to be transmitted across a network distance of up to three degrees of separation; people who are further away in the network have no direct effect on one another.

We can define the instantaneous probability of the officer-involved shooting for each event on the network $v \in V$ using the following intensity (instantaneous probability) function derived from the Hawkes Model:

$$\lambda_v(t) = \lambda_0(t) + \alpha\beta \sum_{i:(v_i,v \in E)} e^{\beta(t_i - t)} \text{ [64]} \tag{1}$$

Recall from the earlier description of this equation that λ_0 is the background rate, α is the social influence rate, and β is the rate at which the social influence decays over time. We can estimate the value of these three parameters—the background rate, the influence rate, and the rate of influence decay--by way of maximum likelihood estimation: searching for a set of parameter values on these three dimensions that maximizes the log-likelihood of the observed data.

---

[63] We shift the origin of time so that [0,T] represents the study period. This period is defined by the interval of time between the first complaint and last complaint in our dataset of complaints.

[64] It is important that our Hawkes process be well-defined such that $\alpha \int_0^\infty \phi(t)dt < 1$. If not, the process will diverge and there can be infinitely many events occurring at all time *t*. For this reason, we normalize the kernel function so that its integral is 1 and we add the constraint $\alpha \in [0,1]$. We include parameters alpha and beta in the historical kernel for purposes of simplifying the calculation—doing so does not affect our results.

We refer the reader to other sources[65] for a formal discussion of the conditional intensity function and its proper interpretation in a Hawkes process. From these we derive the following formula for the log-likelihood function. The log-likelihood function describes the way in which the log-likelihood changes in response to variation in parameter values. Drawing from Equation 1, we can derive the log-likelihood function of observing the shooting events recorded by the data, given parameters λ_0, α and β (background rate, influence rate, and rate of influence decay):

$$\mathcal{L}(\lambda_0, \alpha, \beta) = \sum_{i=1}^{N} \log \lambda_v(t_i) - \sum_{v \in V} \int_0^T \lambda_v(t) dt$$

The first sum calculates the log-likelihood of every infection event that *did* occur, and the second sum calculates the log-likelihood that each individual was *not* infected at all other times.

With our choice of intensity, our log-likelihood function becomes:

$$\mathcal{L}(\lambda_0, \alpha, \beta) = \sum_{i=1}^{N} \log \left(\lambda_0 + \alpha \beta \sum_{\substack{j:(v_j, v_i) \in E \\ t_j < t_i}} e^{-\beta(t_i - t_j)} \right) - T|V|\lambda_0 - \alpha \sum_{v \in V} \sum_{i:(v_i, v) \in E} (1 - e^{-\beta(T - t_i)}) \tag{2}$$

The log-likelihood function is the objective function that informs maximum likelihood estimation. Finding the maximum likelihood estimate of these parameters amounts to solving the following optimization problem:

$$\underset{\alpha, \beta, \lambda 0}{\arg\max} \, \mathcal{L}(\alpha, \beta, \lambda_0) \tag{3}$$

We observed numerically that Equation 3 has many local optima; hence we solve Equation (3) with a combination of techniques.

The search for optimal parameters in parameter-space takes place over a multi-dimensional grid-space that consists of all possible combinations of our three parameters. Brute force grid search is a method in which the researcher incrementally and exhaustively searches the full set of parameter combinations (of three parameters) over a relatively coarse grid of parameter space.

Another search option is "gradient descent," in which the gradient of the log-likelihood function is computed to determine the most efficient local search path—the path of "steepest descent"—to locate the minimum negative log-likelihood. To compute the gradient, we take the derivative of the log-likelihood function with respect to each of the parameters. Iterative step-wise search proceeds along the gradient until the gradient becomes flat, which indicates that log-likelihood of observed data cannot be decreased further.

---

[65] Emmanuel Bacry, Iacopo Masstromateo, and Jean-Francois Muazy, Hawkes Processes in Finance, 1 Microstructure and Liquidity 1 (2015); Daryl J Daley and David Vere-Jones, AN INTRODUCTION TO THE THEORY OF POINT PROCESSES: Vol. II: General Theory And Structure (Springer Science & Business Media 2007) .

We approach parameter estimation in the following way:

1. First, we do a brute force grid search in intervals of .01 to locate parameters that maximize the log-likelihood of the observed data.

2. Second, we optimize over $\{\alpha, \lambda_0, \beta\}$ using a computational search algorithm, Sequential Least Squares Programming (SLSQP), initializing from 100 randomly chosen initial parameter values.

3. Third, we combine the first two. Starting from the 10 best points obtained during the first step, we optimize over $\{\alpha, \lambda_0, \beta\}$ using a search algorithm, Sequential Least Squares Programming (SLSQP).

The SLSQP algorithm is a type of gradient descent that helps us search the parameter space efficiently. All algorithms and grid searches were run until convergence, and all converged to the same values.[66] Searches were also run using a number of alternative algorithmic formulations until parameters converged and all produced the same results.

3.0   Results

   3.1   Definitions of "background shooting" and "contagion shooting"

      a. The probability of a background shooting is defined as the probability that shooting n is influenced only by shooting n:

$$p_{n,n} = \frac{\lambda_0}{\lambda_0 + \alpha\beta \sum_{n' \; : \; t_{n'} \leq t_n} e^{-\beta(t_n - t_{n'})} 1(o_m \sim o_n)}$$

      Below is the distribution of ALL shootings from observed data, with their associated probabilities that they are background shootings. Approximately 350 shootings were associated with a probability of 1 as background shootings.
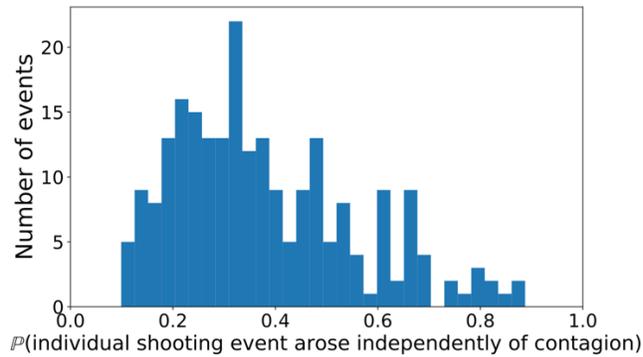

Fig. A1 Distribution of Background Shootings

---

[66] Gradient descent can be used to locate both minima and maxima, the latter by minimizing the negative log-likelihood of observed data. We use both gradient descent and second-order curvature exploration in the SLSQP algorithm.

b. The probability of a triggered event is defined as the probability that shooting n is influenced by any previous shooting (denoted by n prime):

$$\bar{p}_n = \sum_{n' : t_{n'} < t_n} \frac{\alpha\beta e^{-\beta(t_n - t_{n'})} 1(o_{n'} \sim o_n)}{\lambda_0 + \alpha\beta \sum_{m \,:\, t_m \leq t_n} e^{-\beta(t_n - t_m)} 1(o_m \sim o_n)}$$

Below is the distribution of all shootings from observed data with their associated probabilities that they are influenced by contagion. Approximately 300 shootings were associated with insignificant contagion probabilities.
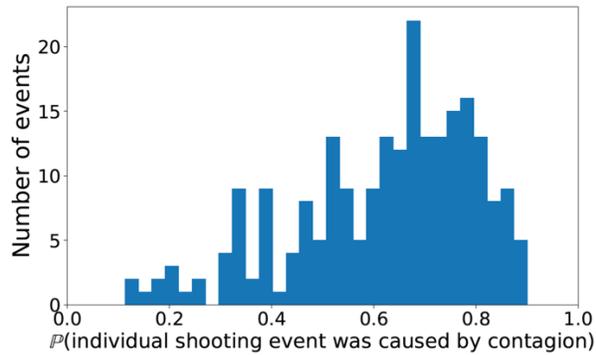
Fig. A2 Distribution of Contagion Shootings

3.2   Parameter values obtained via grid search, gradient descent and combination:

λ_0 (background) = 3.89 X 10^{-2} α (social influence) = 1.9 X 10^{-2} β = 4.388

3.21 Background rate

a.   Background rate-grid search
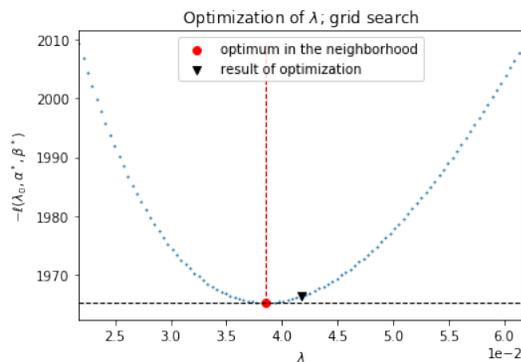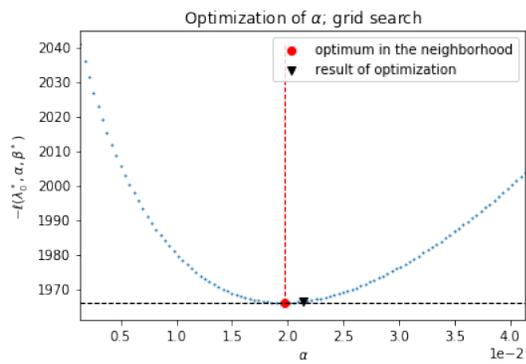
b. Background rate--gradient descent (SLSQP)

Optimization of λ using SLSQP (best of 100 random inits)
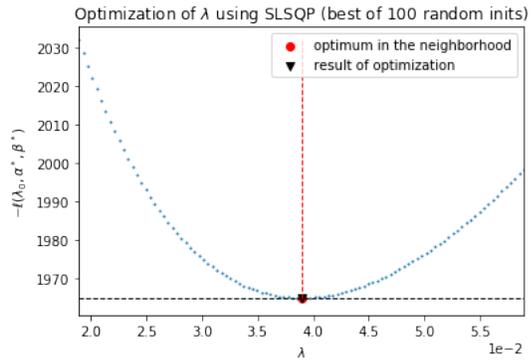
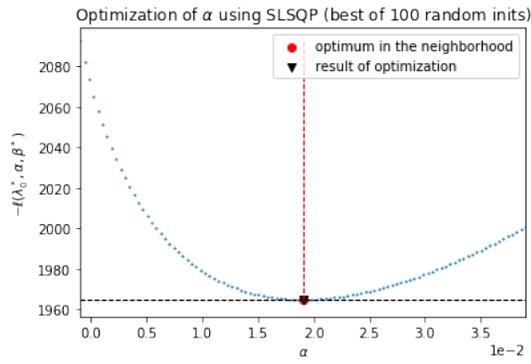c. Background rate—combination grid search and gradient descent (warm start using 10 best points from grid search)
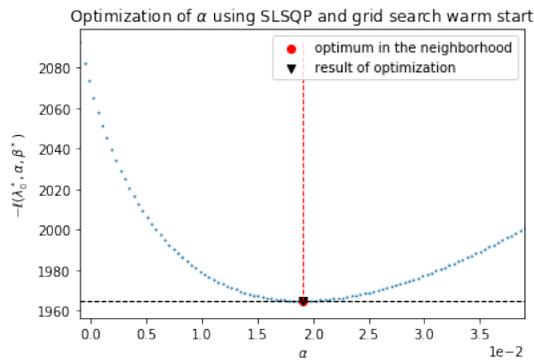

Optimization of λ using SLSQP and grid search warm start
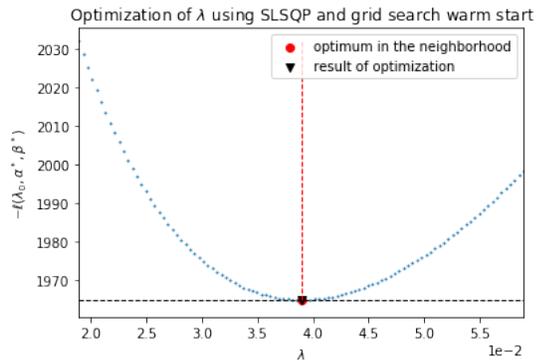
3.22 Social influence

a. Social influence-grid search

Optimization of α; grid search

b. Social influence-gradient descent (SLSQP)

Optimization of α using SLSQP (best of 100 random inits)

c. Social influence-combination grid search and gradient descent (warm start)

Optimization of α using SLSQP and grid search warm start

3.23 Decay rate

a. Decay rate-grid search

Optimization of β; grid search
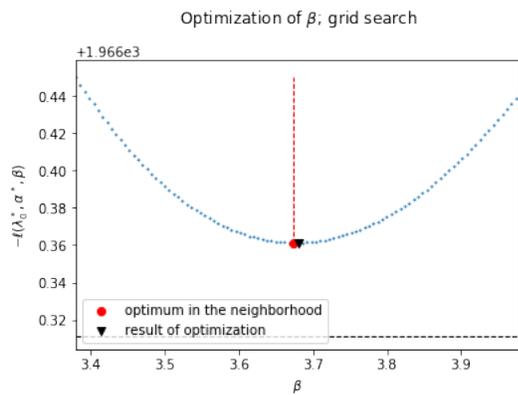
b. Decay rate-gradient descent (SLSQP)

Optimization of β using SLSQP (best of 100 random inits)

+1.964e3
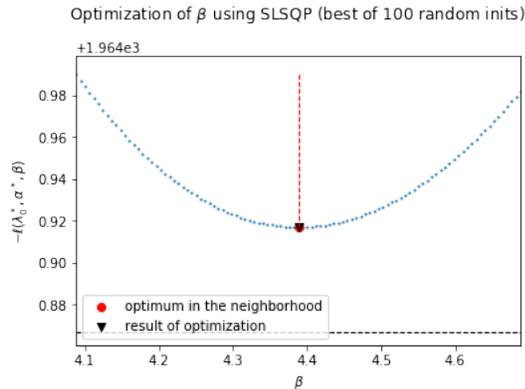
c. Decay rate-combination grid search and gradient descent (warm start)

Optimization of β using SLSQP and grid search warm start

+1.964e3
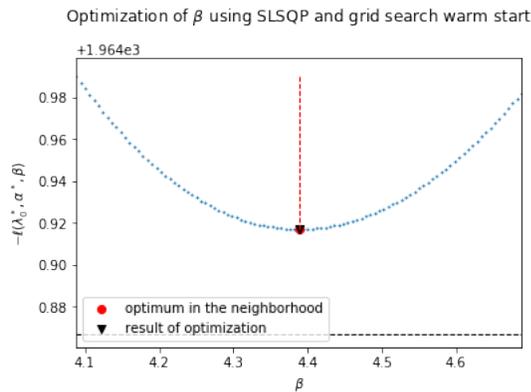
4. Calculating Interpretable Statistics

4.1 Expected Number of Offspring Shootings

Let K∈{0,1}I×I be the binary matrix of ``friendships'' in the police network.

Let s(t+1)∈ R_+^I be a vector-valued random variable of expected number of t-th generation shootings at each node in the network.

We have:

$$s(t+1)=\alpha K s(t-1)$$

and therefore, letting $s^{(0)}i = \frac{1}{I} X \, \mathbf{1}^T$ where 1 is a length-I vector of ones,

$$\mathbb{E}[\text{total number of shootings is the network associated with first}]= \mathbf{1}^T \sum_{t\geq0}(\alpha K)^t s_i^{(0)} = \frac{1}{I}\mathbf{1}^T (\mathcal{I}\mathrm{I} - \alpha K)^{-1}\mathbf{1}$$

$$=1.496986 \text{ offspring shootings (including first shooting)}$$

4.2 Expected Number of Shootings Influenced by Contagion

We calculate the expected number of shootings influenced by earlier shootings:

$$\mathbb{E}[\text{shootings associated with contagion}|\ N \text{ total shootings}] = \ N - \sum_n \mathbb{P}(n - \text{th shooting is from background})$$

$$= 141 \text{ shootings}$$