Review of “Does Violence Interruption Work”

July 18, 2017

There is a lot to like in this paper. The problem of gang violence is real and in a city like LA, a significant public safety issue. Some other large cities like Chicago share this problem. Spinoffs from the original CEASEFIRE intervention have been tried in a number of locales with at best mixed results. Among the many problems is that the intervention is a full court press involving many different agents and organizations so that it is virtually impossible to know what features of the intervention matter and to replicate the intervention in different settings. In effect, there are many variants on the CEASEFIRE idea that in practice are not likely to be comparable. So, having another test in a large city like LA is in principle good.

The authors are also quite careful in how they describe the study and clearly have put in a lot of work. The model is interesting as a statistical formulation. Finally, if the findings can be believed, they are large enough to make a practical difference. Too often these kinds of study will show effects that are far too small to matter, even if they jump the .05 hurdle.

But at the very least, the authors need to be far more circumspect about the theory, the data, the intervention and the research design. And I fear that if they do that, some very serious problems will become apparent. I come very away confused about what has been demonstrated. But perhaps the paper can be saved with a major reframing

1. The binary events being modeled are never really defined. Are we talking about crime events, victims or perpetrators? One shooting can even involve several victims and several perpetrators. One crime can often lead to several charges, all of which can involve violence. This is worrisome because these different definitions imply the need for different models and a consideration of different kinds and amounts of measurement error. It also goes to what the actual experimental units really are (more on that shortly).

2. We also need a definition of a gang or at least how the LAPD defines a gang and determines when crimes are gang-related and, presumably retaliatory. For example, if several bystanders are accidentally shot in a violent incident, is that gang related and retaliatory? I appreciate that a key is what gets reported to the police as gang-related, but presumably the police investigate and apply their own definitions later.

3. As already noted, it is not at all clear what the interventions is. This is vital because (a) one must know the content of any intervention if it is to be replicated; (b) one must know the content of any intervention to know how to theorize about it; and (c) one must know the content of any intervention to distinguish between components of the intervention and confounders that are not components of the intervention. All of these concerns are troubling in this analysis. For example, are summer jobs, or recreational programs, or CBT, or drug testing, or restorative justice included? What about the role of probation or parole officers for those currently under supervision? What about the confiscation of firearms? I could go on for some time because there are decades of gang interventions. What *is* included here?

4. There is a mismatch between the theory proposed and how the study is conducted. I read this paper (and others like it) claiming that the intervention is delivered to *individuals* who have reason to retaliate. But we have no measures on what is delivered to given individuals nor whether or not these individuals were engaged in subsequent crime. So, this analysis seems guilty of the classic ecological fallacy. The intervention is directed at individuals, and the outcome is retaliatory crimes of violence committed (or not) by those individuals. But the intervention and outcome are measured *in time and space locations*. Statisticians and social scientists have been cognizant of the ecological fallacy for decades. The error is clear, and the authors need to argue that in this instance, there are no important conceptual or statistical consequences. I think that will be a hard sell.

5. These issues carry over to the model and analysis. The retaliatory "causal" link is theorized and modeled as stronger when two crimes (or is it individuals or perpetrators?) are closer and time and space. The link gets exponentially weaker with greater distance in either. But, there are no measures of the actual links between crimes. We don't know in fact if a later crime is retaliation for a given earlier crime. We don't even know if the perpetrator of the later crime knew

about the earlier crime or was motivated by revenge. It's all ecological. Arguably, all the model picking up is that a violent gang-related crime is more likely to occur when it is closer in time and space to an earlier violent gang-relate crime. This could result from factors having little or nothing to do with revenge. For example, the violence might be an economic decision for control of the local drug market. Or earlier violence might help to legitimate later violence in the same neighborhood; it is OK to pull a gun in a dispute and use it. Or it may be that prospective offenders learn that that chances of apprehension are very small. (What are the clearance rates for homicides for the LAPD?) You *can* get away with murder. Nowhere is crime motive actually measured.

6. The experimental design violates the assumption of no interference (aka Don Rubin's stable unit treatment value assumption). The gang members interact, and those who work with them do as well. So there cannot be a single treatment effect, but many (formally almost limitless). Each possible random assignment implies another possible treatment effect. To take an extreme example, two members of the same gang can be assigned to the treatment or the control condition or to different ones. And that can matter insofar as they talk to one another.

7. The statistical tests done to evaluate the study design do not solve these kinds of problems. For example, the K-S test comparing the experimental and control distributions does not address the role confounders (in part because apparently none are defined or measured).

8. And finally a minor point. I had a close look at the likelihood function in the appendix which was made difficult because some of the notion is not defined or explained.

In summary, we have an ecological study design to estimate the impact of a poorly defined intervention with individuals. As an empirical matter, all we have is associations conditional on location is time and space. The theory and the model are really just overlays that provide no real insight even though the model apparently fits the data pretty well overall. Put another way, one could motivate the model using another story and there could be other models that fit as well or better. Put still another way, we have a genuinely interesting model fitting exercise (for those who care about such things) but no real scientific or policy payoff. And I fear that too many readers will think that the good fit justifies the substantive conclusions offered. It does not. Finally, we don't know what the intervention was nor its dose, either in the aggregate or individual by individual.

Can the paper be saved? Perhaps. Here's a possible story line. There is a serious policy problem that has been very difficult to solve. LA has given it a try using the CEASEFIRE approach. As an empirical matter, the strength of the associations between proximity in time and space of gang-related different violence crimes is weakened after the intervention is introduced. But we really don't know much about why for all the reasons provided above. The so-called model is really just an estimator. It is not some theory of how violent, gang-related crimes come about. The fit is pretty good, which can suggest that the estimates may have good statistical properties. There are many plausible explanations for these estimates but this study cannot sort them out.

The authors really need to back off from the claim that their estimator is a model of how the world works. It is an ecological model at best hat is far removed from the theory applied and the description of the intervention.

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2017 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ____________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ____________________
Emily W. Loughran, Sr. Director of Licensing
Date: 6·2·17

By: ____________________
Jason Xu, RIPS Academic Mentor
Date: ____________________

By: ____________________
Collin Cademartori, RIPS Student
Date: ____________________

By: ____________________
Xi Chen, RIPS Student
Date: ____________________

By: ____________________
Alistair Letcher, RIPS Student
Date: ____________________

By: ____________________
Jalena Trisovic, RIPS Student
Date: ____________________

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2017 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______________________
Emily Loughran, Director of Licensing
Date: 6.2.17

By: ______________________
Matt Habeland NSF REU Academic Mentor
Date: ______________________

By: ______________________
Hao Li, NSF REU Academic Mentor
Date: ______________________

By: ______________________
Osman Akar, NSF REU Student
Date: ______________________

By: ______________________
Adam Lemuel Dhillon, NSF REU Student
Date: ______________________

By: ______________________
Honglin Chen, NSF REU Student
Date: ______________________

By: ______________________
Alexander Insuk Song, NSF REU Student
Date: ______________________

By: ______________________________
Honglin Chen, NSF REU Student
Date: ____________________________

By: ______________________________
Tiankuang Zhou, NSF REU Student
Date: ____________________________

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______________________
Emily Loughran, Director of Licensing
Date: ______________________

By: ______________________
Matt Habeland NSF REU Academic Mentor
Date: ______________________

By: ______________________
Hao Li, NSF REU Academic Mentor
Date: ______________________

By: ______________________
Osman Akar, NSF REU Student
Date: ______________________

By: ______________________
Adam Lemuel Dhillon, NSF REU Student
Date: ______________________

By: ______________________
Honglin Chen, NSF REU Student
Date: ______________________

By: ______________________
Alexander Insuk Song, NSF REU Student
Date: ______________________

By: ______________________________

Honglin Chen, NSF REU Student

Date: ____________________________

By: ______________________________

Tiankuang Zhou, NSF REU Student

Date: ____________________________

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______________________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ______________________________
Emily Loughran, Director of Licensing
Date: ______________________________

By: ______________________________
Jason Xu, RIPS Academic Mentor
Date: ______________________________

By: ______________________________
Collin Cademartori, RIPS Student
Date: ______________________________

By: ______________________________
Xi Chen, RIPS Student
Date: ______________________________

By: ______________________________
Alistair Letcher, RIPS Student
Date: ______________________________

By: ______________________________
Jalena Trisovic, RIPS Student
Date: ______________________________

	Date	LastName	FirstName	DOB	Desc	Address	RD
1	1/4/2017	boyer	micah	[redacted]	M/W	[redacted]	1415
2	1/4/2017	MORRIS	DARRYL	[redacted]	M/B	[redacted]	1309
3	1/5/2017	WASHINGTON	JEFFREY	[redacted]	M/W	Ellis Ave & S Robertson Blvd, Los Angeles 90034	1409
4	1/5/2017	SMART	DESTIN	[redacted]	M/B	Ellis Ave & S Robertson Blvd, Los Angeles 90034	1409
5	1/5/2017	ZORN	SANDRA	[redacted]	F/H	3RD AV&ROSE AV	1413
6	1/5/2017	DAMBERT	DAVID	[redacted]	M/W	3RD AV&ROSE AV	1413
7	1/6/2017	hickman	stacy	[redacted]	F/B	[redacted]	1431
8	1/6/2017	monroe	beverly	[redacted]	F/B	Breeze Ave & Ocean Front Wk, Venice 90291	1412
9	1/6/2017	Raphael	Patrick	[redacted]	M/B	Navy St & Ocean Front Wk, Venice 90291	1412
10	1/6/2017	hohnstein	dennis	[redacted]	M/W	Sawtelle Blvd & National Blvd, Los Angeles 90064	1402
11	1/7/2017	Riley	Robert	[redacted]	M/B	Ocean Front Wk & Ozone Ave, Venice 90291	1411
12	1/7/2017	Raphael	Patrick	[redacted]	M/B	Ocean Front Wk & Ozone Ave, Venice 90291	1411
13	1/7/2017	Mcgee	Micca	[redacted]	M/W	Ocean Front Wk & Ozone Ave, Venice 90291	1411
14	1/8/2017	Davis	Christelyn	[redacted]	F/B	Paloma Ave & Ocean Front Wk, Venice 90291	1412
15	1/8/2017	flores	nelson	[redacted]	M/H	Speedway & Park Ave, Venice 90291	1412
16	1/9/2017	jones	rebecca	[redacted]	F/B	[redacted]	1487
17	1/9/2017	Townsend	Jason	[redacted]	M/W	Frederick St & Rose Ave, Venice 90291	1415
18	1/10/2017	WASHINGTON	JEFFREY	[redacted]	M/W	Ellis Ave & S Robertson Blvd, Los Angeles 90034	1409
19	1/10/2017	Silen	Marcia	[redacted]	F/W	[redacted]	1413
20	1/10/2017	Tevez	Melanie	[redacted]	F/W	[redacted]	1413
21	1/11/2017	Johnson	William	[redacted]	M/W	[redacted]	1441
22	1/11/2017	WHITTAKER	ALEXIS	[redacted]	F/B	Hampton Dr & Sunset Ave, Venice 90291	1413
23	1/11/2017	Smith	Carlos Larue	[redacted]	M/B	[redacted]	1425
24	1/12/2017	Talley	Shawn	[redacted]	M/B	Market St & Ocean Front Wk	1431
25	1/12/2017	Phegley	Christhoper	[redacted]	M/W	Washington / Pacific	1452
26	1/13/2017	Miller	Andrew	[redacted]	M/B	Ocean Front Wk & Windward Ave, Venice 90291	1431
27	1/14/2017	Fleming	John Jr	[redacted]	M/B	Ocean Front Wk & Horizon Ave, Venice 90291	1431
28	1/14/2017	Garcia	Frank	[redacted]	M/H	Ocean Front Wk & Washington Blvd, Venice 90291	1452
29	1/14/2017	Thilking	John	[redacted]	M/W	Pacific & Windward	1431
30	1/15/2017	maldonado	miguel	[redacted]	M/H	Allin St & Inglewood Blvd, Culver City 90230	1468
31	1/16/2017	JONES	STANLEY	[redacted]	M/W	3RD AV/ROSE AV	1413
32	1/17/2017	Walsh	Naomi	[redacted]	F/W	PARK/OFW	1431
33	1/18/2017	Reutz	Zoe	[redacted]	F/W	Ocean Front Wk & Brooks Ave, Venice 90291	1411
34	1/18/2017	Garcia	Nixon	[redacted]	M/H	Ocean Front Wk & Brooks Ave, Venice 90291	1411
35	1/19/2017	Baxter	Hasan	[redacted]	M/B	Venice Blvd & Globe Ave, Los Angeles 90066	1436
36	1/19/2017	Johnson	Candace	[redacted]	F/W	Venice Blvd & Globe Ave, Los Angeles 90066	1436
37	1/19/2017	Gregory	Gary	[redacted]	M/W	Venice Blvd & Globe Ave, Los Angeles 90066	1436
38	1/19/2017	Liggins	Levell	[redacted]	M/B	Rose Ave & Lincoln Blvd, Venice 90291	1415
39	1/19/2017	Winnen	Michael	[redacted]	M/W	Rose Ave & Frederick St, Venice 90291	1415
40	1/19/2017	Puig	Erica	[redacted]	F/W	Rose Ave & Frederick St, Venice 90291	1415
41	1/19/2017	Secrest	James	[redacted]	M/B	Rose Ave & Frederick St, Venice 90291	1415
42	1/19/2017	White	Raymond	[redacted]	M/B	Washington Blvd & Lincoln Blvd, Venice 90291	1444
43	1/19/2017	bloom	ariana	[redacted]	F/B	W 96th St & Hindry Ave, Los Angeles 90045	1489
44	1/20/2017	Margarito	Freddy	[redacted]	M/H	[redacted]	1407
45	1/21/2017	Sloman	Bonnie	[redacted]	F/W	3rd Ave & Rose Ave, Venice 90291	1413
46	1/21/2017	Thornton	Wendell	[redacted]	M/B	3rd Ave & Rose Ave, Venice 90291	1413
47	1/21/2017	Townsend	jason	[redacted]	M/W	Lincoln Blvd & Flower Ave, Venice 90291	1415
48	1/22/2017	molovinsky	nahmi	[redacted]	F/X	[redacted]	1412
49	1/22/2017	Foss	Jennafer	[redacted]	F/W	3rd Ave & Rose Ave, Venice 90291	1413
50	1/22/2017	Smith	Brian	[redacted]	M/W	Ocean Front Wk & Westminster Ave, Venice 90291	1431
51	1/22/2017	Hester	Daniel	[redacted]	M/B	[redacted]	1469
52	1/22/2017	YATES	DAVID	[redacted]	M/W	Oxford Ave & Dickson St, Marina Del Rey 90292	1444
53	1/24/2017	kimbrogh	brian	[redacted]	M/B	Flower Ave & Lincoln Blvd, Venice 90291	1415
54	1/24/2017	Levizon	Michael	[redacted]	M/W	W 94th St & Hindry Pl, Los Angeles 90045	1489
55	1/24/2017	USSERY	ANTHONY	[redacted]	M/W	Culver, Los Angeles 90066	1454
56	1/24/2017	Canales	Julio	[redacted]	M/H	S Centinela Ave & Sanford St, Culver City 90230\	1465
57	1/24/2017	FUNDERBURK	HEATHER	[redacted]	F/W	Culver, Los Angeles 90066	1454
58	1/24/2017	Berganholi	Tatiana	[redacted]	F/W	3rd Ave & Rose Ave, Venice 90291	1413
59	1/24/2017	burnette	brandon	[redacted]	M/W	[redacted]	1444
60	1/24/2017	schlagel	ronald	[redacted]	M/W	Hindry Ave & W 96th St, Los Angeles 90045	1489
61	1/25/2017	letki	marie	[redacted]	F/W	[redacted]	1463
62	1/25/2017	delangis	hailey	[redacted]	F/W	[redacted]	1489
63	1/26/2017	Dyer	Carry	[redacted]	M/W	Ocean Front Wk & Ozone Ave, Venice 90291	1411
64	1/26/2017	morris	amber	[redacted]	F/W	Flower Ave & Lincoln Blvd, Venice 90291	1414
65	1/27/2017	allen	elzie	[redacted]	M/B	[redacted]	1406
66	1/27/2017	dwarte	irma	[redacted]	M/H	Military Ave & Sardis Ave, Los Angeles 90064	1403
67	1/27/2017	louk	william	[redacted]	M/B	5th Ave & Sunset Ave, Venice 90291	1413
68	1/27/2017	Brown	David	[redacted]	M/B	[redacted]	9999
69	1/29/2017	PENNA	ALEXANDER	[redacted]	M/W	[redacted]	1412
70	1/31/2017	COWDIN	WYLIE	[redacted]	M/W	3RD/ROSE	1413
71	1/31/2017	Reyes	Argelio	[redacted]	M/H	[redacted]	

Journal of Quantitative Criminology

Crime Topic Modeling
--Manuscript Draft--

Manuscript Number:	
Full Title:	Crime Topic Modeling
Article Type:	Original Paper
Keywords:	Machine learning; Non-negative matrix factorization; Text mining; Crime
Corresponding Author:	P. Jeffrey Brantingham University of California Los Angeles Los Angeles, CA UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of California Los Angeles
Corresponding Author's Secondary Institution:	
First Author:	Da Kuang
First Author Secondary Information:	
Order of Authors:	Da Kuang
	P. Jeffrey Brantingham
	Andrea L. Bertozzi
Order of Authors Secondary Information:	
Funding Information:	Army Research Office (ARO MURI grant W911NF-11-1-0332) — Dr. P. Jeffrey Brantingham
	National Science Foundation (NSF grant DMS-1417674) — Dr Andrea L. Bertozzi
	Office of Naval Research (ONR grant N00014-16-1-2119) — Dr Andrea L. Bertozzi

Abstract:

Objectives The classification of crime into discrete categories entails a massive loss of information. Crimes emerge out of a complex mix of behaviors and situations, yet most of these details cannot be captured by singular crime type labels. This information loss impacts our ability to not only understand the causes of crime, but also how to develop optimal crime prevention strategies.

Methods We apply machine learning methods to short narrative text descriptions accompanying crime records with the goal of discovering ecologically more meaningful latent crime classes. We term these latent classes 'crime topics' in reference to text-based topic modeling methods that produce them. We use topic distributions to measure clustering among formally recognized crime types.

Results Crime topics replicate broad distinctions between violent and property crime, but also reveal nuances linked to target characteristics, situational conditions and the tools and methods of attack. Formal crime types are not discrete in topic space. Rather, crime types are distributed across a range of crime topics. Similarly, individual crime topics are distributed across a range of formal crime types. Key ecological groups include identity theft, shoplifting, burglary and theft, car crimes and vandalism, criminal threats and confidence crimes, and violent crimes.

Conclusions Crime topic modeling positions behavioral situations as the focal unit of analysis for crime events. Though unlikely to replace formal legal crime classifications, crime topics provide a unique window into the heterogeneous causal processes underlying crime. We discuss whether automated procedures could be used to cross-check the quality of official crime classifications.

Crime Topic Modeling

Da Kuang
Department of Mathematics
University of California Los Angeles
520 Portola Plaza
Los Angeles, CA 90095-1555
dakuang@math.ucla.edu

P. Jeffrey Brantingham*
Department of Anthropology
University of California Los Angeles
341 Haines Hall
Los Angeles, CA 90095-1553
branting@g.ucla.edu
Tel: 310-267-5251

Andrea L. Bertozzi, UCLA Mathematics
Department of Mathematics
University of California Los Angeles
520 Portola Plaza
Los Angeles, CA 90095-1555
bertozzi@math.ucla.edu

* Corresponding Author

Running Head: Crime Topic Modeling

Crime Topic Modeling

Abstract

Objectives The classification of crime into discrete categories entails a massive loss of information. Crimes emerge out of a complex mix of behaviors and situations, yet most of these details cannot be captured by singular crime type labels. This information loss impacts our ability to not only understand the causes of crime, but also how to develop optimal crime prevention strategies.

Methods We apply machine learning methods to short narrative text descriptions accompanying crime records with the goal of discovering ecologically more meaningful latent crime classes. We term these latent classes 'crime topics' in reference to text-based topic modeling methods that produce them. We use topic distributions to measure clustering among formally recognized crime types.

Results Crime topics replicate broad distinctions between violent and property crime, but also reveal nuances linked to target characteristics, situational conditions and the tools and methods of attack. Formal crime types are not discrete in topic space. Rather, crime types are distributed across a range of crime topics. Similarly, individual crime topics are distributed across a range of formal crime types. Key ecological groups include identity theft, shoplifting, burglary and theft, car crimes and vandalism, criminal threats and confidence crimes, and violent crimes.

Conclusions Crime topic modeling positions behavioral situations as the focal unit of analysis for crime events. Though unlikely to replace formal legal crime classifications, crime topics provide a unique window into the heterogeneous causal processes underlying crime. We

discuss whether automated procedures could be used to cross-check the quality of official crime classifications.

Keywords: Machine learning; Non-negative matrix factorization; Text mining; Crime.

1. Introduction

Upon close inspection, most criminal events arise from subtle interactions between situational conditions, behavioral routines, and the boundedly-rational decisions of offenders and victims (Brantingham and Brantingham 1993). Consider two crimes. In one event, an adult male enters a convenience store alone in the middle of the night. Brandishing a firearm, he compels the store attendant to hand over liquor and all the cash in the register (Wright and Decker 1997:89). This event may be contrasted with a second involving female sex worker who lures a *john* into a secluded location and takes his money at knife point, literally catching him with his pants down (Wright and Decker 1997:68). In spite of the fine-grained differences between these events, both end up classified as armed robberies. As a matter of law, the classification makes perfect sense. The law favors a bright line to facilitate classification of behavior into that which is criminal and that which is not (Casey and Niblett 2015; Glaeser and Shleifer 2002). The loss of information that comes with condensing complex events into singular categories, however, may severely hamper our ability to understand the immediate causes of crime and what might be done to prevent them, though the quantitative tractability gained may certainly offset some of the costs.

The present paper explores novel methods for crime classification based directly on textual descriptions of crime events. Specifically, we borrow methods from text mining and machine learning to examine whether crime events can be classified using text-based latent topic modeling (e.g., Blei 2012). Our approach hinges on the idea that criminal events are composed of mixtures of behavioral and situational conditions that are captured at least partially in textual descriptions of those events. Over a corpus of events, the relative frequency of situational and behavioral conditions is captured by the relative frequency of different words in the text-based descriptions of those events. Topic modeling of the text then allows one to infer the latent behavioral and situational conditions driving those events.

Latent topic modeling offers two unique advantages over standard classification systems. First, latent topic models potentially allow higher-level class structures to emerge autonomously from lower-level data, rather than being imposed *a priori*. Simpler or more complex class structures, relative to the formal system in place, may be one result of autonomous classification. Such emergent classifications may also be ecologically more meaningful. Second, latent topic models allow for soft clustering of events. Common crime classification systems require so-called hard clustering into discrete categories. A crime either is, or is not a robbery. Soft-clustering, by contrast, allows for events to be conceived of as mixtures of different latent components, revealing nuanced connections between behaviors, settings and crime. An event that might traditionally be considered a robbery, for example, may actually be found to be better described as a mixture of robbery and assault characteristics.

The remainder of this paper is structured as follows. Section 1 reviews several long-standing issues surrounding crime classification and causal inference. Section 2 introduces text-based latent topic modeling at a conceptual level. This forms a basis for describing how the models may be applied to the problem of crime classification. Section 3 presents methodological details underlying non-negative matrix factorization as a method for topic modeling (Lee and

Seung 1999). Here we also introduce methods for evaluating topic model classifications using the official classifications as a benchmark. In this context we can measure the distance between different classifications in terms of their underlying topic structure. Section 4 introduces the empirical case and data analysis plan. We analyze all crimes occurring in the City of Los Angeles between Jan 1, 2009 and July 19, 2014 using data provided by the Los Angeles Police Department (LAPD). Section 5 presents results. The paper closes with a discussion of the implications of this work and future research directions.

2. Causal Heterogeneity and Crime Classification

Our starting premise is that situational conditions drive crime events. This is a well-established position linked to both situational crime prevention (Clarke 1983; Clarke 1980) and crime pattern theory (Brantingham and Brantingham 1978, 1984). Situational crime prevention sees offenders as making boundedly-rational choices in response to situational cues about the quality of crime opportunities (Clarke and Cornish 1985). Crime pattern theory goes further to argue that offender decision making, if it produces successful crimes, quickly evolves into behavioral templates or scripts that are triggered with little rational thought at the time of offending (Brantingham and Brantingham 1978; Brantingham and Brantingham 1993). Crime cues are typically localized to relatively small geographic places (Groff, Weisburd, and Yang 2010; Weisburd, Groff, and Yang 2012), but may be variably fleeting or stationary in time (see Belk 1975). Overall, situational crime prevention and crime pattern theory posit a close match between situational conditions and the behavioral repertoires underlying different types of crimes.

While the above perspectives offer a comprehensive theory of situational crime causation, the formal process of crime classification makes it difficult to operationalize in practice. Most if not all situational information is discarded in applying crime type labels to events, leaving behind a bare minimum of behavioral information sufficient to satisfy to narrow legal criteria (but seeBrantingham 2016; Brennan 1987). For example, the California Penal Code defines robbery as “the felonious taking of personal property in the possession of another, from his person or immediate presence, and against his will, accomplished by means of force or fear” (CA PEN § 211). This definition provides few constraints on what property is involved, why that property was seen as a suitable target, what constitutes possession by the victim, or how force or fear was deployed. And these gaps in information concern only the most immediate situational conditions surrounding the criminal act itself.

One recourse for filling the gap in situational information about crime is to emphasize detailed observational or ethnographic studies of offending (e.g., Wright and Decker 1994; Wright and Decker 1997). Rich ethnographic observations provide convincing detail linking situational conditions to crime. However, sampling constraints necessarily limit how statistically representative such studies can ever be. Alternatively, experimental studies can seek to test how offenders make decisions in response to controlled manipulation of environmental cues (Keizer, Lindenberg, and Steg 2008; Wright, Logie, and Decker 1995). The ecological validity of such studies may be questioned depending upon how artificial experimental tasks are.

A majority of studies adopt a third approach emphasizing spatio-temporal patterns of specific crime types in relation to independent measures of crime situations (see Clarke 1980: 139). The regular covariation between specific crime types and measured situational conditions is taken as evidence of a causal process. An advantage of this distributional approach is that sample sizes may be large enough to be representative of the behavioral situations surrounding crime events for a full spectrum of crime types, though there will always be conditions that go unmeasured given the complexity of real-world environments (Brantingham 2016).

Less often appreciated is the fundamental impact that formal classification has on causal inference, though recognition of these challenges is not new (Brennan 1987; Gibbs 1960; Sellin 1938). In an ideal world, crime types would be defined such that all events in a crime type category share a common cause. In other words, ideal crime types are causally homogeneous (Brennan 1987). Both forward prediction and backward inference are straightforward under such circumstances. With causal homogeneity, observation of a situational condition, even if it is done independently of the event itself, makes it easy to predict the corresponding type of crime. Conversely, observing a particular type of crime makes it easy to infer the situational conditions that must have been present at the time the offense was committed.

Most formally recognized crime types are not causally homogeneous, but causally heterogeneous (Brennan 1987). This heterogeneity is not simply a result of classification error where events of one type are incorrectly assigned another type and thereby erroneously mix causes (Gove, Hughes, and Geerken 1985; Maltz and Targonski 2002; Nolan, Haas, and Napier 2011). Rather crime itself arises under an array of overlapping situational conditions. Formal classification only makes crime seem more causally homogenous than it actually is.

The consequences of this apparent homogeneity are profound. Forward and backward causal models are difficult to apply without error. If the relationship between situational conditions and formal crime types is one-to-many, then forward prediction is compromised. Having observed a singular situational condition, many different crime types might be predicted to occur. If the relationship between situational conditions and crime types is many-to-one, then backwards inference is compromised. Having observed a specific crime type, many different situational conditions might be causally responsible for the event. Alas, in real-world settings, the relationship between formal crime types and situational conditions is likely many-to-many meaning that both forwards and backwards causal models are compromised. Mapping formal crime types in relation to larger and larger lists of independently measured situational conditions is unlikely to rectify the problem since causal heterogeneity is an unavoidable byproduct of typological system itself. Indeed, one wonders whether the inability of criminology to make much progress in explaining crime has as much to do with the imperfections in crime typology as the failures of theory (Gibbs 1960:322-323; Weisburd and Piquero 2008). What is needed is an approach to crime classification that allows simultaneous scoring of multiple behavioral and situational conditions (Brennan 1987: 215).

While the broader theoretical challenges here are significant, a more immediate problem concerns how to garner such behavioral and situational information to facilitate the construction of situational crime types. As discussed above, ethnographic methods cannot scale sufficiently to provide a statistically representative picture for crime in general. Mapping official crime types with respect to independent situational measures may simply perpetuate the effects of causally heterogeneous formal crime type categories. Here we turn to novel methods from computational linguistics and apply them to textual narratives associated with crime events. These methods allow crime classifications to emerge naturally from situational information associated with individual crime events. The approach positions the situation as the unit of analysis. It allows crime events to be viewed as overlapping mixtures of situations. The heterogeneous causal connections between different crime types therefore can be more directly measured.

3. Latent Topic Modeling for Text Analysis

Latent topic modeling is a core feature of contemporary computational linguistics and natural language processing. It is a dominant analytical technique deployed in the study of social media (Blei 2012; Hong and Davison 2010). The conceptual motivation for topic modeling is quite straightforward. Consider a collection of Tweets[1]. Each Tweet is a bounded collection of words (and potentially other symbols) published by a user. In computational linguistics, a Tweet is called a document and a collection of Tweets a corpus. When viewed at the scale of the corpus we might imagine that there are numerous conversations about a range of topics both concrete (e.g., political events) and abstract (e.g., the meaning of life). That these topics motivate the social media posts might not be immediately obvious when examining any one individual Tweet. But viewed as a whole corpus the dimensions and boundaries of the topics may be resolvable. Section 3 will introduce the mathematical architecture for how topics are discovered from a corpus of documents. The key point to highlight here is that each topic is defined by a mixture of different words. Each document is therefore potentially a mixture of different topics by virtue of the words present in that document.

We make a conceptual connection between text-based activity and crime at two levels. The more abstract connection envisions individual crimes as the analog of documents. A collection of crimes such as all reported crimes in a jurisdiction during one week is therefore the analog of the documents in a corpus. We might imagine that the environment consists of a range of complex behavioral and situational factors, some very local and others global, which co-occur in ways that generate different types of crimes. These co-occurring factors are the analogs of different topics. How ‘crime topics’ actually generate crime might not be immediately obvious

[1] A Tweet is a discrete text-based post on the social media website Twitter.

when examining any one crime. But when crimes are aggregated into a lager collection, the dimensions and boundaries of ‘crime topics’ might be discernable. Likewise, the specific combination of behavioral and situational factors involved in a single crime are the analog of words in a document. The key point to emphasize here is that ‘crime topics’ are mixtures of behaviors and situations. Each crime is therefore a mixture of ‘crime topics’ by virtue of the situations and behaviors present at the time of the crime.

The more concrete connection appeals directly to text-based descriptions of crimes as a source of information about the complex environmental backcloth of crime (Brantingham and Brantingham 1993). Specifically, we treat text-based descriptions of crime compiled by reporting police officers as a record of some fraction of the behavioral and situational factors deemed most relevant to that crime. As a result, we seek to apply topic modeling directly to the text-based descriptions of crime accompanying many crime records.

4. Methods

The goal of the current section is to describe methods for building latent topic models using text-based descriptions of crimes. We take a linear algebraic approach due to its computational efficiency and scalability to massive data sets, for example the text descriptions of nearly one million crimes discussed in Section 4. The linear algebraic approach contrasts with probabilistic methods such as the popular latent Dirichlet allocation (LDA) (Blei, Ng, and Jordan 2003), which is computationally expensive. Our approach does not yield a probabilistic interpretation and rigorously should be called a “document clustering” method. Recent research, however, has built connections between linear algebraic and probabilistic methods for topic modeling (Arora et al. 2013), supporting the usefulness of linear algebraic methods as an efficient way to compute topic models.

4.1.1. Text Preprocessing

Text-based narratives are typically very noisy, including typos and many forms of abbreviation of a same word. To obtain reliable results that are less sensitive to noise, we run a few preprocessing steps on the raw text accompanying crime events including removal of so-called stop-words (see e.g., Rajman and Besançon 1998). Stop-words refer to the most common words in a language, which can be expected to be present in a great many documents regardless of their content or subject matter. We augment a standard list of stop-words (e.g. a, the, this, her, …) with all the variations of the words "suspect" and "victim", since these two words are almost universally present in all descriptions of crime and do not provide useful contextual information (though they would be useful for other studies). The linguistic variations include all the prefixes such as "S", "SUSP", "VIC" and anything followed by a number (e.g. "V1", "V2"). All the stop-words are then discarded. We also discard any term appearing less than 5 times in the entire corpus. Finally, any document containing less than 3 words in total is discarded. This procedure runs in an iterative manner until no more terms or documents can be discarded.

4.1.2. Term Frequency-Inverse Document Frequency (TF-IDF)

The term-document matrix, denoted as A, plays a central role in our analysis (see Manning, Raghavan, and Schütze 2008). Each row of A corresponds to a unique word in the vocabulary, and each column of A corresponds to a document (Figure 1). The (i, j)-th entry of A is the term frequency (TF) of the i-th word appearing in the j-th document. Note that the term-document matrix ignores the ordering of words in the documents. Following convention, we include inverse document frequency (IDF) weighting for each term in the vocabulary (Manning, Raghavan, and Schütze 2008). This weighting scheme puts less weight on the terms that appear in more documents, and thus less emphasis is given to terms with less discriminative power.

4.1.3. Non-negative Matrix Factorization (NMF)

We focus on a particular linear algebraic method in unsupervised machine learning, namely nonnegative matrix factorization (NMF) (Lee and Seung 1999). NMF is designed for discovering interpretable latent components in high-dimensional unlabeled data such as the set of documents described by the counts of unique words. NMF uncovers major hidden themes by recasting the term-document matrix A into the product of two other matrices, one matrix representing the topics and another representing the documents in the latent topic space (Figure 1) (Xu, Liu, and Gong 2003). In particular, we would like to find matrices $W \in \mathbb{R}_{+}^{m \times k}$ and $H \in \mathbb{R}_{+}^{k \times n}$ to solve the approximation problem $A \approx WH$, where $\mathbb{R}_{+}$ is the set of all nonnegative numbers and *m*, *n* and *k* are the numbers of unique words, documents, and topics, respectively. A is the term-document matrix given as the input, while W and H enclose the latent term-topic and topic-document information.

Numerous algorithms exist for solving $A \approx WH$ (Cichocki et al. 2009; Kim, He, and Park 2014), but most would take several hours to run on large-scale data sets consisting of millions of documents. We employ a highly efficient “hierarchical rank-2 NMF” algorithm that is orders of magnitude faster than directly solving $A \approx WH$ (Kuang and Park 2013). The algorithm first constructs a hierarchy of topics in the form of a binary tree, and then flattens the hierarchy to generate a traditional topic model. While the topic hierarchy is useful for explorative analysis, the flat level of topics enables easier quantitative evaluation. We show both forms in our analysis of crime data. In contrast to the hierarchical LDA (Teh et al. 2006), which is more computationally expensive than LDA, hierarchical NMF can process web-scale data containing millions to billions of documents such as Tweets or the crime narratives used in our study.

4.2. Cosine Similarity & Crime Type Clusters

Text-based topic modeling typically reveals that any one document is a mixture of different topics. Therefore, in principle, the distance between any two documents can be measured by comparing how far apart their topic mixture distributions are. Here we extend this idea to consider officially recognized crime types as mixtures of different crime topics. The distance between any two official crime types can be measured using the topic mixtures observed for those two crime types. We use cosine similarity (Steinbach, Karypis, and Kumar 2000) to compute such measures.

Consider two hypothetical crime types *A* and *B*. Type *A* might represent aggravated assault and type *B* might represent residential burglary. Inspection of all of the events formally classified as assault with a deadly weapon might show that 40% fall into crime topic $i = 1$, 30% fall into topic 4, 20% into topic 9, and 10% into topic 12. Similarly, for all the events formally classified as residential burglary, 5% might fall into topic $i = 9$, 15% into topic 12, 60% into topic 15 and 20% into topic 19. Assault with a deadly weapon and residential burglary are similar only in events falling into topics 9 and 12. More formally, the similarity between any two official crime types *A* and *B* is given as:

$$\cos(\theta) = \frac{\sum_{i=1}^{k} A_i B_i}{\sqrt{\sum_{i=1}^{k} A_i^2}\sqrt{\sum_{i=1}^{k} B_i^2}}$$

where A_i is the frequency at which events formally classified as crime type *A* belongs to topic *i* and equivalently for events formally classified as crime type B_i.

We choose cosine similarity over other measures such as KL-divergence and chi-square distances because cosine similarity is bounded, taking values between -1 and 1, and is a good measure for graph-based crime type clustering (discussed below). Negative values reflect distributions that are increasingly diametrically opposed and positive values distributions that point in the same direction. Values of cosine similarity near zero reflect vectors that are uncorrelated with one another. In our case, cosine similarity will only assume values between 0 and 1 because NMF returns only positive valued matrices.

Viewing the collection of official crime types as a graph, where each crime type is a node and cosine similarities define the weights of the edges between nodes, we use average linkage clustering (Legendre and Legendre 2012) on this graph to partition the crime types into ecologically meaningful groups (see also Brennan 1987: 228). Crime types are clustered in an agglomerative manner. Initially, each crime type exists as its own isolated cluster. The two closest clusters are then merged in a recursive manner, with the new cluster adopting the mean similarity from all cluster members. The process continues until only *C* clusters are left. The number *C* can be chosen automatically by a cluster validation method such as predictive strength (Tibshirani and Walther 2005), or manually for easier interpretation. We manually set the number of clusters.

5. Data and Analysis Plan

The above modeling framework is flexible enough in principle to handle any form of data (e.g., Chen, Wang, and Dong 2010), not just text. In spite of this flexibility, we do not stray far from its most common application in text mining. Here we exploit the presence of short text descriptions associated with individual crime events to compute text-based hierarchical NMF. Table 1 illustrates several examples of individual crime events and the associated text descriptions of the events.

We focus on the complete set of crimes reported to the Los Angeles Police Department (LAPD) from January 1, 2009 and July 19, 2014. The end date of the sample is arbitrary. Los Angeles is a city of approximately 4 million people occupying an area of 503 square miles. The Los Angeles Police Department is solely responsible for policing this vast area, though Los Angeles is both surrounded by and encompasses independent cities with their own police forces.

The total number of reported crimes handles by the LAPD during the sample period was 1,027,168. In a typical year, the LAPD collected reports on 180,000 crimes. On average 509 crimes were recorded per day, with crime reports declining over the entire period. During the first year of the sample, LAPD recorded on average 561.5 crimes per day. During the last year they recorded 463.8 crimes per day.

The crime coding system used by the LAPD includes 226 recognized crime types. This is considerably more finely resolved than either the FBI Uniform Crime Reports (7 Part I and 21 Part II offenses), or National Incident Based Reporting System (49 Group and 90 Group B offenses). Aggravated assault, for example, is associated with four unique crime codes including assault with a deadly weapon, assault with a deadly weapon against a police officer, shots fired at a moving vehicle, and shots fired at a dwelling. These crime types could be considered a type of ground truth against which topic model classifications can be evaluated. We are here interested in the degree of alignment of the LAPD crime types and topic models derived from text-based narratives accompanying those crimes.

In addition to this rich coding system, a large fraction of the incidents recorded in the sample include narrative text of the event. Of the 1,027,168 recorded crimes, 805,618 (78.4%) include some form of text narrative. On average 397.6 events per day contain some narrative text describing the event. The fraction of events containing narrative text increased over time from 76.6% of events, in the first six months of the sample, to 87.0%, in the last six months.

There are pointed differences in the occurrence of narrative text by official crime types (Table 2). Virtually all violent crimes are accompanied by narrative text. Robbery and homicide have associated narrative text for 98.9% and 98.2% of events, respectively. Assault and kidnapping have 97.8% and 97.4% of events associated with narrative text. Burglary shows narrative text occurrence on par with the most serious violence crimes (98.6%). For less serious property crimes, narrative text reporting falls off to 91.1% for theft and 74.3% for vandalism. The lowest occurrence of narrative text is seen for arson (37.8%) and motor vehicle theft (4.3%). In the former case it must be acknowledged that most arson reporting responsibilities lie with the fire department, so low narrative load might be expected. In the latter case, either the vehicles are not recovered (about 40% of the cases) and therefore the circumstances of the theft are not known, or detailed circumstances beyond make, model and year of the car—all recorded in separate fields—are not deemed as relevant to recording of the crime.

Overall, the text narratives associated with crime events total 7,649,164 discrete words, after preprocessing (see above). These are unevenly distributed across events. The mean number of words contained in a single narrative is 18.57 (s.d. 6.72), while the maximum number of words is 41 (see Table 1). Individual words are also unevenly distributed, though not massively so (Table 3). For example, the word “unknown” is the most common word in the corpus appearing 635,099 times. However, this still represents only 8.3% of all words. The next most common word is “property” occurring 305,014 times, but represents only 4% of all words. Words that are strongly indicative of crime type are extremely rare. The word homicide appears only 45 times in the entire text corpus, a frequency of $5.88x10^{-6}$ overall. Burglary appears 252 times, robbery 286 times, assault 457 times, and theft 969 times. When they do appear, diagnostic words are not generally coincident with the corresponding formal classifications. For example, of the 1,593 formally classified homicides in the dataset, only 11 of those events also find the word *homicide* as part of the narrative text. Thus, 1,582 formally classified homicides are not explicitly marked as such in the narrative text. The 34 events that include the word homicide in the narrative, but are not classified as homicides, include 17 events labeled as “other” (primarily threatening letters or phone calls), nine aggravated assaults, seven vandalism events, and one robbery. In general, narrative text provides context rather than strictly redundant typological detail. It is important to note, however, that narrative text and formal crime type classifications are unlikely to be completely decoupled. Ultimately, it is the job of police officers in the field to recognize and record behavioral and circumstantial evidence consistent with legal definitions of different crime types. Thus we should expect that specific narrative words correlate to some degree with formally recognized crime types.

The analysis that follows includes three parts. First, we present results for hierarchical topic models. We do this for all crimes combined and then turn our attention to analyses of the subset of crimes formally classified as aggravated assault and homicide. Second, we explore how formally recognized crime types are found distributed across different topics. The corollary that individual topics are distributed across different crime types is also discussed. Finally, recognizing that different formally recognized crime types are distributed across topics, we measure the 'distance' between different crime types based on the similarity in their topic mixtures.

6. Results

6.1. Hierarchical Models for All Crimes

Figure 2 presents a hierarchical topic model applied to all crime events in the LAPD corpus associated with narrative text. After preprocessing the data set includes 711,119 events. Each node in the tree represents a latent topic characterized by key words appearing in the topic. Summary statistics for the number of events, the percent violent and property crime, and the top-ten words for each topic node are shown in tabular format. The hierarchical structure is shown in graph form. Terminal leaf nodes are highlighted in gray.

The topic tree has three major components. The topics associated with the left branch (Nodes A-O) is linked to property crimes (Figure 2). Words such as *property* and *vehicle* identify key targets of crime, while words such as *window, door, enter, remove,* and *fled* describe the behavioral steps or sequences involved in commission of a crime. The validity of the property crime label for this component may be tested by using the formally recognized crime types in the LAPD ground truth. For example, 93.4% of the events associated with terminal leaf node C are formally recognized by the LAPD as property crimes. None of the intermediate or terminal nodes in the left branch (Nodes A-O) captures less than 89.9% property crimes.

By contrast, the right branch (Nodes P-AG) stands out for its connection to violent crime (Figure 2). Words such as *face, head* and *life* identify key targets of crime, while words such as *approach*, *verbal,* and *punch* identify sequences of behaviors involved in violent actions. The LAPD ground truth supports the broad label of topics P-AG as violent crime. For example, 90.5% of all the events associated with terminal topic S are formally recognized as violent crime types. With the exception of nodes P and Y, no other topic in this component captures less than 70% of formally recognized violent crimes. Terminal node Y appears to be an association of violations of court orders and/or annoying communications, which are reasonable ecological precursors to or consequences of other violent crimes.

Intermediate node P is a bridge between crime topics that are clearly associated with violent crime (Nodes Q-AG) and a series of crime topics we label as deception-based property crime (Nodes AH-AL). Words indicative of shoplifting and credit card fraud stand out in this group of topics. Why such topics trace descent through a branch more closely with violent is unclear.

6.2. Hierarchical Models for Aggravated Assault & Homicide

Figure 3 presents topic modeling results for the subset of crimes formally classified by the LAPD as aggravated assaults (LAPD code 230) and homicide (LAPD code 110). This is a semi-supervised analysis in the sense that we have used information external to narrative data to partition or stratify the collection of events into *a priori* groups. Our goal is to assess topic distinctions that arise within these serious violent crimes. A total of 40,208 events are classified as either aggravated assaults (38,626 events) or homicides (1,582 events). Notionally, these events are separated on the basis of outcome (i.e., death), but such a distinction is not visible within the classification hierarchy. Rather, the key distinction is between topics involving weapons other than firearms (Nodes A-I) and those involving firearms (Nodes J-R). Homicide looms large in terms of legal and harm-based classification (Ratcliffe 2015; Sherman 2011), but is not resolved within the larger volume of aggravated assaults. Homicides never make up more than 2.1% of any of the non-gun violence topics (Nodes A-I) (Figure 3). Homicides never rise above 11.8% in the gun violence topics (Nodes J-R). Notably, the greater lethality of guns is clearly visible when comparing the percent of homicides that are gun-related and those that are not. The most lethal crime topic is terminal node N, with key words *approach*, *handgun*, *multiple*, *shot*, and *fled*. Node P stands out with an emphasis on the use of vehicles as a weapon, but still tracing a pattern of descent linked to gun violence. Inspection of the top 100 words in this topic confirms that gun-related terms do not appear in topic P. The close connection to topic Q, which links guns and vehicles, is clearly through the common element of vehicles not guns.

Figure 4 shows that removing homicides from the subset of events does not fundamentally change the structure of the resulting topics. Indeed, it seems clear that assaults provide the overriding structure for crimes of interpersonal violence. This outcome may reflect the relatively low volume of homicides relative to aggravated assaults, but also the fact that homicides and aggravated assaults are ecologically very closely related (Goldstein 1994). Topic nodes A-I are notable for making fine-grained distinctions between the targets of violence, including *head*, *face*, *hand*, and *arm*, the weapons used, including *metal object*, *bottle*, and *knife*, and the action, including *hit*, *threw*, *punch*, *kick*, *stab*, and *cut*. The topics appear tactically very exacting. For example, the topics consistently show knives being used to target the body, while bottles/blunt object are used to target the head (Ambade and Godbole 2006; Webb et al. 1999).

6.2.1. Hierarchical model for homicides

Figure 5 presents the results of hierarchical NMF analysis of text narratives associated with formally classified homicides. There are clear distinctions that surface within formally classified homicides in spite of the much smaller numbers of events (1,414 with more than three words). The primary split is between homicides involving firearms (Node A and all of its daughters) and those where firearms are not indicated (Node R). Node R in fact features words *stab* and *head*, which we know from the broader analysis of aggravated assaults are two terms associated with knife violence and blunt-force violence, respectively (see Figure 3 and Figure 4). Node H implicate *gangs* exclusively in relation to gun violence. Nodes D, F and G highlight the central role of *vehicles* in gun violence. In each of these latter topics, words showing people emerging to attack or being attacked in cars, lending much behavioral and situational nuance to gun violence. By contrast, the adjacent branch (Nodes I-Q) appears to capture street-based homicides where the offender *approached* and *fled* on *foot*.

6.3. Crimes as Mixtures of Topics

The above discussion points to key terms such as *knife*, *gun*, and *glass*, or *stab*, *shot*, *hit*, that are useful in discriminating types of events from a range of behaviors and settings associated with different crimes. However, terminal topics are not themselves discrete. Rather, there is considerable overlap in the words or terms that populate different topics. This observation leads to a conceptualization of crimes as mixtures of crime different topics.

Table 4 shows a confusion matrix for formal crime types assigned by the LAPD against the topics associated with each crime event. A confusion matrix is typically used for evaluating the performance of a predictive algorithm (Fielding and Bell 1997). Here a confusion matrix is used to illustrate both how official crime types exist as mixtures of topics and how individual topics are associated with many different official crime types. We use a refined version of the leaf nodes from hierarchical clustering for all crime types and number the topics from 1 to 20 (see Figure 2). We also restrict the confusion matrix to the thirty most common crime types in the dataset for readability. Clustering analyses below restrict the analysis to the forty most common crime types.

Official crime types mix topics in unique ways. Row counts in Table 4 give the number of events of a given official crime type that are assigned to different discovered crime topics. For example, 29,497 (32.94%) of the 89,552 events officially classified by the LAPD as burglary from vehicle are assigned to Topic 1. This topic is marked by words *smash*/*broke*, *rear*/*passenger*/*side*/*driver*/*front*, *window*, and *remove*, all of which provide clear target and behavioral information intuitively consistent with the official crime type. However, other topics also grab significant numbers of burglary from vehicle events. Topics 3 (7.25%), 5 (5.02%), 8 (14.14%), 10 (10.87%), 14 (8.79%), and 19 (9.09%) each represent at least 5% of total events (Table 4). Topic 8 shares a connection on property crime with Topic 1, but otherwise emphasizes a very different focus, marked by words such as *force*/*gain*, *access*/*entry*, *tool*, *remove* and *property*. Topic 8 sounds considerably more generic and is consistent with burglary in general. Similarly, Topic 10 also grabs a large number of burglary from vehicle events, but here the focus is more clearly on vandalism, marked by words such as *kei* ([sic] i.e., *key*), *scratch* and *tire*. A more formal analysis of mixture characteristics is presented below.

Topic mixtures also characterize violent crimes. For example, aggravated assault (or assault with a deadly weapon) has events distributed evenly across Topics 2 (7,689 events or 18.02%), 6 (8041 events, 18.84%) and 9 (8,038 events, 18.83%). Topic 2 is characterized by words such as *punch*/*kick*, *hit*/*struck*, *face*/*head*, without prominent occurrence of words related to weapons. Topic 6, by contrast, features words such as *gun*/*handgun* as well as *approach*, *demand* and *money*. Topic 9 involves words such as *verbal*, *argument*/*dispute*, *grab*, *push*, and *hand*. While aggravated assaults appear to be evenly divided among these three topics, the topics themselves suggest heterogeneity in crime contexts. Topic 8 clearly stands out as related to robbery.

Crime topics are also not exclusively linked to individual crime types (Table 4). Rather single topics are spread across crime types at different frequencies. For example, 58.63% (24,497) of the Topic 1 events fall within burglary from vehicle. However, 12.99%, 10.77% and 9.7% of Topic 1 events are classified as petty vandalism under $400, vandalism over $400 and burglary, respectively. Topic 1 thus reveals connections among three different crime types. Such is the case for each topic. For example, 14.3% (8,041) of Topic 6 events are aggravated assaults, though robbery is the single most common crime type attributed to this topic (41.15% or 23,112 events). Battery (9.17% or 5,147 events), attempted robbery (6.8% or 3,820 events) and theft from person (5.3% or 2,979 events) are all also heavily represented within Topic 6.

Overall, the confusion matrix gives the sense that crimes may be related to one another in subtle ways and that these subtle connections can be discovered in the narrative descriptions of those events. A more formal way to consider such connections is to measure the similarities in their topic mixtures. The premise is that two crime types are more similar to one another if their distribution of events over topics is similar. For example, burglary from vehicle and petty vandalism show similar relative frequencies of events within Topic 3 (7.3% and 5.0%, respectively), Topic 5 (5.0% and 7.8%) and Topic 10 (10.9% and 12.2%) (Table 4). This gives the impression that burglary from vehicle and petty vandalism are closely related to one another.

6.4. Distances Between Crime Types & Crime Topic Clustering

To develop a more rigorous quantitative understanding of the relationships among formally recognized crime types we turn to the cosine similarity metric (Steinbach, Karypis, and Kumar 2000). Figure 6 shows the cosine similarity between formally recognized crime types as a matrix plot where the gray-scale coloring reflects the magnitude of similarity. The matrix is sorted in descending order of similarity. The darkest matrix entries are along the diagonal confirming that any one crime type is most similar to itself in the distribution of events across topics. More revealing is the ordering of crime types in terms of how far their similarities extend. For example, the rank 1 crime type, ‘other miscellaneous crimes’, has a topic distribution that is broadly similar to the topic distributions for every other crime type (Figure 6). The classification ‘other miscellaneous crime’ is a grab-bag for events that do not fit well into other categorizations. It is reasonable to expect that such crimes will occur randomly with respect to setting and context and therefore share similarities with a wide array of other crime types. What is astonishing is that this broad pattern of connections is picked up in the comparison of topic profiles.

More surprising perhaps are the widespread connections shared by shots fired (rank 2) and aggravated assault (assault with a deadly weapon) (rank 3) with other crimes. Guns appear to mix contextually with many other formally recognized crime types. By contrast, robbery and attempted robbery show a more limited set of connections. Both of these latter crime types display particularly weak connections to burglary and vandalism. Identity theft appears to be largely isolated in its topic structure from other crimes (rank 20).

Figure 7 goes one step further to identify statistical clusters, or communities within similarity scores using average linkage clustering (Legendre and Legendre 2012). We focus on a six cluster solution using this method. Consistent with Figure 6, identity theft is clustered only with itself (pink). This is also the case for shoplifting (brown). The first major cluster (purple) includes burglary, petty and grand theft, attempted burglary, trespassing, bike theft, and shots fired at an inhabited dwelling. The second cluster (red) includes burglary from vehicle, petty and serious vandalism, petty and grand theft from vehicle, embezzlement, and vehicle stolen. The third cluster (green) includes criminal threats, forged documents, other miscellaneous crimes, annoying behavior, violation of a court or restraining order, child endangering, bunco and disturbing the peace. The final and largest cluster (orange) incudes violent crimes such as battery, robbery, aggravated assault (assault with a deadly weapon), attempted robbery, theft from person, brandishing a weapon, battery on a police officer, shots fired, homicide, resisting arrest and kidnapping.

7. Discussion and Conclusions

The application of formal crime classifications to criminal events necessarily entails a massive loss of information. We turn to short narrative text descriptions accompanying crime records to explore whether information about the complex behaviors and situations surrounding crime can be automatically learned and whether such information provides insights in to the structural relationships between different formally recognized crime types.

We use a foundational machine learning method known as non-negative matrix factorization (NMF) to detect crime topics, statistical collections of words reflecting latent structural relationships among crime events. Crime topics are potentially useful for not only identifying ecologically more relevant crime types, where the behavioral situation is the focal unit of analysis, but also quantifying the ecological relationships between crime types.

Our analyses provide unique findings on both fronts. Hierarchical NMF is able to discover a major divide between property and violent crime, but below this first level the differences between crime topics hinge on quite subtle distinctions. For example, six of eight final topics within the branch linked to property crime involve crimes targeting vehicles or the property therein (see Figure 2). Whether entry is gained via destructive means, or non-destructive attack of unsecured cars seems to play a key role in distinguishing between crimes. Such subtleties are also seen in the topics learned from arbitrary subsets of crimes. For example, among those crimes formally classified as aggravated assault and homicide shows a clear distinction between topics associated with knife/sharp weapon and gun violence (see Figures 3, 4 and 5). A distinction is also seen between violence targeting the body and that targeting the face or head. Few would consider knife and gun violence equivalent in a behavioral sense. That this distinction is discovered and given context is encouraging.

Individual crime types are found distributed across different topics, suggesting subtle variations in behaviors and situations underlying those crimes. Such variation also implies connections between different formally recognized crime types. Specifically, two events might be labeled as different crime types, but arise from very similar behavioral and situational conditions and therefore be far more alike than their formal labels might suggest. Clustering of crimes by their topic similarity shows that this is the case. As presented in Figure 7, some crime types stand out as isolated from all other types (e.g., identity theft, shoplifting). Other crime types cluster more closely together. For example, the formal designation ‘shots fired’ does connect more closely with other violent crime types such as assault, battery and robbery, even though ‘shots fired’ is found widely associated with many other crimes as well. Burglary from vehicle clusters more closely with vandalism and embezzlement than it does with residential or commercial burglary.

The similarity clusters confirm some aspects of intuition. Violent crimes are naturally grouped together. Burglary and theft are grouped together. Burglary from vehicle, car theft and vandalism are grouped together. Less intuitive perhaps is the group that combines criminal disturbance with 'confidence' crimes such as forged documents and bunco.

7.1. Implications

We can think of the clusters identified in Figure 7 as ecological groups that are close to one another in the behaviors and situations that drive the occurrence of those crimes. This observation has potential implications for understanding causal processes as well as designing avenues for crime prevention. It is possible that crimes that are closer together in terms of their topic structure share common causes, while those that occupy different clusters are separated along causal lines. For example, it is intriguing that burglary occupies a separate cluster (i.e., is topically more distant) from burglary from vehicle (Figure 7). Clearly the differences between targets (i.e., residence vs vehicle) plays a key role here, but other behavioral and situational differences might also prove significant. For example, the tools and methods for gaining entry to each type of target are quite different, and words associated with such tools-of-the-trade and stand out for their discriminative value (see Figure 2). Other hidden structures might also tie crimes together. The grouping of burglary with theft suggests a focus on loss of property, while the grouping of burglary from vehicle with vandalism suggests a focus on property destruction. It is also possible that degrees of professionalism or skill are part of the structural mapping. Vandalism is reasonably considered a crime requiring a bare minimum of skill and therefore presents very few barriers to entry. Burglary from vehicle requires perhaps only a small increase in skill above this baseline. Theft and burglary, by contrast, may require a minimum degree of expertise and planning (Wright, Logie, and Decker 1995), though it would be a stretch to describe these as high-skill activities.

Several distinctions also stand out with respect to violent crimes. Notably, several crimes that might be thought of as precursors of violence do not cluster directly with violent crime. For example, criminal threats, violations of court and restraining orders, and threatening phone calls all occupy a cluster along with the catch-all ‘other crime’. Conversely, theft from person (i.e., theft without threat of force) clusters with violent crimes, though in a technical sense it is considered a non-violent crime. Robbery is a small step away from theft from person and one wonders whether routine activities that facilitate the less serious crime naturally lead to the more serious one.

The clustering shown in Figure 7 may also imply something about the ability to generalize crime prevention strategies across crime types. It may be the case that crimes that cluster together in topical space may be successfully targeted with a common set of crime prevention measures. The original premise behind ‘broken windows policing’ was that efforts targeting misdemeanor crimes impacted the likelihood of felony crime because the same people were involved (Wilson and Kelling 1982). It is also possible that policing efforts targeting certain misdemeanor crime types may have an outsized impact on certain felony crime types because they share similar behavioral and situational foundations, whether or not the same people are involved. Figure 7 suggests, for example, that targeting the conditions that support theft from person might impact robberies. Efforts targeting vandalism might impact burglary from vehicle. In general, we hypothesize that the diffusion of crime prevention benefits across crime types should first occur within crime type clusters and only then extend to other crime clusters.

7.2. Limitations

There are several limitations to the present study. The first concerns unique constraints on text-based narratives associated with crime event records. These narratives are unlikely to be completely free to vary in a manner similar to other unstructured text systems. Tweets are constrained in terms of the total number of characters allowed. Beyond this physical size constraint, however, there is literally no limit to what can be expressed topically in a Tweet. Additional topical constraints are surely at play in the composition of narrative statements about crime events. For example, the total diversity of crime present in an environment likely has some upper limit (Brantingham 2016). Thus, narratives describing such crimes may also have some topical upper limit. In addition, we should recognize that the narrative text examined here has a unique bureaucratic function. Text-based narratives are presumably aimed at providing justification for the classification of the crime itself. As alluded to above, this likely means that there is a preferred vocabulary that has evolved to provide minimally sufficient justification. Thus we can imagine that there has been a co-evolution of narrative terms and formal crime types that impacts how topics are ultimately resolved. The near complete separation of property from violent crimes in topic space may provide evidence that such is the case.

A second limitation surrounds our ground truth data. We assumed that the official crime type labels applied to crime events are accurate. However, crime type labels may harbor both intentional and unintentional errors (Gove, Hughes, and Geerken 1985; Maltz and Targonski 2002; Nolan, Haas, and Napier 2011). The application of a crime type label is to some extent a discretionary process and therefore the process is open to manipulation. Additionally, benign classification errors both at the time of report taking and data entry are certainly present. If such mislabeling is not accompanied by parallel changes in the event narrative text, then there are sure to be misalignments between official crime types and discovered crime topics. What would be needed is a ground truth crime database curated by hand to ensure that mislabeling of official crime types is kept to a minimum. Curation by hand is not practical in the present case with ~1 million crime records.

The challenge of mislabeling suggests a possible extension of the work presented here. It is conceivable that a pre-trained crime topic model could be used as an autonomous "cross-check" on the quality of official crime type labels. We envision a process whereby a new crime event, consisting of an official crime type label and accompanying narrative text, is fed through the pre-trained topic model. The event is assigned to its most probable topic based on the words occurring in the accompanying narrative text. If there is a mismatch between the officially assigned crime type and the one determined through crime topic assignment, then a alarm might be set for additional review.

More ambitious is the idea that a ground-truth topic model could be used for fully autonomous classification. Here a new event consisting only of narrative text would be evaluated with an official crime type assigned based on the most probable classification from the topic model. No human intervention would be needed. Exploratory work on this process shows, however, that the narrative texts accompanying crime events in our data sample provides too little information for autonomous classification to be accurate at the scale of individual crime types. Police will almost always have more complete information at the time of assigning official crime type labels. While text-based topic models exploit novel information in a novel way, we must conclude for the moment that the crime topic model presented here is insufficient for fully autonomous classification, especially given the legal demands that would be placed on assigned crime types.

Nevertheless, the analyses presented here suggest that larger scale crime classes can be learned automatically from unstructured text descriptions of those crimes. Individual crimes existing as mixtures of different crime topics and, simultaneously, individual crime topics being distributed across nominally different crime types. Reiterating the conceptual connection with traditional topic modeling methods, the situation with crime parallels the idea that a single Tweet may draw on a mixture of different topics, while a single topic may be distributed across many quite distinctive Tweets. Our view is that latent ‘crime topics’ capture features of the behaviors and situations underlying crimes that are often impractical to observe and almost completely lost when adopting formal crime classifications. Crime topics also hold potential for greater understanding of the situational causes of crime less constrained by the byproducts of formal crime type classifications. Extending causal inferences using crime topics will be the subject of future work.

8. References Cited

Ambade, Vipul Namdeorao, and Hemant Vasant Godbole. 2006. Comparison of wound patterns in homicide by sharp and blunt force. *Forensic Science International* 156 (2–3):166-170.

Arora, Sanjeev, Rong Ge, Yonatan Halpern, David M Mimno, Ankur Moitra, David Sontag, Yichen Wu, and Michael Zhu. 2013. A Practical Algorithm for Topic Modeling with Provable Guarantees. Paper read at ICML (2).

Belk, Russell W. 1975. Situational Variables and Consumer Behavior. *Journal of Consumer Research* 2 (3):157-164.

Blei, David M. 2012. Probabilistic topic models. *Communications of the ACM* 55 (4):77-84.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3:993-1022.

Brantingham, P. Jeffrey. 2016. Crime Diversity. *Criminology* 54:553-586.

Brantingham, P. L., and P. J. Brantingham. 1993. Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology* 13.

Brantingham, Patricia L., and Paul J. Brantingham. 1993. Environment, Routine, and Situation: Toward a Pattern Theory of Crime. *Advances in Criminological Theory* 5 (2):259-94.

Brantingham, Paul J., and Patricia L. Brantingham. 1978. A Theoretical Model of Crime Site Selection. In *Theoretical Perspectives on Crime and Criminal Justice*, edited by M. Krohn and R. L. Akers. Beverly Hills: Sage.

———. 1984. *Patterns in Crime*. New York: Mackmillan.

Brennan, Tim. 1987. Classification: An Overview of Selected Methodological Issues. *Crime and Justice* 9:201-248.

Casey, Anthony J, and Anthony Niblett. 2015. The Death of Rules and Standards. *Available at SSRN 2693826*.

Chen, Yanhua, Lijun Wang, and Ming Dong. 2010. Non-negative matrix factorization for semisupervised heterogeneous data coclustering. *IEEE Transactions on Knowledge and Data Engineering* 22 (10):1459-1474.

Cichocki, Andrzej, Rafal Zdunek, Anh Huy Phan, and Shun-ichi Amari. 2009. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. Chichester, UK: John Wiley & Sons.

Clarke, R. V. G. 1980. "Situational" Crime Prevention: Theory and Practice. *The British Journal of Criminology* 20 (2):136-147.

Clarke, Ronald V. 1983. Situational Crime Prevention: Its Theoretical Basis and Practical Scope. *Crime and Justice* 4:225-256.

Clarke, Ronald V., and Derek B. Cornish. 1985. Modeling Offenders' Decisions: A Framework for Research and Policy. *Crime and Justice* 6:147-185.

Fielding, Alan H., and John F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24 (1):38-49.

Gibbs, Jack P. 1960. Needed: Analytical Typologies in Criminology. *The Southwestern Social Science Quarterly* 40 (4):321-329.

Glaeser, Edward L., and Andrei Shleifer. 2002. Legal Origins. *The Quarterly Journal of Economics* 117 (4):1193-1229.

Goldstein, Arnold P. 1994. *The Ecology of Aggression*. New York: Plenum.

Gove, Walter R., Michael Hughes, and Michael Geerken. 1985. ARE UNIFORM CRIME REPORTS A VALID INDICATOR OF THE INDEX CRIMES? AN AFFIRMATIVE ANSWER WITH MINOR QUALIFICATIONS*. *Criminology* 23 (3):451-502.

Groff, Elizabeth, David Weisburd, and Sue-Ming Yang. 2010. Is it Important to Examine Crime Trends at a Local "Micro" Level?: A Longitudinal Analysis of Street to Street Variability in Crime Trajectories. *Journal Of Quantitative Criminology* 26 (1):7-32.

Hong, Liangjie, and Brian D Davison. 2010. Empirical study of topic modeling in twitter. Paper read at Proceedings of the first workshop on social media analytics.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg. 2008. The Spreading of Disorder. *Science* 322:1681-1685.

Kim, Jingu, Yunlong He, and Haesun Park. 2014. Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework. *Journal of Global Optimization* 58 (2):285-319.

Kuang, Da, and Haesun Park. 2013. Fast rank-2 nonnegative matrix factorization for hierarchical document clustering. Paper read at Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining.

Lee, Daniel D, and H Sebastian Seung. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401 (6755):788-791.

Legendre, Pierre, and Louis Legendre. 2012. *Numerical Ecology*. 3 ed. Oxford: Elsevier.

Maltz, Michael D, and Joseph Targonski. 2002. A note on the use of county-level UCR data. *Journal of Quantitative Criminology* 18 (3):297-318.

Manning, Christopher D, Prabhakar Raghavan, and Hinrich Schütze. 2008. Introduction to information retrieval: Cambridge University Press.

Nolan, James J, Stephen M Haas, and Jessica S Napier. 2011. Estimating the impact of classification error on the "statistical accuracy" of uniform crime reports. *Journal of Quantitative Criminology* 27 (4):497-519.

Rajman, M., and R. Besançon. 1998. Text Mining: Natural Language techniques and Text Mining applications. In *Data Mining and Reverse Engineering: Searching for semantics. IFIP TC2 WG2.6 IFIP Seventh Conference on Database Semantics (DS-7) 7–10 October 1997, Leysin, Switzerland*, edited by S. Spaccapietra and F. Maryanski. Boston, MA: Springer US.

Ratcliffe, Jerry H. 2015. Towards an Index for Harm-Focused Policing. *Policing* 9 (2):164-182.

Sellin, Thorsten. 1938. Culture Conflict and Crime. *American Journal of Sociology* 44 (1):97-103.

Sherman, Lawrence W. 2011. Al Capone, the Sword of Damocles, and the Police–Corrections Budget Ratio. *Criminology & Public Policy* 10 (1):195-206.

Steinbach, Michael, George Karypis, and Vipin Kumar. 2000. A comparison of document clustering techniques. Paper read at KDD workshop on text mining.

Teh, Yee Whye, Michael I. Jordan, Matthew J. Beal, and David M. Blei. 2006. Hierarchical Dirichlet Processes. *Journal of the American Statistical Association* 101 (476):1566-1581.

Tibshirani, Robert, and Guenther Walther. 2005. Cluster Validation by Prediction Strength. *Journal of Computational and Graphical Statistics* 14 (3):511-528.

Webb, E., J. P. Wyatt, J. Henry, and A. Busuttil. 1999. A comparison of fatal with non-fatal knife injuries in Edinburgh. *Forensic Science International* 99 (3):179-187.

Weisburd, David L., Elizabeth R Groff, and Sue-Ming Yang. 2012. *The criminology of place: Street segments and our understanding of the crime problem*: Oxford University Press.

Weisburd, David, and Alex R. Piquero. 2008. How Well Do Criminologists Explain Crime? Statistical Modeling in Published Studies. *Crime and Justice* 37 (1):453-502.

Wilson, J. Q., and G. Kelling. 1982. Broken Windows: The Police and Neighborhood Safety. *The Atlantic Monthly* 127:29-38.

Wright, Richard, Robert H. Logie, and Scott H. Decker. 1995. Criminal Expertise and Offender Decision Making: An Experimental Study of the Target Selection Process in Residential Burglary. *Journal of Research in Crime and Delinquency* 32 (1):39-53.

Wright, Richard T., and Scott H. Decker. 1994. *Burglars on the Job: Streetlife and Residential Breakins*. Boston: Northeastern University Press.

———. 1997. *Armed Robbers in Action: Stickups and Street Culture*. Boston: Northeastern University Press.

Xu, Wei, Xin Liu, and Yihong Gong. 2003. Document clustering based on non-negative matrix factorization. Paper read at Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.

1. Figures

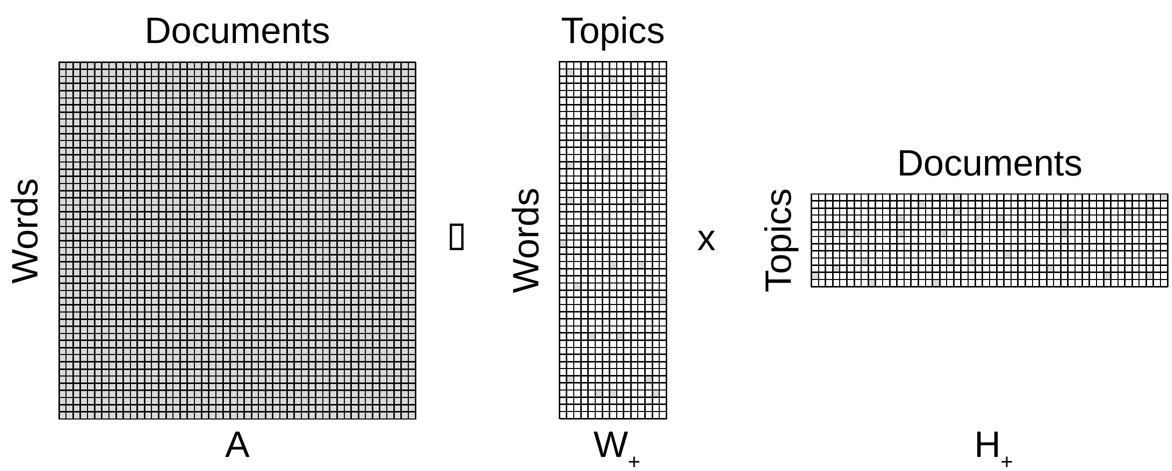


Figure 1. Conceptual illustration of non-negative matrix factorization (NMF) decomposition of a matrix consisting of *m* words in *n* documents into two non-negative matrices of the original *n* words by *k* topics and those same *k* topics by the *m* original documents.

Property Crimes

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
357,473	102,600	78,045	24,555	254,873	94,146	160,727	42,152	118,575	14,302	104,273	66,414	31,638	34,776	37,859
(5.7%, 94.3%)	(6.1%, 93.9%)	(6.6%, 93.4%)	(4.4%, 95.6%)	(5.5%, 94.5%)	(5.0%, 95.0%)	(5.8%, 94.2%)	(0.9%, 99.1%)	(7.5%, 92.5%)	(1.9%, 98.1%)	(8.3%, 91.7%)	(7.3%, 92.7%)	(5.9%, 94.1%)	(8.5%, 91.5%)	(10.1%, 89.9%)
properti	window	smash	object	properti	door	properti	lock	properti	unlock	properti	properti	remov	locat	direct
unknown	smash	window	hard	door	resid	unknown	entri	remov	vehicl	direct	locat	properti	enter	unknown
remov	vehicl	vehicl	unknown	remov	open	remov	gain	direct	enter	unknown	remov	vehicl	properti	fled
vehicl	object	passeng	scratch	unknown	front	direct	tool	unknown	remov	locat	enter	unsecur	fled	mean
window	passeng	properti	vehicl	enter	rear	vehicl	secur	locat	door	remov	fled	unknown	unknown	vehicl
fled	unknown	unknown	damag	locat	enter	locat	cut	enter	properti	fled	unknown	permiss	direct	tool
direct	side	remov	sharp	fled	pri	fled	vehicl	fled	unknown	enter	vehicl	left	remov	remov
enter	hard	side	break	direct	ransack	enter	unknown	vehicl	open	vehicl	unsecur	ent	vehicl	properti
door	rear	rear	direct	open	window	lock	remov	unlock	possibl	mean	left	purs	resid	open
locat	properti	fled	fled	resid	properti	entri	bike	mean	park	resid	permiss	room	lock	door

Violent Crimes

P	Q	R	S	T	U	V	W	X
353,726	258,395	109,207	48,141	61,066	29,593	31,473	149,188	64,409
(61.6%, 38.4%)	(81.7%, 18.3%)	(90.2%, 9.8%)	(90.5%, 9.5%)	(89.9%, 10.1%)	(88.4%, 11.6%)	(91.4%, 8.6%)	(74.8%, 25.2%)	(74.6%, 25.4%)
store	punch	verbal	verbal	face	struck	punch	approach	phone
pai	verbal	face	disput	punch	fist	face	demand	kill
item	face	punch	involv	time	close	time	monei	call
punch	approach	involv	argument	fist	caus	kick	state	state
approach	involv	disput	engag	struck	head	ground	phone	threaten
face	argument	argument	push	head	hit	approach	fear	cell
verbal	disput	time	angri	caus	injuri	argument	kill	im
exit	time	struck	alterc	close	time	fled	grab	fear
locat	push	push	grab	hit	face	began	fled	life
time	struck	fist	enrag	injuri	visibl	push	point	text

Y	Z	AA	AB	AC	AD	AE	AF	AG
28648	35761	84779	31831	52948	16090	36858	32315	4543
(54.7%, 45.3%)	(86.9%, 13.1%)	(74.9%, 25.1%)	(73.8%, 26.2%)	(75.6%, 24.4%)	(70.0%, 30.0%)	(78.0%, 22.0%)	(75.2%, 24.8%)	(97.1%, 2.9%)
phone	kill	approach	monei	approach	grab	approach	approach	knife
cell	state	demand	demand	grab	neck	foot	foot	stab
call	threaten	monei	point	foot	purs	fled	fled	pull
order	im	grab	handgun	fled	necklac	punch	punch	produc
violat	fear	fled	gun	hand	hand	ask	ask	attempt
text	call	point	fear	neck	pull	knife	properti	brandish
messag	life	foot	gave	purs	chain	state	locat	cut
court	told	handgun	approach	pull	arm	properti	hand	pocket
time	gonna	gun	properti	properti	fled	locat	state	approach
annoi	safeti	properti	give	locat	walk	pocket	push	arm

Deception-based Property Crime

AH	AI	AJ	AK	AL
95,331	41,328	54,003	31,289	22,714
(8.2%, 91.8%)	(0.7%, 99.3%)	(14.1%, 85.9%)	(11.3%, 88.7%)	(17.8%, 82.2%)
store	card	store	item	store
pai	credit	pai	pai	merchandis
item	info	item	conceal	pai
exit	account	conceal	select	exit
conceal	purchas	exit	exit	enter
select	obtain	select	locat	walk
enter	check	merchandis	busi	conceal
merchandis	person	enter	enter	properti
locat	permiss	locat	regist	select
walk	bank	walk	fail	remov

node
events
(v%, p%)

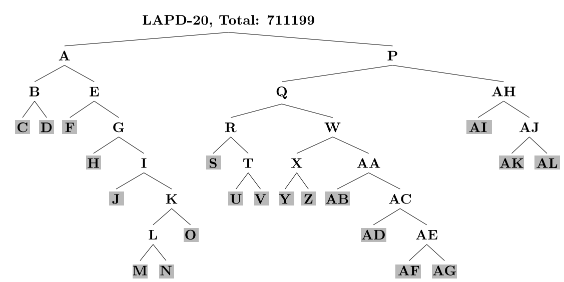


Figure 2. Hierarchical NMF topic structure for the entire corpus of events. The left branch captures property crimes. The right branch captures violent crimes. Deception-based property crimes form a distinct tree in the right branch. Tables show topic labels, number of events in each topic, number of events of the top 40 most frequent crime types in each topic, the percent of events for the topic that are formally classified as violent crime (v%) or property crime (p%), and the top-ten topic words. Terminal leaves of the topic model are marked in gray.

Non-gun Violence

A	B	C	D	E	F	G	H	I
21,875	10,600	5,960	4,640	11,275	4,114	7,161	2,197	4,964
(1.6%, 98.4%)	(1.9%, 98.1%)	(1.8%, 98.2%)	(2.1%, 97.9%)	(1.2%, 98.8%)	(0.9%, 99.1%)	(1.4%, 98.6%)	(0.3%, 99.7%)	(1.8%, 98.2%)
verbal	knife	knife	verbal	head	punch	bottl	bottl	injuri
knife	verbal	stab	disput	caus	kick	injuri	glass	caus
involv	stab	cut	involv	struck	ground	caus	head	struck
argument	involv	attempt	argument	injuri	face	head	threw	head
disput	disput	produc	engag	face	time	struck	beer	visibl
stab	argument	argument	alterc	punch	approach	glass	hit	hit
head	cut	pull	angri	bottl	head	hit	struck	metal
caus	engag	kitchen	enrag	hit	began	threw	face	object
struck	alterc	arm	hit	kick	fall	visibl	argument	argument
face	attempt	hand	struck	glass	fell	argument	verbal	bat

Gun Violence

J	K	L	M	N	O	P	Q	R
18,333	5,730	12,603	9,367	4,201	5,166	2,208	2,958	3,236
(5.9%, 94.1%)	(4.9%, 95.1%)	(6.4%, 93.6%)	(7.9%, 92.1%)	(11.8%, 88.2%)	(4.7%, 95.3%)	(1.5%, 98.5%)	(7.0%, 93.0%)	(2.1%, 97.9%)
unknown	unknown	vehicl	vehicl	fire	vehicl	vehicl	vehicl	point
vehicl	direct	fire	fire	round	drove	drove	exit	gun
shot	fled	shot	shot	strike	exit	intention	shot	handgun
fire	locat	round	round	shot	shot	hit	fled	state
fled	object	strike	strike	unknown	hit	attempt	fire	approach
locat	approach	handgun	drove	approxim	fled	collid	passeng	produc
direct	shot	gun	unknown	handgun	drive	ram	locat	pull
approach	stab	drove	fled	fled	locat	run	stop	kill
strike	time	point	exit	approach	stop	att	drive	fled
round	foot	exit	locat	multipl	attempt	jump	approach	fear

node
events
(h %, a %)

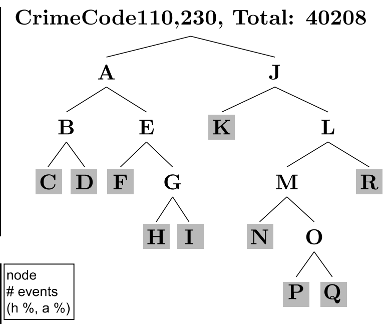


Figure 3. Hierarchical NMF for subset of crimes formally classified as aggravated assault and homicide. Terminal leaves of the topic model are marked in gray.

Non-gun Violence

A	B	C	D	E	F	G	H	I
21,514	11,147	7,018	4,704	2,314	4,129	10,367	4,490	5,877
verbal	head	bottl	injuri	bottl	punch	knife	verbal	knife
knife	caus	injuri	caus	glass	kick	verbal	disput	stab
involv	struck	caus	struck	head	ground	stab	involv	cut
argument	injuri	head	visibl	threw	face	involv	argument	attempt
disput	face	struck	head	beer	time	disput	engag	produc
head	punch	glass	metal	struck	head	argument	alterc	argument
stab	bottl	hit	hit	hit	approach	cut	angri	pull
caus	hit	threw	object	face	began	engag	enrag	kitchen
struck	kick	visibl	argument	argument	fall	attempt	hit	arm
face	glass	argument	time	verbal	fell	alterc	struck	hand

Gun Violence

J	K	L	M	N	O	P	Q	R
17,270	5,383	11,887	3,172	8,715	4,756	1,002	3,754	3,959
unknown	unknown	vehicl	point	vehicl	vehicl	drove	vehicl	fire
vehicl	direct	fire	gun	fire	drove	vehicl	exit	round
shot	fled	shot	handgun	shot	exit	shot	shot	strike
fled	locat	round	state	round	hit	unknown	fled	shot
fire	object	strike	approach	strike	shot	strike	hit	unknown
locat	approach	handgun	produc	drove	fled	locat	drive	approxim
direct	shot	gun	pull	unknown	drive	fire	locat	handgun
approach	stab	drove	kill	fled	locat	avoid	intention	fled
round	time	point	fled	exit	attempt	jump	attempt	approach
strike	struck	exit	fear	locat	stop	hit	stop	gang

node
events

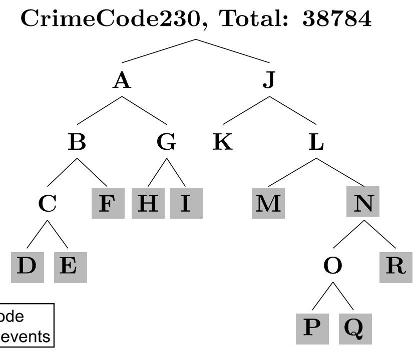


Figure 4. Hierarchical NMF for subset of crimes formally classified as aggravated assaults. Terminal leaves of the topic model are marked in gray.

Gun-Vehicle Homicide

A	B	C	D	E	F	G	H
996	387	243	48	195	80	115	144
vehicl	vehicl	vehicl	seat	vehicl	exit	vehicl	kill
fire	kill	exit	passeng	exit	vehicl	drove	gang
strike	walk	seat	front	drove	walk	drive	shoot
unknown	exit	drive	vehicl	drive	stop	shot	shot
multipl	seat	passeng	park	walk	shoot	fire	walk
approach	shot	drove	driver	fled	round	struck	car
fled	shoot	fire	kill	shot	fire	fled	reason
round	drive	walk	open	fire	fled	unknown	member
time	gang	park	shot	struck	street	walk	unknown
shot	st	front	sat	round	approach	st	confront

Non-gun Homicide

R
418
death
caus
stab
disput
involv
demis
verbal
head
unknown
time

node
events

Gun Homicide

I	J	K	L	M	N	O	P	Q
609	412	190	102	88	222	197	81	116
strike	strike	strike	round	strike	fled	di	injuri	di
approach	approach	round	numer	multipl	unknown	injuri	succumb	wound
multipl	unknown	fire	fire	fire	locat	hospit	scene	hospit
fire	fled	numer	strike	time	approach	succumb	multipl	gunshot
unknown	round	approach	unknown	approach	foot	wound	time	transport
time	fire	multipl	locat	apprchd	time	gunshot	struck	result
round	locat	vehicl	approach	gun	direct	transport	pronounc	fire
fled	time	stand	fled	stand	multipl	multipl	shot	multipl
locat	foot	time	vehicl	head	shot	scene	dead	unknown
foot	multipl	unknown	caus	handgun	enter	shot	unknown	shot

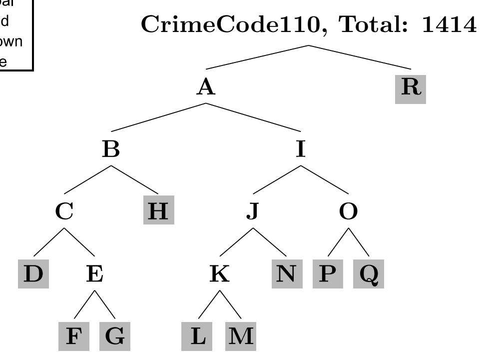


Figure 5. Hierarchical NMF for subset of crimes formally classified as homicides. Terminal leaves of the topic model are marked in gray.

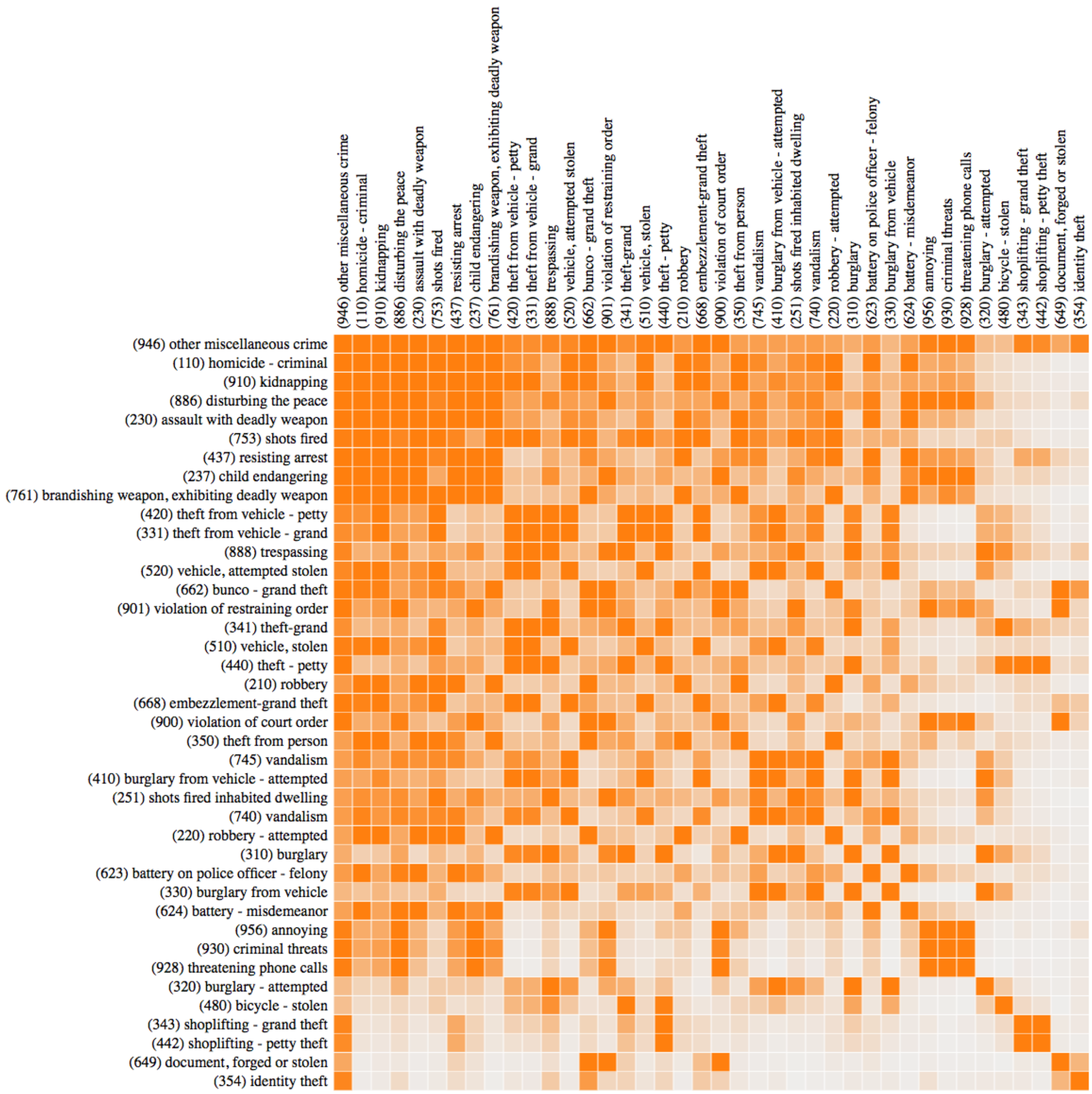


Figure 6. Cosine similarity between crime type pairs sorted in descending order of similarity.

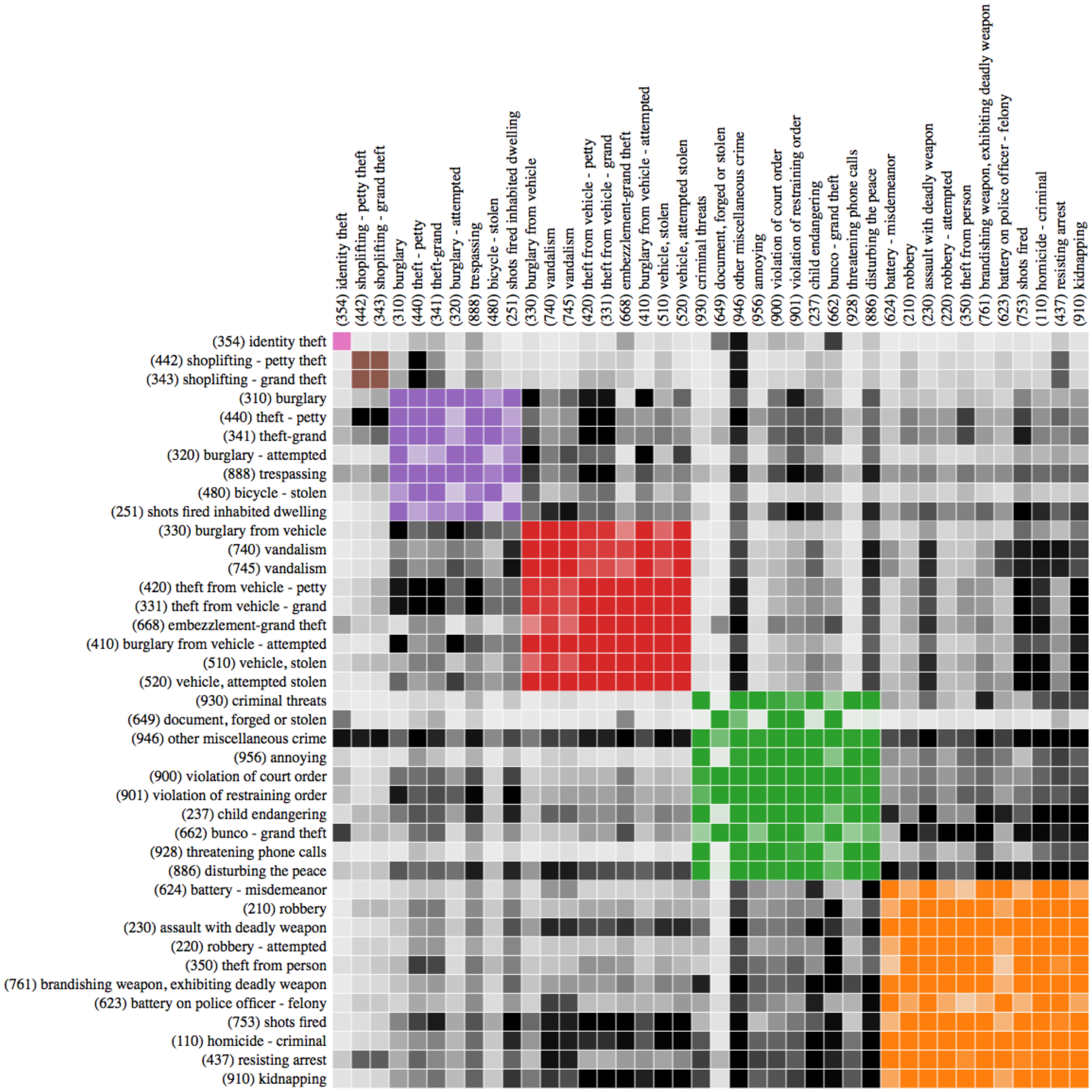


Figure 7. Average linkage clustering for cosine similarity between crime type pairs sorted by cluster proximity.

1. Tables

Table 1. Examples of official crime classifications and the narrative text tied to the event.

Official Crime Classification	Accompanying Narrative Text
Homicide	VICT IS A [GANG NAME] GANG MEMBER WAS STANDING ON SIDEWALK SPRAY PAINTING GRAFFITI SUSPS DROVE BY THE VICT FIRED SHOTS AT VICT
Assault	VICT AND SUBJ ARE MTHR DAUGHTER VICT ATTPT TO DISCIPLINE SUBJ SUBJ BECAME ANGRY AND ATTPT TO CUT VICT
Robbery	SUSP ENTERED LOCATION PRODUCED HANDGUN DEMANDED MONEY FROM REGISTER REMOVED PROPERTY FROM LOCATION AND FLED TO UNKNOWN LOCATION
Burglary	UNK SUSP ENTERED VICS RESID BY BREAKING SCREEN ON WINDOW WALKED THROUGHTHE RESID EXITED REAR DOOR AND ENTERED DETACHED GARAGE SUSP EXITED WITH PROPERT
Burglary-theft from Vehicle	SUSP USING PORCELAIN CHIPS BROKE VEHS WINDOW PRIOR TO SUSP GAINING ENTRY SUSP FLED THE LOC
Motor Vehicle Theft	SUSP ENTERED VIC VEH WITH UNK PRY TOOL AND REMOVED PROP FROM VEH SUSP PUNCHED IGNITION SWITCH
Theft	S ENTERED CLOTHING STORE AND TOOK APPROX 20 BLUE TSHIRT AND THEN FLED LOCATION WITHOUT PAYING

Table 2. Counts of events with and without accompanying narrative text by official crime type.

	No Narrative Text	Narrative Text	Total	Fraction with Narrative Text
Robbery	597	53,379	53,976	0.989
Burglary	1,320	91,260	92,580	0.986
Homicide	28	1,565	1,593	0.982
Assault	1,032	45,665	46,697	0.978
Kidnapping	45	1,707	1,752	0.974
Grand Theft Person	230	7,754	7,984	0.971
Theft	13,326	136,117	149,443	0.911
Burglary-theft from Vehicle	20,192	126,912	147,104	0.863
Other Miscellaneous Crime	72,518	256,816	329,334	0.780
Vandalism	27,630	80,038	107,668	0.743
Arson	1,111	675	1,786	0.378
Motor Vehicle Theft	83,521	3,730	87,251	0.043

Table 3. *The t*op twenty-five most common words in *the full text corpus consisting of 7,649,164 discrete words*.

Word	Count	Proportion
unknown	635,099	0.0830
property	305,014	0.0399
fled	277,770	0.0363
vehicle	255,609	0.0334
location	202,661	0.0265
removed	197,171	0.0258
entered	143,602	0.0188
window	106,461	0.0139
direction	106,412	0.0139
door	96,918	0.0127
residence	66,576	0.0087
front	57,912	0.0076
open	55,413	0.0072
approached	55,261	0.0072
rear	50,794	0.0066
smashed	45,553	0.0060
left	45,155	0.0059
entry	40,341	0.0053
store	36,515	0.0048
stated	36,068	0.0047
object	35,696	0.0047
money	33,608	0.0044
punched	33,317	0.0044
items	32,354	0.0042
face	31,653	0.0041

Table 4. Confusion matrix for official crime types by topics. Dominant words in each topic are shown across the top. Row totals reflect the total number of crimes formally classified under each crime type. Column total reflect the total number of crimes clustered within each topic. Boldface numbers are column maxima.

	'window'	'punch'	'door'	'card'	'direct'	'approach'	'store'	'entri'	'verbal'	'vehicl'	'kill'	'item'	'damag'	'lock'	'check'	'phone'	'object'	'left'	'properti'	'resid'	
	'smash'	'face'	'open'	'credit'	'unknown'	'demand'	'pai'	'gain'	'involv'	'park'	'state'	'busi'	'caus'	'secur'	'cash'	'cell'	'hard'	'purs'	'locat'	'ransack'	
	'passeng'	'time'	'front'	'info'	'fled'	'monei'	'exit'	'properti'	'disput'	'unlock'	'threaten'	'select'	'paint'	'cut'	'forg'	'call'	'unknown'	'return'	'remov'	'enter'	
	'side'	'fist'	'pri'	'account'	'tool'	'foot'	'conceal'	'remov'	'argument'	'driver'	'fear'	'pai'	'threw'	'bike'	'account'	'hand'	'sharp'	'miss'	'enter'	'rear'	
	'rear'	'struck'	'pry'	'permiss'	'mean'	'point'	'merchandis'	'forc'	'push'	'kei'	'im'	'conceal'	'sprai'	'bicycl'	'bank'	'order'	'scratch'	'wallet'	'fled'	'window'	
	'front'	'head'	'rear'	'obtain'	'properti'	'grab'	'select'	'access'	'engag'	'unknown'	'call'	'enter'	'injuri'	'tool'	'monei'	'violat'	'break'	'insid'	'unknown'	'bedroom'	
	'driver'	'kick'	'unlock'	'purchas'	'broke'	'handgun'	'enter'	'unknown'	'angri'	'scratch'	'life'	'locat'	'wall'	'park'	'busi'	'ask'	'smash'	'discov'	'unlock'	'poe'	
	'broke'	'close'	'tool'	'person'	'remov'	'gun'	'walk'	'made'	'grab'	'tire'	'told'	'purchas'	'kick'	'return'	'order'	'grab'	'type'	'unattend'	'poe'	'screen'	
	'unknown'	'hit'	'side'	'make'	'pry'	'fled'	'attempt'	'tool'	'alterc'	'insid'	'safeti'	'bag'	'visibl'	'garag'	'deposit'	'text'	'windshield'	'observ'	'busi'	'insid'	
	'remov'	'multipl'	'driver'	'ident'	'prop'	'hand'	'regist'	'locat'	'hand'	'drove'	'knife'	'shelf'	'scratch'	'miss'	'attempt'	'messag'	'fled'	'shop'	'ent'	'exit'	
Formal Crime Classification	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**T8**	**T9**	**T10**	**T11**	**T12**	**T13**	**T14**	**T15**	**T16**	**T17**	**T18**	**T19**	**T20**	**TOTAL**
BURGLARY FROM VEHICLE	**29,497**	148	6,495	267	4,499	36	22	**12,663**	7	**9,735**	25	1,239	243	**7,872**	81	371	**7,547**	594	8,138	73	89,552
BURGLARY	4,879	51	**19,294**	128	3,032	77	914	11,604	16	148	89	2,050	246	4,482	308	251	1,704	584	**16,463**	**16,705**	83,025
BATTERY	183	**28,972**	893	197	1,565	5,147	610	172	**26,721**	1,427	1,512	241	5,384	415	103	959	517	2,196	717	1,276	79,207
PETTY THEFT	115	172	1,726	1,872	**5,943**	1,423	9,969	699	330	1,110	574	**4,131**	170	5,378	1,258	**4,914**	102	**10,521**	14,737	2,259	67,403
IDENTITY THEFT	12	36	242	**45,672**	195	84	193	1,561	17	156	535	598	43	950	5,087	1,251	5	554	103	104	57,398
GRAND THEFT	84	138	1,241	1,435	4,946	1,543	1,526	1,310	178	1,173	561	1,666	144	3,773	1,821	1,099	108	5,581	13,289	3,081	44,697
ROBBERY	79	4,928	442	77	746	**23,112**	1,949	139	639	1,231	1,406	669	483	277	451	2,593	577	1,407	2,718	435	44,358
VANDALISM ($400 & over)	5,419	362	1,589	84	2,925	310	168	320	962	7,712	159	321	**12,595**	788	87	149	6,852	444	998	927	43,171
ASSAULT WITH DEADLY WEAPON	393	7,689	421	82	1,783	8,041	331	74	8,038	5,537	2,655	91	3,136	445	34	157	1,332	1,088	469	883	42,679
VANDALISM (less than $400)	6,534	375	1,834	85	2,845	318	142	428	1,184	4,454	210	300	8,318	1,186	83	247	5,423	486	820	1,103	36,375
CRIMINAL THREATS	53	497	216	64	168	925	131	55	3,434	242	**25,035**	62	182	68	58	1,234	53	223	61	595	33,356
THEFT FROM VEHICLE - PETTY	362	9	1,640	235	2,241	76	24	1,331	17	7,280	27	376	69	478	50	278	91	787	4,980	70	20,421
SHOPLIFTING	1	14	72	41	80	48	**12,353**	10	7	39	6	3,405	9	70	46	143	5	227	1,103	6	17,685
THEFT FROM VEHICLE - GRAND	184	3	1,038	68	1,825	36	8	891	6	4,899	13	312	53	452	28	96	77	501	3,625	58	14,173
FORGED OR STOLEN DOCUMENT	4	11	8	1,265	53	66	118	57	12	61	51	151	4	7	**9,099**	26	0	45	65	32	11,135
OTHER MISCELLANEOUS CRIME	77	343	303	1,249	211	625	1,280	220	648	1,000	1,350	125	377	376	721	345	45	393	347	487	10,522
ANNOYING/LEWD/OBSCENE PHONE CALLS/LETTERS	4	730	28	171	35	102	27	13	202	35	3,713	19	301	12	39	3,212	4	107	14	147	8,915
VIOLATION OF COURT ORDER	35	324	205	85	85	275	69	38	160	186	1,446	15	56	21	2,592	1,102	5	288	225	1,165	8,377
ATTEMPTED BURGLARY	771	10	2,174	4	262	21	19	1,861	3	17	14	37	111	274	14	10	213	33	250	945	7,043
ATTEMPTED ROBBERY	28	749	67	11	69	3,820	172	30	61	166	263	49	79	45	64	304	85	243	206	26	6,537
THEFT FROM PERSON	17	42	39	61	245	2,979	66	5	125	223	50	45	22	43	37	1,181	9	762	548	28	6,527
TRESPASSING/LOITERING ON PRIVATE PROPERTY	129	214	640	294	162	129	107	283	96	96	200	119	92	332	76	23	17	417	1,558	975	5,959
VIOLATION OF RESTRAINING ORDER	42	128	225	57	77	202	38	47	162	161	612	10	32	10	1,007	875	10	111	164	1,268	5,238
BRANDISHING WEAPON	25	62	71	3	82	1,249	49	10	761	266	524	8	38	31	3	21	19	70	37	115	3,444
CHILD ENDANGERING/NEGLECT	8	195	51	27	15	79	118	19	362	194	905	10	264	19	49	21	10	413	76	402	3,237
BICYCLE - STOLEN	6	4	43	12	156	31	31	19	2	11	3	3	3	1,365	1	1	8	151	401	58	2,309
EMBEZZLEMENT-GRAND THEFT	0	3	0	43	1	18	16	0	2	1,802	16	0	1	0	10	31	0	279	1	1	2,224
FELONY BATTERY ON POLICE OFFICER	12	822	33	12	9	97	66	26	163	73	95	3	381	6	30	8	10	137	19	5	2,007
BUNCO - GRAND THEFT	0	5	3	228	33	462	61	10	14	90	142	19	6	7	575	123	0	33	64	44	1,919
SHOTS FIRED	29	57	28	13	411	446	9	3	56	325	51	3	91	6	5	0	7	44	117	75	1,776
TOTAL	50,313	48,937	42,787	56,078	36,159	56,162	33,729	35,189	45,923	55,982	46,582	17,012	34,995	30,147	25,571	22,355	25,416	30,630	74,287	35,441	803,695

Los Angeles Police Department Community Safety Partnership
Evaluation Proposal Executive Summary
UCLA Luskin School of Public Affairs

Overview:
In law enforcement today, there is a profound need for a new paradigm accompanied by replicable models to promote public safety and truly expand the meaning of "police-community partnerships." To answer that need, the Los Angeles Police Department Community Safety Partnership (CSP) has emerged as an innovative model that pushes beyond traditional policing frameworks. Former Police Chief William J. Bratton created the foundation for this model; his successor, Chief Charlie Beck built on that foundation and created CSP to demonstrate what he termed "relationship-based policing." CSP was designed to reduce rampant gang violence that had long impacted residents of four City public housing developments, increasing public safety and quality of life. However, this innovative model has never been scientifically evaluated by an external researcher.

Purpose of Proposal:
The proposed research study will rigorously examine this new model of partnership policing, drawing upon the viewpoints of both law enforcement and the community to evaluate the effectiveness of the CSP over the past three years, capturing both its history and outcomes. As part of this, the research will objectively assess whether the CSP model actually works and -- if it is determined to be effective -- how the key elements of this new model of law enforcement can be implemented nationally. It will examine whether the building of public trust is directly related to the reduction of crime or whether these are two unrelated, co-occurring processes. Determining both efficacy and interrelationship is critical to understanding the CSP and its ultimate impact on reducing crime and violence.

Leadership Team:
Jorja Leap, PhD, UCLA Professor and Executive Director Social Justice Research Project
Jeff Brantingham, PhD, UCLA Professor and Founder, PredPol-Predictive Policing
Todd Franke, PhD, UCLA Professor and Chair, Luskin School of Public Affairs
Gerald Chaleff, Special Assistant for LAPD Constitutional Policing and Legal Counsel (Ret.)
Karrah Lompa, Director, Luskin Social Justice Research Project-Watts Leadership Institute

Advisory Group:
Along with the Leadership Team, the evaluation will be guided by an advisory committee to be comprised of Los Angeles Police Department Staff, key stakeholders and residents of the

involved settings in Watts and Boyle Heights. There will also be an additional advisory group, composed of Community Intervention Workers who, as part of the Los Angeles City Gang Reduction and Youth Development (GRYD) play an integral role in CSP efforts.

Questions to be Answered:
The key questions and focus of this evaluation effort will be developed in collaboration with the LAPD and evaluation advisory board. Some of the main points that may be used to guide this participatory research effort include:

- Understanding the Community Safety Partnership;
  - What is the CSP?
  - What are its roots and dynamics?
- Examining the effectiveness of CSP;
  - Does it actually work?
  - What are the measurable outcomes of the CSP?
  - What are its challenges?
- Mapping CSP history: focusing on its development as an initiative and intervention;
- Creating a history of engagement and accomplishment at different sites;
- Engaging and interviewing LAPD staff with different perspectives and experiences;
- Delineating initiative resources and mission;
- Examining disparate levels of relationships with community based organizations and how these are addressed;
- Documenting the process of relationship building – successes and challenges – in the varied settings in Los Angeles, applying these to differing US urban regions.

Methodology
To ensure scholarly rigor as well as validity and reliability, the proposed research will employ a mixed methods design, using both quantitative and qualitative measures that draw upon innovative approaches including real-time crime analysis, photo-voice, and continuous statistical analysis. Additionally, program sites will be contrasted with matched comparison sites to accurately assess the specific impact of the Community Safety Partnership. This combination of approaches will enable researchers to integrate traditional analysis of crime data and trends along with measuring outcomes that cannot be captured in crime statistics. It is critical for research to evaluate outcomes beyond decreases in crime – to understand the process of trust building within communities of color as well as the transformation of the narrative that traditionally dominates law enforcement.

Budget
$500,000 - $800,000

Dependent on number of sites and comparison sites surveyed

Los Angeles Police Department Community Safety Partnership
Evaluation Proposal

UCLA Luskin School of Public Affairs
Social Justice Research Partnership

> Community policing cannot be a program, unit, strategy or tactic. It must be the core principle that lies at the foundation of a police department's culture. The only way to significantly reduce fear, crime, and disorder and then sustain these gains is to leverage the greatest force multiplier: the people of the community (DOJ, 2015: 43).
>
> -- Police Chief J. Scott Thomson, Camden County, New Jersey

In the United States, the relationship between police and the communities they serve has been unstable and frequently fractious – giving rise to political advocacy, civil litigation, policymaking and policy revision. Contributing to these complexities, aside from a few exceptions, Partnership Policing has been virtually non-existent in urban settings nationally. The use of social media and the rise of often militant community advocacy has further intensified historically based conflicts between law enforcement and communities of color. There is a profound need for both a new paradigm and replicable models to promote public safety and truly expand the meaning of "police-community partnerships."

To answer that need, the Los Angeles Police Department Community Safety Partnership (CSP) has emerged as an innovative model of law enforcement that pushes beyond traditional policing frameworks. Former Police Chief William J. Bratton created the foundation for this model by reducing Department alienation from poor, high crime neighborhoods, ultimately increasing their safety and strengthening the trust between them and law enforcement. Once Bratton moved on, his successor, Chief Charlie Beck built on that foundation and created CSP to demonstrate what he termed "relationship-based policing" to radically reduce rampant gang violence that had impacted residents of four City public housing developments.) While CSP is a police-driven model, overseen by LAPD leadership, from the onset it was based on equal partnership and engagement with community leaders, schools, nonprofits, gang interventionists, philanthropic foundations, and the Urban Peace Institute. Most significantly, it was implemented in public housing developments dominated by multi-generational and violent street gangs in two diverse Los Angeles communities: Boyle Heights and Watts. The proposed research study offers a rigorous examination of this new model of partnership policing, evaluated from the points of view of both police and community.

To ensure scholarly rigor as well as validity and reliability, the proposed research will employ a mixed methods design, using both quantitative and qualitative measures that draw upon innovative approaches including real-time crime analysis, photo-voice, and continuous statistical analysis. Additionally, program sites will be contrasted with matched comparison sites to accurately assess the specific impact of the Community Safety Partnership. This combination of approaches will enable researchers to integrate traditional analysis of crime data and trends along with measuring outcomes that cannot be captured in crime statistics. It is critical for research to evaluate outcomes beyond decreases in crime – to understand the process of trust building within communities of color as well as the transformation of the narrative that traditionally dominates law enforcement.

In examining the Community Safety Partnership, this study will conduct a comprehensive review of the Los Angeles Police Department's stated changes in its practice and its policies, asking whether CSP's trust and relationship based policing works to:

1. Reduce violent crime and increase perceptions of safety?
2. Help residents reduce toxic neighborhood conditions that fuel violent crime and increase conditions that address and reduce trauma?

And, if yes to the two questions above:

3. Does CSP offer a more effective crime fighting and violence reduction model
4. Can the CSP model be taken to scale to transform police culture from warrior policing to partnership policing?

Literature Review

The research literature demonstrates that community violence is symptomatic of social marginalization and inequity. These factors are not the focus of the CSP, but it is important to understand how these factors are hovering "on the margins" and indirectly affect law enforcement interventions. Primarily occurring in communities of color, this marginalization engenders underfunded schools, economic deprivation, unequal access to mental and physical health services – a perfect storm of risk factors leading to the prevalence of violence and criminal activity (Cohen and Swift, 1993). Scholars also note that this can become a self-perpetuating cycle, with violence often preventing children's ability to walk to and from school, blocking new businesses from establishing themselves in communities, decreasing property values, and exacerbating the breakdown of social networks (Pinderhughes et al., 2015). Additionally, while property and violent crime has decreased nationally, there are still individual communities that have not experienced this diminution of crime (Department of Justice, 2015a) remaining locked in a cycle of violence. However, these same scholars argue that this cycle of violence can be interrupted, often most effectively through a community-based public health approach (Cohen and Swift, 1993; Vine, 2010; Pinderhughes et al., 2015). This holistic approach to community health, which works to ensure all have safe spaces to live, work, learn, eat, and play, catalyzes the relationship-based model of policing that seeks to reduce crime by building community alliances that work on enhancing neighborhood assets and health. The understanding and implementation of this model has been a critical part of a major shift in the narrative of modern law enforcement.

This perspective is central to the report issued by President Obama's Task Force on 21st Century Policing, which shifted the focus of crime reduction methods and units of measurement from crime statistics to "procedural justice, authentic relationships with community members, and sustained commitment to improve the health and well-being of the community" (Rice and Lee, 2015). Increased tension between police departments and communities coupled with the unequal distribution of crime, have brought law enforcement to a "tipping point." Many police departments have begun incorporating relationship-based policing principles into training and practice, most notably the Los Angeles Police Department. While the LAPD's past efforts at community policing have been well intended, but often limited in scale, insufficiently supported or in certain cases, inadequately implemented, in 2012 the creation of the Community Safety Partnership (CSP) marked the beginning of a much more extensive commitment to and investment in transforming LAPD's policing culture. Originally conceived as a pilot project, this relationship-based policing program was instituted at three housing developments in Watts and one in Boyle Heights, two of the city's most violent and gang-dominated neighborhoods that have both endured historically troubled relationships with law enforcement. The origins and development of the Community Safety Partnership are covered in the 2015 report for the President's Task Force on 21st Century Policing authored by Constance Rice and Susan Lee, entitled "Relationship Based Policing Achieving Safety in Watts."

In the CSP model, officers are responsible for working with myriad stakeholders, including residents, community based organizations, the Housing Authority of the City of Los Angles, local schools, and gang interventionists. The CSP model, which The President's Task Force on 21st Policing cited as a basis for its recommendations, draws upon scholarly writings, sources united in the view that officers' ability to build relationships with community members is crucial to actually protecting them (Bullock, 2013; DOJ, 2015). This is a developmental process as officers participate in community life: attending local meetings and church services while constantly engaging in conversations with residents that focus on how officers partner with and serve residents (and not primarily on criminal activity). These efforts with residents help build the trust necessary for relationship-based partnership policing to succeed (Torres, 2017). Additionally, officer demographics currently reflect community demographics, which studies reveal helps to further build trust. (Jennings et al., 2015). Officers involved in CSP develop the skills necessary to build and foster this trust through an innovative and on-going training program, 40% of which is led by the Urban Peace Institute (UPI). As UPI and multiple scholars note, the LAPD's decision to outsource training to an external agency is indicative of its commitment to program success, demonstrating its faith in the expertise of community organizations, especially because law-enforcement agencies are traditionally quite protective of their training programs (Goetz, 2003; Rice and Lee, 2015).

Law enforcement agencies are stereotypically organizations of tradition. Thus, it can often be difficult to introduce a new culture of policing to officers as well as leadership. However, integrating community-based policing initiatives into every level of a police department is essential to the endeavor's success. The need to institutionalize the culture of community partnership policing is well demonstrated in the research literature, although scholars also warn that this must be done incrementally and organically, to avoid the backlash and rejection that results from transformation by edict (Goetz, 2003; Torres, 2017). Therefore, the lessons learned from the incremental introduction of the CSP can serve as a blueprint for introduction and acceptance by police departments across the county. A rigorous evaluation of the efficacy of the Community Safety Partnership is vital to this process.

Key Research Questions

The research will be guided by certain key questions – designed to provide a rigorous evaluation of the Community Safety Partnership. However, there is added value to this research project, because its findings will help determine whether CSP is effective enough to warrant replicating in other LAPD units and other police departments. The questions will be further developed in partnership with the Los Angeles Police Department as well as a community advisory board, but the following serve as a working guide to the foci of the evaluation research.

What is the history of the development of the Community Safety Partnership?
    How as it evolved over time in different sites?

What are the major goals and objectives of the CSP?
How have these goals and objectives been implemented?
    How has implementation varied at each site?
    What has happened to implementation over time?

How does CSP work programmatically?
    Who are the key partners?

What are the costs – both direct and indirect of the program?
What is the evidence of collaboration?
What are obstacles to collaboration?

How is leadership enacted within the CSP?
What is done to ensure continuity as leadership changes?
How is the program being culturally sustained and institutionalized?
What is being done to build community-based leadership?

What are residents' perceptions of and reaction to CSP?
How do these compare to past police-community relationships?
How have these perceptions and reactions changed over time?
What are residents' perceptions and beliefs regarding the future?

How is CSP affecting various measures of crime, violence and community health?
Review of past and present crime statistics, to include multiple indicators such as arrests, use of force, officer involved shootings and other unobtrusive measures
Review of past and present school attendance and engagement
Review of past and present youth involvement in juvenile justice system
Review of past and present measures of community health

Evaluation Overview

The Community Safety Partnership evaluation envisioned by the UCLA Social Justice Research Partnership will be both high participatory and dynamic. All study materials will be designed to uncover LAPD and resident experience. The UCLA team will work to examine both the CSP narrative and outcomes. The focus of this evaluation effort will be developed in collaboration with the LAPD and evaluation advisory board. Some of the main points that may be used to guide this participatory research effort include:

- Understanding the Community Safety Partnership;
  - What is the CSP?
  - What are its roots and dynamics?
  - Does CSP work?
- Mapping a history of the CSP: focusing on its development as an initiative and intervention;
- Creating a history of engagement and accomplishment at different sites;
- Engaging and interviewing LAPD staff with different perspectives and experiences;
- Delineating initiative resources and mission;

- Examining disparate levels of relationships with community based organizations and how these are addressed;
- Documenting the process of relationship building – successes and challenges – in the varied settings in Los Angeles, applying these to differing US urban regions.

Based on these points and drawing upon network theory, our work will be guided by the effort to gain a better, more holistic understanding of how the CSP functions, utilizing:

- Depth interviews, and/or focus groups, with leadership and officers from CSP as well as outside leaders and residents.
    - The number of depth interviews will be dependent on budget constraints, depth interviews are often effective augmented with focus groups that address cultural relationships and trust issues. Evaluation team members will develop relationships with both LAPD and residents, an optimal method for data collection.

The research process will be both participatory and intentional. Interview and documentary narrative protocol questions will be constructed to elicit both LAPD and resident experience. This applied research endeavor will be designed to determine whether CSP helps residents change their view of the LAPD and their own safety. The evaluation team will also work to link these efforts to developing CSP interventions in other settings. Once again, dependent on budget, the CSP Evaluation Project will include the following foci:

- During the interviewing phase, UCLA anticipates collecting multiple accounts of the CSP that will merit further discussion and development. Once initial depth interviews and focus groups are conducted, there should be facilitated group conversations with CSP leadership and staff to discuss initial research results and to develop next steps for examination, with the ultimate evaluation report in mind.
- Interviews and/or focus group discussion will include but will not be limited to questions about CSP structure, roles, responsibilities, its leadership and recruitment model, community based strategies and LAPD-resident interactions.
- Once the depth interviews and focus groups are completed, the UCLA research team in collaboration with CSP and residents, will develop ideas for future questions and outcome measures that, dependent upon budget and timing, will be discussed in follow-up focus groups.
- The potential use of Photovoice and other media methods will be explored with CSP and residents. If there is a useful, “fiscally sensitive” way to augment the documentary narrative and evaluation with media, the LA team will conduct media-based research to include with the evaluation.

- The documentary narrative will be developed through a collaborative process, involving the UCLA team and CSP leadership staff in an ongoing exchange. This will be facilitated through in-person as well as virtual meetings and dialogue.
- Concurrent with these processes, surveys of residents will be conducted to collect data on community attitudes, responses to CSP, observations surrounding public safety, sense of well-being, cultural conflict and related topics.
- Quantitative data surrounding crime rates, gang activity, demographic trends, school participation and drop alongside related indicators will be collected and analyzed continuously throughout the evaluation period.

The scope of work to be developed for CSP evaluation will consist of multiple potential products. These may include but are not limited to an *Ethnographic Interview and Focus Group Analysis,* a *Documentary Narrative, Analysis of Crime Trends* and *Outcome Evaluation*. The methodologies are described in the next section and include:

- Documentary Narrative
- Ethnographic/Depth Interviews[1]
- Focus Groups
- Photovoice[2]
- Crime Statistics
- Quantitative Measures

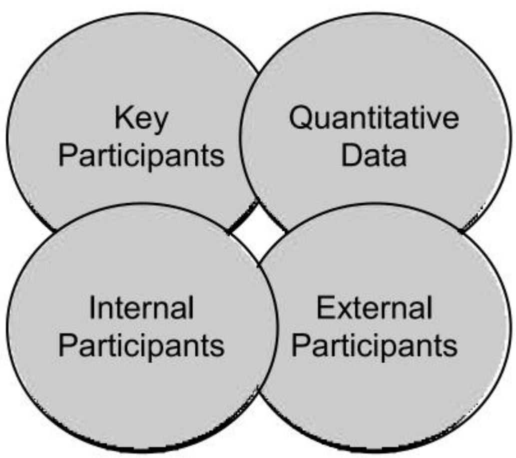
Information Categories

[1] These interviews are based on unstructured interview protocols composed of "open ended" (as compared with close-ended, forced choice) questions that allow individuals to express thoughts and offer narratives of their experiences and observations. Such interviews are then coded and analyzed for key themes.

[2] Photovoice is an innovative research method developed specifically for use in community-based research. Participants express their thoughts or represent their communities by photographing scenes that highlight research themes. These photographs are collaboratively interpreted through group discussions and narratives are developed to explain how photos highlight research themes.

Key participants will consist of those individuals and/or groups who are part of organizations or jurisdictions that have received services, consultation or technical assistance as part of the CSP. This group will consist primarily of community residents in the four housing developments.
Internal Research Participants are individuals who are currently serving or have served as part of CSP leadership and staff or who have been affiliated with CSP efforts. This group will consist primarily of the LAPD leadership and staff. It will be equally important to interview LAPD officers, other agency participants and residents who oppose CSP.
External Participants are individuals with knowledge and experience with CSP but are neither current nor former LAPD leadership and staff nor community residents. This group will be composed of individuals and organizations that possess insights and knowledge integral to CSP efforts, most significantly the Urban Peace Institute.
Quantitative Data will consist of statistical and demographic information regarding CSP, residents and settings the intervention has served.

The relationship between these Information Categories is represented in the following diagram:



Proposed Methodology

To explore the research questions, understand the CSP implementation process and record community change and meaningful outcomes as well as demonstrate the actual relationship between trust building and crime reduction, it will be critical to employ a broad range of research methods. While some of these strategies (analysis of crime statistics, interviews and focus groups) represent traditional methods used in examining and evaluating citywide public safety efforts (Cite GRYD materials here), there are additional, innovative approaches that will be part of the research. These will enrich and extend the research descriptions and findings, ensuring that this evaluation effort will be both comprehensive and meaningful.

Depending on the budget, data will be collected at either two or all four of the CSP sites in Los Angeles. In Watts, this will consist of the three housing developments operated by the Housing Authority of the City of Los Angeles at Imperial Courts, Jordan Downs, and Nickerson Gardens. In Boyle Heights, this will consist of the housing development operated by HACLA at Ramona Gardens. In addition to the four sites, the evaluation team will collect data at two *comparison* sites that that match the four sites closely but that do not receive any CSP or GRYD intervention. By matching the two comparison sites for gang activity, school climate, resources available, population composition and other demographic variables, the evaluation can assess meaningful differences attributable to the implementation of the Community Safety Partnership. Every effort will be made to find comparison sites within Los Angeles County – even if these lie outside the city of Los Angeles. Additionally, if comparison sites cannot be found in Los Angeles County, the research team will then work to find geographically accessible comparison sites – either in Orange County or in the tri-county "Inland Empire" which closely mirrors Los Angeles County both demographically and in terms of its problems with gang violence.

Documentary Narrative

To systematically record and present the history of the Community Safety Partnership (CSP) it is critical to create a documentary narrative. Using extensive document review, depth interviews and construction of an oral history with all individuals involved with the formation and implementation of the CSP, a narrative will be completed. All documents will be coded and integrated with interviews and oral history materials. It will be essential to record and create an archive of the genesis of the CSP. The narrative that is produced from this effort will offer both a history and a preliminary blueprint for scaling CSP interventions both within and outside of Los Angeles.

Ethnographic/Depth Interviews

Ethnographic or depth interviews will be conducted over a one year period. Key individuals from the Los Angeles Police Department (LAPD), involved stakeholders and residents will be interviewed over this one year period to track their experiences, chronicle change within the CSP, document their perceptions of public safety and well-being and other measures to be determined. The residents who are interviewed will be a combination of those who have been actively involved with CSP at each site as well as uninvolved residents who have been external to CSP.

Focus Groups

Focus groups will be conducted at each of the four sites with youth, families and other residents. In addition, focus groups will be conducted with service providers including community-based nonprofit organizations, medical practitioners, GRYD prevention and intervention workers and entities such as the Watts Gang Task Force. These focus groups will be conducted with cultural competence, linguistic fluency, guarantees of confidentiality and of safety, including off-site arrangements where needed.

Photovoice

This innovative research method combines photography with social action. Participants will be asked to express their view of or history with CSP by photographing scenes that represent research themes including community health, community assets, CSP impact, LAPD roles and sense of safety and well-being. These grass roots photographers are then interviewed to explain how their photographs portray a research theme, offering unique insights about CSP effectiveness and outcomes

Community Surveys

To ensure the collection of large-scale community data, community surveys will be developed with the help of the community advisory group. Residents in both the Watts and Boyle Heights communities will be surveyed to assess their attitudes regarding public safety, the LAPD, and the impact of the CSP.

Crime Statistics and Quantitative Measures

Drawing upon statistics available through the LAPD as well as (potentially) County-involved agencies including the Department of Children and Family Services, Department of Mental Health, and Probation Department, measures of crime and community health will be charted and analyzed before and after the implementation of the CSP. Additionally, measures from various educational agencies and departments (e.g., LACOE, LAUSD, and Charter School Organizations) that serve the CSP sites will also be collected, coded, and analyzed. While these findings may be largely suggestive, combined with the qualitative research materials, they will create a multi-faceted portrait of the impact of the CSP and its potential for change if brought to scale throughout Los Angeles.

Community Involvement and Representativeness

The primary goal of this evaluation will be to create a collaborative research effort involving the LAPD and the Watts and Boyle Heights communities to assess the impact of the Community Safety Partnership and the developing relationship between the community and the LAPD. There are multiple steps involved this approach which are detailed below.

Sampling

To ensure that the data collected as part of this project is representative of the community, the evaluation needs to understand what will constitute ‘representativeness’ in the results. This can be accomplished through a variety of methods including the use of existing census track data for each of the affected communities. Additionally, the communities may have race and ethnicity and additional demographic data at the city level that is more current than the census. Once this information is available a more detailed plan for data collection will be developed in collaboration with the community and law enforcement.

Survey Instrument design

This will proceed initially along two parallel tracks involving interested community groups and law enforcement. The goal is to identify the domains that different stakeholders in each group believe are important to assess. Once each track has progressed to the point of having concrete ideas and questions, stakeholders (designated as a research advisory group) will meet to decide on the final set of questions and relevant details including whether some of the questions might differ by community. The actual survey must be sensitive to length and the context of the community. This will entail translations of the instrument into Spanish and, if necessary, an additional language.

Data collection Methods

The research will be sensitive to reaching a representative sample of the people in the community; to achieve this, active participation by the various stakeholder groups will be essential. As part of this, we will engage in several different data collection methods. The survey can certainly be made available online and circulated by community partners through social media. This may be an appealing format for young adults and others since they will have the ability to respond using a smart phone, tablet or computer. The electronic version of the instrument can also be emailed to participants allowing them to participate that way or be listed on community organization websites or distributed through the local faith community. However, there is the very real problem of the digital divide in both Watts and Boyle Heights. To collect data from "harder to reach" groups –particularly for interviews and focus groups -- the evaluation team will work closely with nonprofit organizations and resident leaders to ensure that the study is authentically representative. Additionally, community members may also be involved in the research process. Experience has demonstrated that evaluation is most effective when residents and community members are partners in the data collection process, including helping to ask questions, administer surveys door-to-door, and co-facilitate focus groups. Additionally, this strategy has been shown to enhance the response rate of those that are being surveyed. This process, while supporting the collection of the necessary information, has the distinct advantage of providing the community a chance to be vested in the process and by extension, the outcome.

Data Sharing

The research team will develop an easy to use, dynamic website for the evaluation that will serve two on purposes. First, through a private intranet site, it will facilitate interaction among the evaluation team and key leaders of the CSP, which may include members of the Los Angeles Police Department and community stakeholders as they monitor the impact of the programs. Second, the research team will design data dashboards, some for private (internal) use and others for public consumption as CSP deems appropriate. The data dashboard is intended to provide information about successes and identify areas for improvement by aligning important metrics of success with the project goals. Users will be able to select information for specific sub-groups such as age, gender, and generate reports based on an array of performance indicators. At the end of the evaluation, the public will have access to a defined set of dashboard data, which will be determined by the evaluation team and the CSP. A password-protected segment of the dashboard, accessible only by authorized users (e.g., CSP staff, funders etc.) will contain more sensitive data.

The internal site will be accessible only to key CSP team members. It will help to centralize research processes and reduce ambiguity about the implementation of the evaluation. The site will serve as collaborative workspace, to vet data collection instruments, reports, and other evaluation products for comment and revision by all stakeholders. Authorized CSP team members will have access to their data that are collected as part of the evaluation, which can be used for their own internal evaluation and research purposes.

UCLA Social Justice Research Partnership

The UCLA Social Justice Research Partnership (SJRP) is a cross-disciplinary research, evaluation and policy collaborative, composed of a multi-ethnic staff with varied educational backgrounds and areas of expertise. As a collaborative, SJRP possesses extensive experience in research and evaluation in the fields of criminal justice, public health, education, and social welfare focusing on community wellness, violence prevention, government-community collaboration, and social policy. The Project Leadership Staff and advisory groups will all draw from the considerable resources of both UCLA and the Social Justice Research Partnership. Its approach is described in detail in Appendix A. The following individuals will comprise the Project Leadership Staff:

Jorja Leap, PhD, MSW

Dr. Jorja Leap has been a member of the faculty of the UCLA Luskin School of Public Affairs since 1992. Dr. Leap's research examines gangs, the criminal justice system and the dilemmas faced by the formerly incarcerated. As part of these efforts, Dr. Leap is currently the Qualitative Research Director of the Los Angeles Mayor's Office Gang Reduction Youth Development (GRYD) program and a member of the California Board of State and Community Corrections (BSCC) Gang Standing Committee. Additionally, Dr. Leap is affiliated with The California Endowment, as an Evaluation and Learning Specialist for its Building Healthy Communities (BHC) and Sons and Brothers Initiatives. She is also involved in research and community building efforts in South Los Angeles and is currently a lead member of the multi-disciplinary team implementing the parenting program, Project Fatherhood, in the Jordan Downs housing project of Watts. Most recently, along with Karrah Lompa, she established the Watts Leadership Institute, a 10-year initiative dedicated to community capacity building and violence prevention. Dr. Leap has conducted numerous community-based participatory research studies and evaluations. In her recent writings, Dr. Leap has completed a chapter on community-based responses to gangs in the book *Changing Course: Preventing Gang Membership,* published jointly by the National Institute of Justice and the Center for Disease Control as well as a chapter on Gangs, Violence and Drugs for the volume, *Violence: A Global*

Health Priority published by Oxford University Press in November 2014. Her most recent book *Project Fatherhood: A Story of Courage and Healing in One of America's Most Troubled Communities* was published by Beacon Press (June 2015). All proceeds from this book go to Project Fatherhood Jordan Downs. She also authored *Jumped In: What Gangs Taught Me about Violence, Love, Drugs and Redemption* published by Beacon Press (2012), with all proceeds going to Homeboy Industries.

P. Jeffrey Brantingham, PhD

Dr. Brantingham is Professor of Anthropology at the University of California Los Angeles and an expert in criminal behavior and policing. His research interests lie in the study of human behavior in complex environments, including offender mobility, offender target selection and the organization of criminal street gangs. Jeff directs the UC Mathematical and Simulation Modeling of Crime Project (UC MaSC), a collaboration between mathematicians, social scientists and law enforcement agencies aimed at understanding criminal behavior and crime hotspot formation. The goal of this research is to seek model-driven crime prevention and policing strategies. Dr. Brantingham is currently part of the evaluation research team assessing the Los Angeles City Gang Reduction and Youth Development (GRYD) Initiative. He has published more than 50 academic journal articles. His work on the mathematics of crime and predictive policing has received widespread media coverage including features by the Associated Press, New York Times, Los Angeles Times, The Economist, and BBC, NBC and CBS News. Additionally, His work on real-time crime prediction has led to deployments in more than 60 police departments around the world. Dr. Brantingham is a co-founder of PredPol, a company dedicated to delivering real-time crime predictions to law enforcement agencies.

Todd Franke, PhD, MSW

Dr. Franke is Professor of Social Welfare at the University of California Los Angeles and is currently the Chair of the UCLA Department of Social Welfare. He is a nationally recognized expert in research and evaluation methodology, with a specialized focus on quantitative analysis. Dr. Franke has worked with Dr. Leap and Ms. Lompa on previous and ongoing community-based research efforts, including the Homeboy Industries evaluation, the California Community Foundation Building a Lifetime of Options and Opportunities for Men (BLOOM) Initiative and The California Endowment Building Healthy Communities (BHC) Initiative. Throughout his tenure at UCLA, Dr. Franke has been involved with agencies that serve thousands of youth and families representing unique geographic and cultural communities in California, particularly Southern California counties. He has numerous years of experience conducting cross-sectional and longitudinal research in the fields of education (Los Angeles Unified School District –School Mental Health), child welfare (Los Angeles County Department of Children and Family Services) and juvenile justice (Los Angeles County Probation

Department), particularly for gang involved youth and reentry populations. He is currently examining mental health services in the State of California for the State Department of Mental Health.

Gerald Chaleff

Gerald Chaleff served as the Special Assistant for Constitutional Policing and Legal Counsel to William Bratton and Charlie Beck, Chiefs of Police for the Los Angeles Police Department. As a leading expert in Federal Consent Decrees, he oversaw the LAPD Consent Decree from inception to conclusion, over nearly a decade. Mr. Chaleff was appointed to the LAPD by Chief William J. Bratton in January 2003 and served as the Commanding Officer of the Consent Decree Bureau, overseeing the implementation of the more than 200 provisions of the LAPD Consent Decree. Prior to his professional work with the LAPD, Mr. Chaleff was appointed to the Los Angeles Board of Police Commissioners in 1997, and elected as President of the Board from 1999 to 2001. While President of the Police Commission, Mr. Chaleff negotiated the terms of the Consent Decree with the U.S. Department of Justice, the City of LA Mayor's Office, the LA City Attorney and members of the LA City Council. Mr. Chaleff is also a former President of the Los Angeles County Bar Association and served as Deputy General Counsel to the Webster Commission, which examined the LAPD's response to the civil unrest of 1992. He previously worked for the Los Angeles County District Attorney's Office and the Public Defender's Office, followed by several years in private litigation. In private practice, Mr. Chaleff became a nationally recognized expert in criminal defense, in State and federal court, and has been elected to the American College of Trial Lawyers and the Chancery Club. He is a former partner of Orrick, Herrington & Sutcliffe, LLP, where he chaired the white-collar criminal defense practice group. Mr. Chaleff has extensive trial experience and has provided legal analysis to a variety of news media, providing commentary and legal analysis on high-profile criminal cases. Mr. Chaleff received his Bachelors' of Science from the University of California, Los Angeles, and his Juris Doctor from Harvard Law School.

Karrah Lompa, MSW/MNPL

Ms. Lompa's work focuses on research, policy and evaluation in the fields of juvenile justice reform, program evaluation, community-based research and organizational development with an emphasis in nonprofit capacity building. She is currently a lecturer at the UCLA Luskin School of Public Affairs. Alongside Dr. Leap, in January 2016, she established the Watts Leadership Institute, a ten-year initiative to develop leadership, fundraising capacity, policy advocacy and communication technology among the indigenous leaders and small and struggling nonprofit agencies of Watts. Ms. Lompa remains an active member of a research-advocate effort to advise the Los Angeles County Probation Department on the rebuilding and reprogramming of Camp Kilpatrick, traveling to best practice jurisdictions across the country to inform this LA-based reform. Furthermore, Ms. Lompa led a team of qualitative researchers evaluating the implementation of the Children's Defense Fund's Freedom Schools in six LA County Probation camps. Ms. Lompa splits her time between her research and evaluation work in the community, serving as the chief administrator for the larger research and evaluation team and volunteering. Ms. Lompa has extensive experience in nonprofit leadership and administration. In her role as Executive Director of Free Arts for Abused Children she was responsible for significantly stabilizing the agency's financial position, diversifying streams of income, developing the board of directors and strengthening and growing programs.

Evaluation Advisory Group and Support

In addition to the evaluation research leadership team, the evaluation will be guided by an advisory committee to be comprised of Los Angeles Police Department Staff, key stakeholders and residents of the involved settings in Watts and Boyle Heights. In turn, the success of SJRP relies on having professional, positive and honest relationships with innumerable thought leaders and institutional representatives. Based on our past efforts, the research team will draw upon past key relationships with individuals who include:

- James Bell, The Burns Institute
- Father Greg Boyle, Founder, Homeboy Industries
- Susan Burton, Founder & Executive Director, A New Way of Life
- Dr. Christina T. Christie, UCLA School of Education and Information Studies
- Dr. Denise Herz, CSULA School of Criminal Justice and Criminalistics
- Supervisor Sheila Kuehl, LA County Board of Supervisors
- Julio Marcial, Program Director, The California Wellness Foundation
- Terri MacDonald, Chief Probation Officer, LA County Probation Department
- Mark Ridley-Thomas, LA County Board of Supervisors
- Dr. Robert Ross, President & CEO, The California Endowment
- Dr. Beatriz Solis, Regional Director, The California Endowment
- Anne Tremblay, Director, Mayor's Office of Gang Reduction and Youth Development (GRYD)
- Fernando Rejon, Executive Director, Urban Peace Institute
- Susan Lee, Founding Partner, Urban Peace Institute

It is critical to note that this is a preliminary list of individuals to be involved with the Advisory Group. It will also include additional members drawn from both law enforcement and the involved communities. There will also be an additional advisory group, composed of Community Intervention Workers who are part of the Los Angeles City Mayor's Office Gang

Reduction and Youth Development (GRYD) and who play an integral role in the Community Safety Partnership efforts.

Community Safety Partnership Evaluation
Preliminary Budget[3]

Personnel

$321,750

Jorja Leap, PhD Principal Investigator (20% FTE)

[3] This represents a base budget amount, to be adjusted dependent on the number of CSP sites to be studied in the final evaluation plan.

Jeffrey Brantingham, PhD Co-Principal Investigator (20% FTE)
Todd Franke, PhD Director of Research (20% FTE)
Karrah Lompa, MSW Director of Operations (20% FTE)
Gerald Chaleff, JD Project Advisory Consultant (5% FTE)
TBN, Project Director (75% FTE)
TBN, Research Associate (75% FTE)
Community Outreach and Interview Squad (5 people, 50% FTE each)

Personnel Benefits and Taxes **$96,525**

Computers/Technology/Software **$16,325**
iPads for Interview Squad (5)
Mobile "Hot Spots" (5)
Analytical Software

Research Incentives/Gift Cards **$10,000**

Mileage **$4,950**
Average 750/miles/month at .55

Supplies and Materials **$5,000**

Administration and Operations (10% of total) **$45,450**

TOTAL **$500,000**

References

Bullock, K. (2013). Community, intelligence-led policing and crime control. *Policing & Society, 23*(2), 125. Retrieved from https://search.proquest.com/docview/1356500392?accountid=14512

Cohen, L., & Swift, S. (1993). A public health approach to the violence epidemic in the united states. *Environment & Urbanization, 5*(2), 50-66. Retrieved from https://search.proquest.com/docview/61349630?accountid=14512

Davis, R. C., & Miller, J. (2002). Immigration and integration: Perceptions of community policing among members of six ethnic communities in central queens, new york city. *International Review of Victimology, 9*(2), 93-111. Retrieved from https://search.proquest.com/docview/60462289?accountid=14512

Goetz, B., & Mitchell, R. (2003). Community-building and reintegrative approaches to community policing: The case of drug control. *Social Justice, 30*(1), 222. Retrieved from https://search.proquest.com/docview/231910348?accountid=14512

Pelfrey, William Virgil, Jr. *The inchoate nature of community policing: Examining the differences between community and traditional police officers* Available from Sociological Abstracts. (60434651; 200220233). Retrieved from https://search.proquest.com/docview/60434651?accountid=14512

Scheider, M. C., Chapman, R., & Schairo, A. (2009). Towards the unification of policing innovations under community policing. *Policing: An International Journal of Police Strategies & Management, 32*(4), 694-718. doi:http://dx.doi.org/10.1108/13639510911000777

Torres, J. A. (2017). Predicting perceived police effectiveness in public housing: Police contact, police trust, and police responsiveness. *Policing & Society, 27*(4), 439-459. doi:http://dx.doi.org/10.1080/10439463.2015.1077837

Vine, M. M., Elliott, S. J., & Keller-Olaman, S. (2010). To disrupt and displace: Placing domestic violence on the public health agenda. *Critical Public Health, 20*(3), 339-355. doi:http://dx.doi.org/10.1080/09581590903419586

APPENDIX A

UCLA Social Justice Research Partnership

The mission of the UCLA Social Justice Research Partnership is to serve as a social justice research partner, conducting meaningful community-based participatory research to provide children, youth and families with the best opportunities for health and well-being, while strengthening communities and public policy.

At the local, state, national and international level, SJRP designs and develops participatory research and evaluation initiatives to help prevent violence, foster community capacity building, develop innovations in service delivery and positive systems change as well as inform and expand public policy. As part of these efforts, SJRP conducts applied research and evaluation to examine community practices and social policy to support the evidence base so best practices can thrive in frameworks that focus on results-based accountability.

Supported by a team of researchers trained in qualitative and mixed-methods, the SJRP research and evaluation draws upon a community-based model that is consistent with the philosophy, mission and aims of participatory action research. Our work emphasizes an approach that:

- Utilizes the unique insights and skills of those involved without overburdening direct service staff or stakeholders with excessive data demands;
- Helps document processes, significant themes and outcomes;
- Offers the most rigorous and culturally sensitive research strategy possible.

As part of this approach, the Social Justice Research Partnership relies upon the following key strategies:

1. *Best Practice Focus:* Our research and evaluation designs consistently support the use of quantitative and qualitative instruments and analysis that allow for both discovery and confirmation of findings. This focus is critical to producing information useful for improving local program effectiveness, and for providing evidence-based information concerning best practices and the contexts in which they are most effective to all organizations and stakeholders involved in the project. As part of these best practices, SJRP creates and conducts ethnographic observation, population surveys, case studies, data review and analysis, interviews, focus groups, “big data” analysis and other related research activities.

2. *Commitment to Usefulness:* This best practice focus goes hand-in-hand with a dedication to ensuring that findings are *actually used*. To ensure this, SJRP’s ongoing work is grounded in the real-world situations facing stakeholders, practices, policies and the larger system. Our team presents data and research findings in practical language with clear implications for future policy and program decisions. Consistent communication and practical feedback to stakeholders is central to usefulness.

3. *Participation and Responsiveness:* Useful evaluation must be grounded in the actual experience – a need we address in two ways. First, we ensure *meaningful participation in research and evaluation* through a strategy that actively engages a diverse group of

stakeholders involved in the effort. Second, we structure data collection and analysis to *measure and assess the strength and effectiveness of not only the outcomes but also the process and implementation challenges* relevant to the study. With this approach in mind, the evaluation team has long emphasized flexibility in practice, with a willingness to revise and adjust in the project plan as the situation may demand. SJRP does not believe in simply accepting a project and turning in a final report – we view ourselves as partners with the agencies or organizations that contract with us.

4. *Mixed-Methods and Measures:* Valid measurement is the bedrock of useful evaluation. Accurate, complete and comparable quantitative data is necessary to identify similarities and differences in process and outcomes. Support analysis identifies important contributors to outcomes across these different settings. Qualitative information revealing the perspectives, challenges and conclusions of stakeholders is critical to the evaluation endeavor. The evaluation team has extensive experience blending different data sources in mixed-methods measurement and analysis in a culturally sensitive manner. Just as importantly, the team has an ongoing record of effective collaboration with public, private and community partners. All evaluation plans and instruments are developed in collaboration with our community partners invested in the work.

5. *Using Variation to Strengthen Analysis:* A constant theme throughout our work is that there is complexity and diversity of social initiatives operating in a “real world” environment. This must be treated as an opportunity to generate findings that are relevant to users and stakeholders. Accordingly, SJRP’s research and evaluation design and methods are built to account for and explain the effects of this complexity, rather than to control or ignore it through simple aggregate analyses.

Because our work is not being implemented in a sterile laboratory – but rather in a complex social environment with numerous features that affect its success – SJRP utilizes a mixed-methods framework for our work. Our work is highly sensitive to, and values, diversity and culture in all aspects of the conceptualization and implementation of any project we engage in. Due to the diversity of organizations, partners and programs we collaborate with, every project and partner is different; working in collaboration is what determines the final strategy and plan. Clear communication and collaboration is essential not only during the planning phase of our work, but throughout. This forms the foundation for a rigorous, well-integrated, and responsive evaluation or research project as well as helping with future policy development and capacity building. All of our work is grounded in these collaborations, conversations and planning.

Population and Geography

With this purpose in mind, our evaluation and community-based research work is with nonprofit organizations, philanthropic institutions and public entities who want to better understand elements of their programming and/or their impact. The core of this work takes place throughout Los Angeles County, although we have also engaged in and continue to serve as a research partner on statewide efforts led by The California Endowment and The California Wellness Foundation. We are also currently involved in projects based in Central and Northern California as well as national policy efforts based in Washington, D.C. Additionally, SJRP has been engaged with projects based at University of St. Andrews in Scotland as well as the United Nations Violence Prevention Alliance. Our projects are often initiated by an organization, such as The California Endowment or by nonprofit organizations, such as Homeboy Industries, that seek to better understand an initiative, program or need. Always, the initiative or program is intended to understand the need of, or provide services to, marginalized and underserved communities. Therefore, the first level of our target population consists of nonprofit organizations, public agencies and institutions serving the community, but the secondary population served is comprised of the communities and their residents.

Strengths

The Social Justice Research Partnership team synthesizes the strengths of individual researchers who each bring unique training, experience and talents to our work and who are all committed to working collaboratively and in a participatory manner on all projects and assignments. The team has a combined record of over 75 years of community-based research, evaluation and nonprofit leadership, with most of those efforts in Southern California and focused on examining the efforts of various nonprofit organizations, philanthropic institutions and Los Angeles County departments, including Homeboy Industries, Project Fatherhood, Brotherhood Crusade, A New Way of Life, Beit T'Shuvah, The California Endowment, The California Wellness Foundation, the Los Angeles County Department of Probation and the Department of Children and Family Services, as well as the Mayor's Office of Gang Reduction and Youth Development.

Over the last four years, the number of esteemed institutions SJRP has engaged with has greatly expanded. Projects over this period have included developing, conducting and reporting on myriad process and outcome evaluations; serving as the Boyle Heights Learning and Evaluation Specialists for The California Endowment's Building Healthy Communities initiative; completing a robust landscape analysis and two case studies of The California Endowment's Sons and Brothers Initiative; providing ongoing technical assistance and evaluation of the implementation of the Board of State and Community Correction's Racial and Ethnic Disparities (RED) grant implementation in the Stanislaus County Probation Department; and establishing the Watts Leadership Institute, a ten-year initiative with inaugural funding from The California Wellness Foundation, to develop leadership, fundraising capacity, policy advocacy and communication technology among the indigenous leaders and small and struggling nonprofit agencies of Watts.

SJRP's fundamental philosophy is that those individuals with lived experience are the experts – they are the stakeholders with the greatest depth of knowledge regarding need, opportunity and solution. As a result, there is no aspect of this work that does not rely heavily on community engagement and input. Virtually every component of SJRP's work is built from ethnographic participant observation meaning the research team is immersed in the community – this level of community engagement is the strength of our work.

Violence interruption directly reduces gang retaliation

Authors: P. Jeffrey Brantingham[1]*, Baichuan Yuan[2], Nick Sundback[4], Frederic Paik Schoenberg[3], Andrea Bertozzi[2], Joshua Gordon[3], Jorja Leap[4], Kristine Chan[5], Molly Kraus[5], Denise Herz[5].

Affiliations:
[1]Department of Anthropology, University of California Los Angeles, Los Angeles, CA 90095, USA.
[2]Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA.
[3]Department of Statistics, University of California Los Angeles, Los Angeles, CA 90095, USA.
[4]Luskin School of Public Policy, University of California Los Angeles, Los Angeles, CA 90095, USA.
[5]School of Criminal Justice and Criminalistics, California State University Los Angeles, Los Angeles, CA 90032, USA.

*Correspondence to: branting@ucla.edu

Abstract:

Gang violence is propelled by retaliation. Spontaneous attacks resulting from chance encounters between rivals, or situational interactions that challenge gang territory or reputation can trigger cycles of tit-for-tat reprisals. Yet it has been difficult to determine if interventions that seek to reduce the likelihood of retaliation translate into lower rates of gang crime. Here we find spatial-temporal point process models can quantify the degree of retaliatory behavior of crimes in Los Angeles due to their joint causal dynamics. Quasi-experimental interventions in Los Angeles appear to cut in half the number of gang retaliations, above and beyond the effects of policing. Moreover, this model reveals the effect of treatment without needing control group in experiment.

One Sentence Summary:

Hawkes point process models fitted to violent crime occurrences in Los Angeles, CA, suggest that systematic interventions aimed to reduce the likelihood of retaliation appear to have substantial effect in diminishing the incidence of gang-related violent crime retaliations.

Main Text:

Gang violence is distinctive for its ability to trigger clusters of retaliatory crimes (*1, 2*). Challenges between gangs that threaten geographic territory or gang reputation can easily escalate to a shooting, while a shooting or homicide often demands retribution (*3, 4*) ultimately driving a sequence of tit-for-tat reciprocal attacks (*5-7*). Retaliation adds to the cumulative volume of violent crime and stronger retaliatory forces add more total crime (*8*). It is important therefore to evaluate whether targeted interventions intended to reduce gang violence can do so by disrupting the process of retaliation. Although several studies have examined the aggregate impacts of anti-violence programs, including efforts to disrupt street violence (*9-11*), the results have been mixed (*12-14*). Scant attention has been devoted to detecting and quantifying the direct impacts of interventions on the dynamics of gang retaliation. Here we show that such impacts can be estimated directly from crime event data using a unique multivariate statistical model.

Recent advances in statistical modeling of point processes reveal the dynamics of randomly occurring events characterized by self-excitation or contagion (*15, 16*){REF}.. Such models have conditional intensity (*17*):

$$\lambda(t, x, y) = \mu(x, y) + \sum_{t_i < t} Kg(t - t_i, x - x_i, y - y_i), \quad (1)$$

where λ is the infinitesimal rate at which events accumulate at any point in space-time, given the entire history of the process. The point process is controlled by a spatially inhomogeneous background rate μ, a productivity parameter K indicating the expected number of subsequent events triggered by any one event, and a triggering kernel g governing the rate of self-excitation as a function of the spatial-temporal distance from preceding events. Both μ and g are typically assumed to be non-negative, with g a spatial-temporal density function. In many applications, g is assumed to be exponential or Pareto {REF}. Model (1) provides an intuitive characterization of gang violence (*18*) that partitions the causes of crime into baseline structural processes such as simmering gang rivalries and contagious spread dominated by retaliation (Fig. 1A).

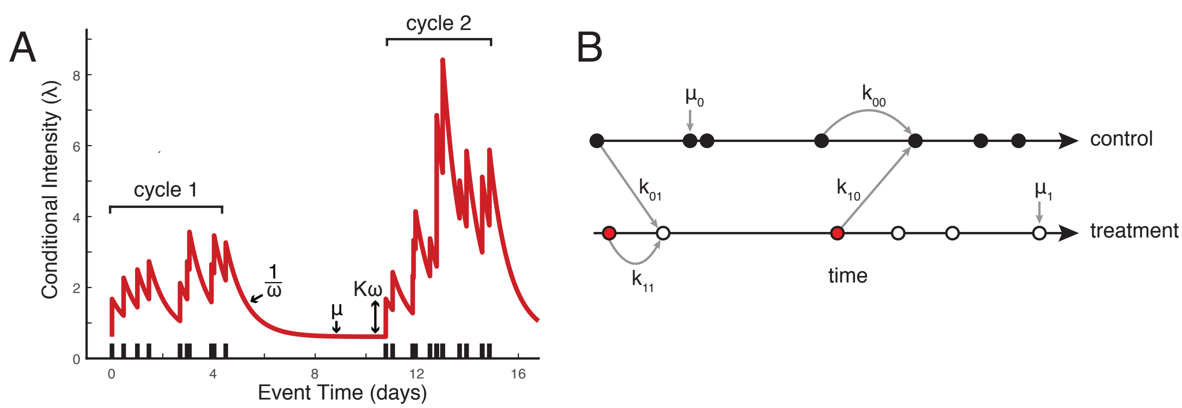


Fig. 1. Self-exciting point process models capture the dynamics of gang violence. (A). A temporal self-exciting point process model $\lambda(t) = \mu + \sum_{t_i<t} Kg(t - t_i)$ with exponential kernel $g(t) = \omega e^{-\omega(t-t_i)}$ fit to a sample of gang aggravated assaults and homicides in South Los Angeles from 2014-2015. Two tit-for-tat cycles of gang violence occur within a period of eighteen days. The conditional intensity λ reflects the instantaneous rate of gang crime. The background rate μ is expected rate of gang violent crime in the absence of retaliation. A crime causes the instantaneous rate to jump by an amount Kω, increasing the risk of retaliation. When elevated above μ, the risk of retaliation decays at a constant rate ω, with a mean lifetime of $1/\omega$. (B). Gang crimes assigned to two different experimental conditions are modeled as two interacting point processes. Non-retaliatory gang crimes assigned to each experimental condition arise spontaneously at a background rate described by μ_j. Retaliations assigned to each experimental condition may be triggered through separate pathways. Pathway k_{11} links one previous treatment crime to treatment retaliations. Pathway k_{01} a links one previous control crime to treatment retaliations. Pathway k_{10} links one previous treatment crime to control retaliations. Pathway k_{00} links one previous control crime to control retaliations. Events where treatment interventions can have an impact are marked in red. If treatment interventions reduce the risk of gang retaliation, then we expect the average number of events to satisfy $k_{11} < k_{01}$ and $k_{10} < k_{00}$.

We extend the above model to a multivariate framework (*19*) useful for describing real-world interventions. In such quasi-experimental field settings there is often imperfect separation between experimental conditions {REF}. To account for interactions we propose the conditional intensity:

$$\lambda_u(t,x,y) = \mu_u(x,y) + \sum_{t_i < t} K_{u_i u}\, g(t - t_i, x - x_i, y - y_i). \tag{2}$$

Here u_i is the type of event i where $u = 0$ represent an event assigned to a non-intervention control condition and $u = 1$ an event assigned to an intervention treatment condition. The model is easily modified to accommodate more than two interacting experimental conditions. The spatially inhomogeneous background rate of events is now partitioned according to condition u (supplemental text). The parameter $K_{u_i u}$ is the expected number of retaliations of type u triggered by an event of type u_i. Thus we have four productivity parameters to estimate $K_{u_i u}$ = k_{11}, k_{01}, k_{10}, and k_{00}, representing the four possible interactions between treatment and control conditions (Fig. 1B). If treatment interventions are effective, then estimated parameters should satisfy the inequalities $k_{11} < k_{01}$ and $k_{10} < k_{00}$. We estimate the model parametrically using an expectation-maximization (EM) algorithm (*20*) and confirm the chosen model form with non-parametric methods (*15, 21*) (supplemental text). The model fit is evaluated using Voronoi residuals (*22, 23*) (supplemental text).

We analyzed gang violence in a unique quasi-experimental setting in Los Angeles where there is approximately random assignment of gang crimes between two different, but interacting intervention conditions. The random assignment between the two conditions arises naturally out of the crime reporting system (see below). The control condition consists of gang violent crimes responded to by the Los Angeles Police Department (LAPD-only). The treatment condition consists of crimes responded to by the LAPD, but with additional notification of the Los Angeles Gang Reduction and Youth Development (GRYD) Intervention Response program (LAPD + GRYD IR). Upon receiving notification, GRYD IR tasks community intervention workers with disrupting opportunities for retaliation through rumor control and proactive peacemaking. We focus on crimes occurring in an 87.2 km2 (33.7 sq miles) area of South Los Angeles during 2014-2015 (Fig. S1). The ten GRYD IR Zones in South Los Angeles represent only 6.7% of the total land area of Los Angeles (1302 km2), but accounted for 45.3% of serious gang crimes city-wide in 2014-2015. We limit our consideration to aggravated assaults and criminal homicides, crimes which entail a greater risk of retaliation compared to other crime types (supplementary text).

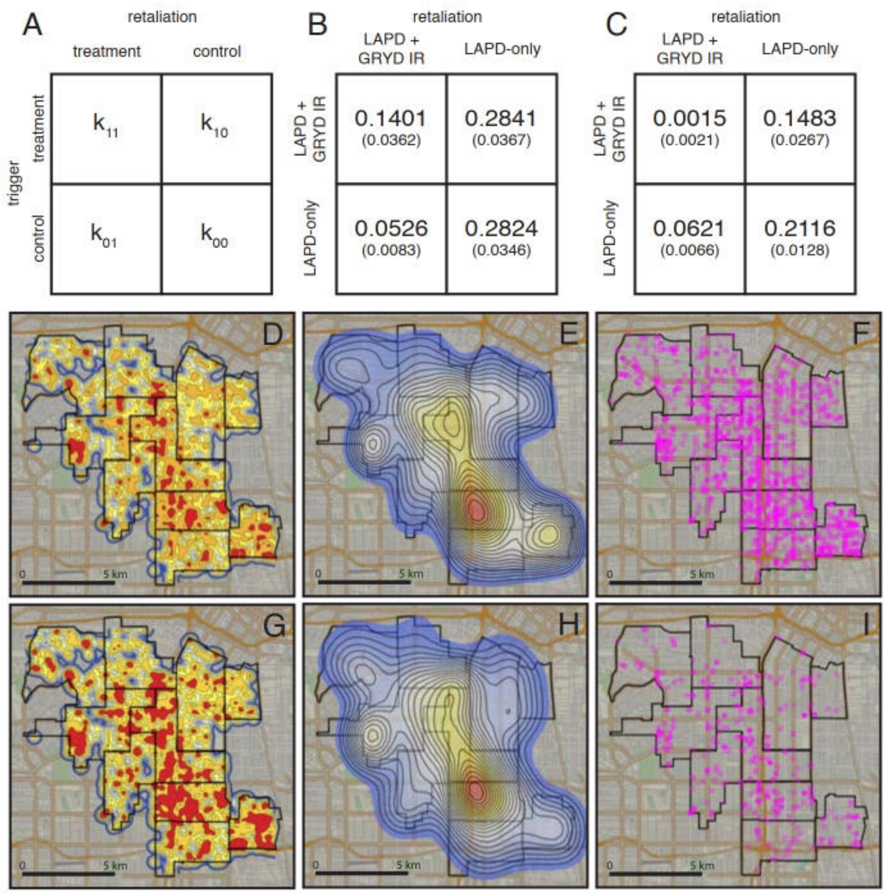


Fig. 2. (A) Matrix representation of productivity K_{ij} with the corresponding triggering pathways noted. Matrix entries are the average number of retaliations assigned to an experimental condition i triggered by an event assigned to condition j. (B). The productivity for combined gang and non-gang aggravated assaults and homicides in South Los Angeles for 2014-15. (C). The productivity for gang aggravated assaults and homicides in South Los Angeles for 2014-15. The control condition includes violent crimes known to the LAPD (LAPD-only). The treatment condition includes crimes known to the LAPD that were also reported to GRYD IR (LAPD + GRYD IR). Standard errors of parameter estimates are shown in parentheses. (D). The log of spatial-temporal background intensity function μ for gang violent crimes mapped over space. (E) Contour plot of the density of background gang aggravated assaults and homicides determined by declustering. (F) Point locations of background gang aggravated assaults and homicides determined by declustering. (G) The log of spatial-temporal triggering kernel g mapped over space. (H) Contour plot of the density of retaliatory gang aggravated assaults and homicides determined by declustering. (I) Point locations of retaliatory gang aggravated assaults and homicides determined by declustering. Boundaries for the ten GRYD IR Zones in South Los Angeles are outlined in black.

In 2014-15, a total of 5,928 aggravated assaults and homicides were reported to the LAPD in the South Los Angeles GRYD Zones, including both gang and non-gang crimes (Table 1). GRYD IR was notified in 9.9% of all aggravated assaults, but 71.5% of all homicides. Thirty-two percent of the total crime volume was identified as gang-involved. GRYD IR was notified in 27.3% of the gang aggravated assaults and 78.8% of gang homicides (Table 1). A key feature facilitating our analysis is that assignment of crimes to each of the two test conditions approximates a randomized experimental protocol. To verify we performed runs tests ($Z = 0.609$, $p=0.54$) and two-sample Kolmogorov–Smirnov (KS) tests (K-S=0.069, $p=0.20$ in time and K-S=0.054, $p=0.17$ in space), which found no discernable departures from the null hypothesis that the GRYD IR notifications are random, independent draws from the same population of gang-related crimes (supplementary text). Upon receiving a report of a gang crime, notification of GRYD IR proceeds as if a biased coin is flipped. If the crime is a gang aggravated assault, the coin is biased toward not notifying GRYD IR. If it is a gang homicide, it is biased towards notifying GRYD IR. Therefore potential treatment effects are not confounded with the process of GRYD IR notification.

Table 1. Gang and non-gang crimes reported only to the LAPD and to the LAPD and GRYD IR in 2014-15.

	LAPD-only				**LAPD + GRYD IR**				**GRAND TOTAL**	
	Gang N	**Non-Gang N**	**TOTAL N**	**% Gang**	**Gang N**	**Non-Gang N**	**TOTAL N**	**% Gang**	**TOTAL N**	**% Gang**
Aggravated Assault	1,249	3,918	5,167	24.2%	470	99	569	82.6%	5,736	30.0%
Homicide	41	29	70	58.6%	152	24	176	86.4%	246	78.5%
TOTAL	1,290	3,947	5,237	24.6%	622	123	745	83.5%	5,982	32.0%

GRYD IR starts at a significant disadvantage as an intervention strategy against violent crime in general. This is clear when we fit model (2), with an exponential kernel for g, to the combined gang and non-gang crimes (Fig. 2B). The estimate of parameter k_{11} indicates each aggravated assault or homicide exposed to the treatment triggered on average 0.1401 retaliations subsequently known to both LAPD and GRYD IR, whereas k_{01} indicates each control assault or homicide triggered on average 0.0526 retaliatory crimes known to both LAPD and GRYD IR. The 62.5% higher rate of retaliation accompanying GRYD IR notification is statistically significant ($p = 0.0092$) (supplemental text). Pathway k_{10} also shows that treatment crimes on average triggered 0.2841 retaliations known only to the LAPD, which is of equivalent magnitude to pathway k_{00} with 0.2824 retaliations known only to the LAPD ($p = 0.486$). In practical terms, every 100 LAPD + GRYD IR (treatment) aggravated assaults and homicides triggers on average 42.4 retaliatory violent crimes ($k_{11} + k_{01}$). Every 100 LAPD-only (control) violent crimes triggers an average 33.5 retaliations ($k_{10}+ k_{00}$). Table 1 shows that the mix of crimes faced by GRYD IR includes many more that are gang-related. GRYD IR may be able to reduce retaliation among the crimes it does confront, but not to the levels characteristic of violent aggravated assaults and homicide in general.

We therefore restricted analyses to gang aggravated assaults and homicides to ensure that the two test conditions confront events with similar potential for spawning retaliation. Against this set of crimes GRYD IR had a substantial impact (Fig. 2C). Pathway k_{11} triggered an average of 0.0015 retaliations for any one treatment gang crime. By contrast, pathway k_{01} triggered 0.0621 retaliations for any one control gang crime. This represents a 97.6% reduction in retaliation associated with GRYD IR notification ($p < 10^{-6}$). Pathway k_{10} triggered and average of 0.1483 retaliations for any one treatment gang crime. Pathway k_{00} triggered 0.2116 retaliations for any one control gang crime. The 29.9% reduction in retaliations is not as large, but is nevertheless significant ($p = 0.0163$). In practical terms, every 100 LAPD + GRYD IR gang crimes triggers and average of 15.0 retaliations ($k_{11} + k_{01}$). Every 100 LAPD-only gang crimes triggers and average of 27.4 retaliations ($k_{10}+ k_{00}$). Overall, the notification of GRYD IR was associated with a 45.3% decrease in retaliatory gang crimes.

To better understand the spatial dynamics of retaliatory gang violence we mapped the background intensity μ and triggering kernel g along with the distributions of background and retaliatory crimes determined via stochastic declustering (*24*) (supplemental text) (Fig. 2D-I). The background risk of gang violence is characterized by numerous compact, but widely distributed hot spots (Fig. 2D), consistent with the observation that the opportunities for violence and the strengths of gang rivalries are geographically variable (*25, 26*). The risk of retaliation is concentrated in more continuous bands (Fig. 2G), bridging discrete areas of background risk. Notably there is a prominent North-South corridor or retaliatory risk that maps to an area locally known as 'death alley' (*27*). The distribution of gang violent crime reflects these patterns of risk (Fig. 2F and 2I). The density of background crimes forms five distinct hot spots (Fig. 2E) suggesting that background crimes are of local, neighborhood origin. The density of retaliatory crimes occupies only two distinct hot spots (Fig. 2H) suggesting that retaliation spreads contagiously beyond immediate local environments.

Stochastic declustering (*24*) also allows us to evaluate differences in the frequency of retaliation by crime type across test conditions (Table 2). Background crimes make up 76.6% of all gang aggravated assaults and homicides for both test conditions combined. However, retaliatory crimes are proportionally more common among events assigned to the LAPD-only control condition. This imbalance is pronounced for gang aggravated assaults, (46.3% retaliation for LAPD-only vs. 10.3% for LAPD + GRYD IR), but particularly extreme for homicides (24.2% LAPD-only vs. 0.7% for LAPD + GRYD IR).

Table 2. Number of retaliatory and background aggravated assaults and homicides in South Los Angeles in 2014-2014 separated by test condition.

	LAPD-only				LAPD + GRYD IR			
	Retaliation	Background	TOTAL	% Retaliation	Retaliation	Background	TOTAL	% Retaliation
Aggravated Assault	395	854	1249	46.3%	44	426	470	10.3%
Homicide	8	33	41	24.2%	1	151	152	0.7%
TOTAL	403	887	1290	45.4%	45	577	622	7.8%

We use the estimated treatment effects along with the results of stochastic declustering in South Los Angeles to compute numbers of prevented crimes (Table 2). The sum ($k_{11} + k_{01}$) is the average number of retaliations known to both the LAPD and GRYD IR produced by two types of triggers (Fig. 1B). Similarly, ($k_{10} + k_{00}$) is the average number of retaliations known only to the LAPD produced by two types of triggers. Note that ($k_{11} + k_{01}$) and ($k_{10} + k_{00}$) are measured directly from data and therefore are observed outcomes. We now define two counterfactual cases. Let ($k_{01} + k_{01}$) be the average number of retaliations that would have been triggered in the absence of GRYD IR notification for those retaliations assigned to LAPD + GRYD IR. Let ($k_{00} + k_{00}$) be the average number of retaliations that would have been triggered in the absence of GRYD IR notification for those retaliations assigned to LAPD-only. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. From stochastic declustering, the observed number of gang retaliations arising from pathways ($k_{11} + k_{01}$) is 45 and from ($k_{10} + k_{00}$) is 403. The counterfactual conditions suggest that retaliatory gang crimes would have been 48.8% and 15.0% higher in the absence of GRYD IR for observed pathways ($k_{11} + k_{01}$) and ($k_{10} + k_{00}$), respectively (supplemental text). GRYD IR prevented an estimated total 82.2 retaliatory gang crimes, of which 77.8 are expected to have been aggravated assaults and 4.4 are expected to have been homicides. The benchmark estimates in (*28*) project the overall cost of a single aggravated assault at $240,000 and a single homicide at $8.98 million. Over the two-year period in 2014-15, the potential savings from GRYD IR in South Los Angeles alone is estimated at $49.0 million.

The multivariate self-exciting point process model presented here makes it possible to detect and quantify causal pathways connecting mixtures of events in unique ways. Instead of determining the causal structure of each event, productivities $K_{u_i u}$ are averaged over the entire time series, which produces very robust and stable estimates. This is particularly important in real-world empirical settings where putative control conditions are rarely, if ever isolated from putative test conditions. Rather control and test conditions are often mixed by structural and practical conditions beyond the control of the observer. The multivariate modeling framework embraces the fact that these mixtures exist and allows for interactions between conditions to proceed as part of the analysis.

References and Notes:

1. S. H. Decker, Collective and normative features of gang violence. *Justice Quarterly* **13**, 243-264 (1996).
2. M. W. Klein, C. L. Maxson, *Street gang patterns and policies*. (Oxford University Press, New York, 2006).
3. L. A. Hughes, J. F. Short, DISPUTES INVOLVING YOUTH STREET GANG MEMBERS: MICRO-SOCIAL CONTEXTS*. *Criminology* **43**, 43-76 (2005).
4. B. A. Jacobs, R. T. Wright, *Street justice: Retaliation in the criminal underworld*. (Cambridge University Press, 2006).
5. B. Bjerregaard, A. J. Lizotte, Gun Ownership and Gang Membership. *The Journal of Criminal Law and Criminology (1973-)* **86**, 37-58 (1995).
6. R. Rosenfeld, T. M. Bray, A. Egley, Facilitating violence: A comparison of gang-motivated, gang-affiliated, and nongang youth homicides. *Journal Of Quantitative Criminology* **15**, 495-516 (1999).
7. J. C. Howell, *Gangs in America's communities*. (Sage, 2012).
8. B. Green, T. Horel, A. V. Papachristos, Modeling contagion through social networks to explain and predict gunshot violence in chicago, 2006 to 2014. *JAMA Internal Medicine*, (2017).
9. W. G. Skogan, S. M. Hartnett, N. Bump, J. Dubois, *Evaluation of CeaseFire-Chicago*. (U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington, D.C., 2009).
10. D. W. Webster, J. M. Whitehill, J. S. Vernick, F. C. Curriero, Effects of Baltimore's Safe Streets Program on Gun Violence: A Replication of Chicago's CeaseFire Program. *Journal of Urban Health* **90**, 27-40 (2013).
11. J. M. Whitehill, D. W. Webster, S. Frattaroli, E. M. Parker, Interrupting Violence: How the CeaseFire Program Prevents Imminent Gun Violence through Conflict Mediation. *Journal of Urban Health* **91**, 84-95 (2014).

12. J. A. Butts, C. G. Roman, L. Bostwick, J. R. Porter, Cure violence: a public health model to reduce gun violence. *Annual review of public health* **36**, 39-53 (2015).
13. A. V. Papachristos, Too big to fail. *Criminology & Public Policy* **10**, 1053-1061 (2011).
14. M. W. Klein, Comprehensive gang and violence reduction programs. *Criminology & Public Policy* **10**, 1037-1044 (2011).
15. D. Marsan, O. Lengline, Extending earthquakes' reach through cascading. *Science* **319**, 1076-1079 (2008).
16. G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, G. E. Tita, Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106**, 100-108 (2011).
17. A. G. Hawkes, Spectra of Some Self-exciting and Mutually Exciting Point Processes. *Biometrika* **58**, 83-& (1971).
18. M. B. Short, G. O. Mohler, P. J. Brantigham, G. E. Tita, Gang Rivalry Networks Via Coupled Point Process Networks. *Discrete and Continuous Dynamical Systems* **19**, 1459-1477 (2014).
19. E. Lai *et al.*, Topic Time Series Analysis of Microblogs. *IMA Journal of Applied Mathematics,* **81**, 409-431 (2016).
20. A. Veen, F. P. Schoenberg, Estimation of Space? Time Branching Process Models in Seismology Using an EM? Type Algorithm. *Journal of the American Statistical Association* **103**, 614-624 (2008).
21. E. W. Fox, M. B. Short, F. P. Schoenberg, K. D. Coronges, A. L. Bertozzi, Modeling E-mail Networks and Inferring Leadership Using Self-Exciting Point Processes. *Journal of the American Statistical Association* **111**, 564-584 (2016).
22. A. Bray, K. Wong, C. D. Barr, F. P. Schoenberg. (2014).
23. J. S. Gordon, R. A. Clements, F. P. Schoenberg, D. Schorlemmer, Voronoi residuals and other residual analyses applied to CSEP earthquake forecasts. *Spatial Statistics* **14, Part B**, 133-150 (2015).
24. J. Zhuang, Y. Ogata, D. Vere-Jones, Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association* **97**, 369-380 (2002).
25. G. Tita, G. Ridgeway, The Impact of Gang Formation on Local Patterns of Crime *Journal of Research in Crime and Delinquency* **44**, 208-237 (2007).
26. A. V. Papachristos, Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. *Am. J. Sociol.* **115**, 74-128 (2009).
27. N. Santa Cruz, K. Schwencke, in *Los Angeles Times*. (2014), chap. The Homicide Report.
28. K. E. McCollister, M. T. French, H. Fang, The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and Alcohol Dependence* **108**, 98-109 (2010).
29. A. A. Braga, D. L. Weisburd, The Effects of 'Pulling Levers' Focused Deterrence Strategies on Crime. *Campbell Systematic Reviews* **6**, (2012).
30. G. Mohler, M. Short, P. J. Brantingham, F. Schoenberg, G. Tita, Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106**, 100-108 (2011).
31. M. B. Short, P. J. Brantingham, A. L. Bertozzi, G. E. Tita, Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences* **107**, 3961 (2010).

32. G. Mohler, Marked point process hotspot maps for homicide and gun crime prediction in Chicago. *International Journal of Forecasting* **30**, 491-497 (2014).

Acknowledgments:

This work was supported by the

Supplementary Materials:
Materials and Methods
Figures S1-S#
Tables S1-S#
Movies S1-S#
Audio Files S1-S#
External Databases S1-S#
References (##-##)

Supplementary Materials:

Materials and Methods:

Table of Contents

Materials

Data Sources. The analyses presented rely on data collected by the Los Angeles Police Department (LAPD) as well as data entered into the Gang Reduction Youth Development (GRYD) Efforts and Outcomes (ETO) Incident Response database. Data provided by the LAPD include only officially reported crimes that that have been through the Department's standard process of verification and quality control. Neither calls for service data, nor suspect and arrest data were used. The LAPD data includes records for all reported crime types ranging from public disorder to homicide. Most of these crime types are not directly relevant understanding the dynamics of gang violence and the impact of GRYD Incident Response (GRYD IR) (see below).

The GRYD ETO database tracks all instances where the GRYD Office was notified of a crime that could warrant GRYD IR intervention. The GRYD ETO database includes the following crime types:

- Homicide
- Multiple Victim Shooting
- Single Victim Shooting
- Stabbing
- Shots fired
- N/A or Unknown
- Other

These crime types do not align perfectly with LAPD (or California Penal Code) crime types. Single victim shooting, multiple victim shootings and stabbings in GRYD ETO data are all classified as aggravated assault (assault with a deadly weapon) in the LAPD data. GRYD crime types N/A or Unknown and Other can align in multiple ways with the LAPD. For events that have been successfully aligned with the LAPD data (see below) we adopt the LAPD crime type classification to ensure fair comparison with the entire LAPD dataset.

In addition to these basic pieces of information, the ETO database includes the GRYD's own evaluation whether an event is a gang crime. Such an evaluation is separate from, but likely influenced by the details about the event relayed to GRYD by the LAPD. We take an inclusive approach wherein we treat any event as a gang crime that either LAPD or GRYD labels as such.

Defining Gang Violence. The California Penal Code (CPC § 186.22) provides guidelines for attaching a gang enhancement to any crime. Any crime that is committed for the benefit of a gang is eligible to be labeled as a gang crime, and individuals convicted of committing such crimes may be subject to enhanced sentences. How a gang label is applied in practice by police to individual crimes, however, is not specified by the criminal code. In general, whether a crime is labeled as gang-related much depends upon the circumstances of the crime.

The gang-related label is complicated both by variation in the ways in which gang members and gang activity may be associated with an individual crime events and by potential variation in the criteria used by raters to determine if a crime qualifies as gang related. In general, crimes may be simply gang-affiliated, meaning that the victims and/or perpetrators of the crime are known or suspected gang members, but the crime itself is not a direct results of gang activities (*6*). By contrast, crimes may be gang-motivated, meaning that the crime was committed to further some social or economic goal of the gang, including establishing or maintaining reputation.

A range of criteria might be used by police or intervention workers to determine whether a specific crime is gang-related including basic intelligence about whether: (1) the event was gang motivated; (2) the event occurred in a gang area; (3) the event featured gang involved or affiliated individuals; (4) recent activity occurred between the victim's and/or suspect's gangs, or (5) the event has the potential for retaliation. Different raters may weigh these criteria to different degrees.

Among the crime events known to the Los Angeles Police Department (LAPD) in 2014-2015, 90 different crime types were labeled as gang crimes. A majority of the crime types so labeled may be simply gang-affiliated crimes where the offender happens to also be a gang member. Table S 1 shows the top 15 crime types sorted in descending order by the percentage that bear a gang label. Gang homicides made up 59.9% of all homicides in Los Angeles in 2014-2015. Gang aggravated assault (also sometimes referred to as assault with a deadly weapon) represented a greater number of events by volume (n =3,805), but constituted only 20.8% of all aggravated assaults recorded during this time frame. By contrast, 171 burglaries (0.6%) were labeled as gang crimes in 2014-2015. Burglaries are thus rarely labeled as gang crime and, in fact, there is little reason to think that even this handful of events had much to do with gang rivalries and the risk of retaliatory gang violence. Rather the suspects or victims in the reported crimes may have been gang-involved, which prompted the application of the label. Our analyses therefore focus on gang aggravated assaults and homicides, which are more frequently committed for the benefit of the gang and are particularly prone to initiating retaliation. We adopt an inclusive approach and treat any aggravated assault or homicide as gang-related if the LAPD or GRYD program labels it as such.

Table S 1. Top 15 gang crime types by volume in 2014-2015 for all LAPD data.

Rank †	Crime Type	Gang		Non Gang		TOTAL
		N	%	N	%	N
1	Criminal Homicide	325	59.9%	218	40.1%	543
2	Shots Fired at Inhabited Dwelling	174	35.9%	311	64.1%	485
3	Aggravated Assault	3,805	20.8%	14,500	79.2%	18,305
4	Discharge Firearms/Shots Fired	144	19.2%	605	80.8%	749
5	Attempted Robbery	294	13.8%	1,834	86.2%	2,128
6	Robbery	1,904	12.9%	12,889	87.1%	14,793
7	Violation of Court Order	419	11.1%	3,355	88.9%	3,774
8	Brandish Weapon	177	9.6%	1,659	90.4%	1,836
9	Criminal Threats - No Weapon Displayed	912	8.0%	10,551	92.0%	11,463
10	Other Miscellaneous Crime	171	4.5%	3,591	95.5%	3,762
11	Vandalism - Felony	639	3.0%	20,586	97.0%	21,225

12	Vandalism - Misdemeanor	307	1.7%	17,261	98.3%	17,568
13	Battery - Simple Assault	538	1.5%	35,550	98.5%	36,088
14	Spousal Abuse - Simple Assault	270	1.1%	24,115	98.9%	24,385
15	Burglary	171	0.6%	28,526	99.4%	28,697

GRYD and GRYD IR. The Los Angeles Gang Reduction Youth Development program is a comprehensive strategy aimed at reducing gang involvement and gang violence through the provision of gang prevention and intervention services, violence interruption activities and proactive peace-making. Prevention services are aimed at providing alternatives for youth before they join gangs. Intervention service are aimed at provide pathways out of gang life for youth that have already become involved. Violence interruption and proactive peace-making, are both aimed at mitigating the consequences of gang violence, including seeking to disrupt retaliation in the aftermath of gang violent crimes. This later function of GRYD is termed GRYD Incident Response and is the focus of our analysis. Established in 2009, the GRYD program design overlaps in part with Chicago's Operation Ceasefire (now Cure Violence) (*9*). GRYD is not a law enforcement strategy and therefore differs from the "pulling levers" model (*29*).

GRYD IR uses a near real-time notification system connecting the LAPD, the GRYD Program Office and intervention workers embedded in the communities. Notification and information sharing facilitates each of the principal parties taking largely independent actions in response to gang crime incidents. The LAPD follows standard law enforcement and investigative protocols in responding to gang crimes. The GRYD Program Office aggregates information and makes allocation decisions for non-law enforcement intervention resources. Community intervention workers (CIWs) engages the community through direct social contacts. Rumor control is seen by CIWs as a critical tool for reducing retaliation.

The use of CIWs to interrupt gang violence is controversial. Many CIWs are themselves former gang members. This known history provides CIWs credibility with the gangs, which they describe as a 'license-to-operate'. Typically, CIWs work with the gang sets or cliques from their home neighborhood and rarely with gangs that would be considered rivals. Thus, a CIW who was formerly affiliated with a Blood set would be very unlikely to provide violence interruption for a Crips set. Because of their credibility with their home gangs, CIWs argue that they are in a unique position to reduce the risk of retaliation. If a CIW asks a gang to 'stand down' there is a real chance that they will. The CIW 'license-to-operate' is a source of controversy, however. A reasonable concern expressed by law enforcement is that CIWs are *too close* to the gangs and therefore may be inclined to turn a blind eye to some gang activities even if they are criminal in nature. Some in law enforcement worry that CIWs are themselves involved in criminal activity with the gangs. The controversy surrounding CIWs means that the LAPD and CIWs, though aware of one another, operate largely independently in response to gang crimes. Our analyses do not address the concerns about CIWs. We simply ask whether the addition of GRYD IR, and any attendant CIW interventions, has a detected impact on gang retaliation.

South Los Angeles gang violence & GRYD IR Zones.

We focus our analysis on violent crimes occurring in an 87.2 km^2 (33.7 sq mi) area of South Los Angeles during 2014-2015 (Fig. S1). As of mid-2015, ten GRYD Zones were in operation in South Los Angeles forming a contiguous area of attention. Prior to mid-2015, only seven of the ten regions shown were formally recognized. Nevertheless, GRYD IR received notifications and responded to crimes over this entire geographic area throughout 2014 and 2015. For example, between July 1, 2014 and December 31, 2014, GRYD IR was notified of 109 events over the South Los Angeles region (Fig. S2). Between July 1, 2015 and December 31, 2015, GRYD IR was notified of 129 events over this region, only an 18% increase in notifications that occurred against a backdrop of increasing violent crime across the region. The 77th 3 GRYD Zone is the only possible exception. This GRYD Zone received few notifications prior to its formal addition. In all other locations, GRYD IR crimes were recorded regardless of whether there was a formal GRYD Zone in place or not.

We therefore treat the South Los Angeles GRYD Zones as a single continuous study region for 2014 and 2015. This region is well bounded, but still expansive enough to understand the spatial dynamics of gang retaliatory violence. There is no official tally of the number of gangs present in the area. The GRYD ETO database notes 54 unique gang names in association with the suspects and victims of gang crimes in the area. This estimate is likely a lower bounds as many events lack information about suspect and victim gang affiliation. Taken at face value, the area hosts 0.6 gangs per km^2. During this period, there were 12,905 violent crimes reported to the LAPD strictly inside the South Los Angeles GRYD Zones. Of these, 3,054 were flagged as gang related. GRYD IR was notified about 809 of these violent crimes, with 666 of the notifications

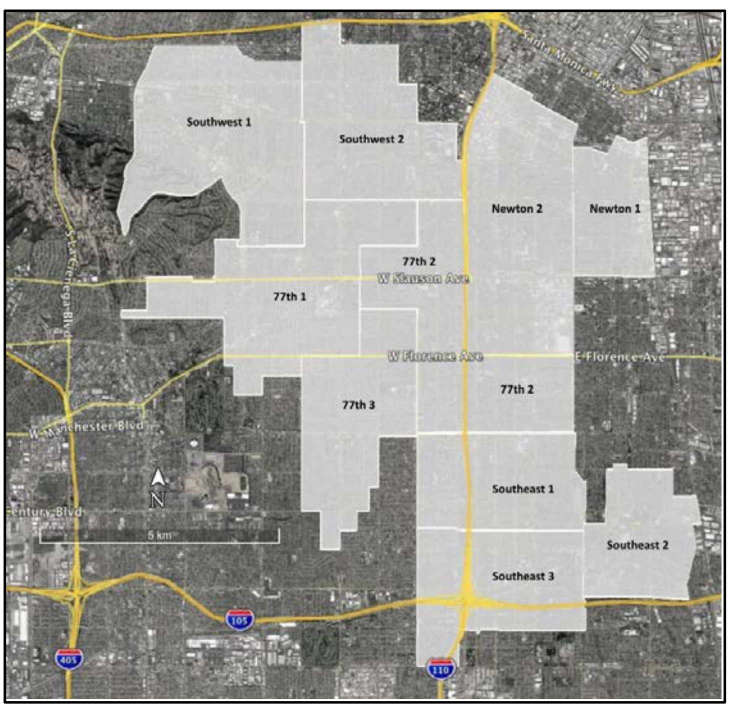


Fig. S1. Google Earth map of South Los Angeles showing the ten GRYD IR Zones in operation as of mid-2015.

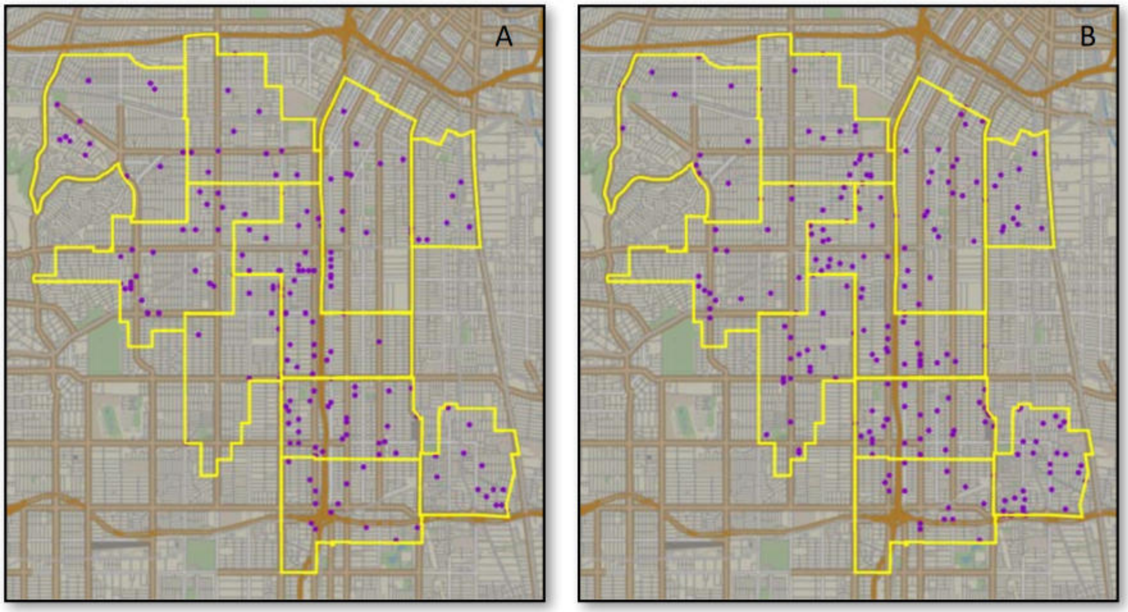


Fig. S2. Locations of GRYD IR Violent Crimes (purple) in a comparable six-month period before (A) and after (B) the July 2015 expansion of GRYD Zones in South Los Angeles. The time periods cover July-December 2014 (A) and July-December 2015 (B).

Runs tests and K-S test. GRYD IR was informed of only a fraction of the total number of gang crimes occurring gin the South Los Angeles area. We hypothesized that the notifications were random and independent, establishing test conditions approximating a randomized controlled trial. We performed one-sample runs tests to evaluate whether notifications received by GRYD IR occur at random. Two-sample K-S tests were run to test whether GRYD IR and LAPD events have the same spatial-temporal distribution.

Consider a gang crime reported to the LAPD. With probability ρ, GRYD IR is notified of the event. We use LAPD + GRYD IR as a short-hand to identify these events. With probability $1 - \rho$ the crime remains known only to the LAPD. These events are noted as LAPD-only. Over a large sequence of reported gang crimes, we expect a fraction of crimes proportional to ρ to fall into condition LAPD + GRYD IR and a fraction proportional $1 - \rho$ to fall into condition LAPD-only. However, we also expect the assignment of any one event to be random and independent of other assignments. To test this hypothesis, reported gang crimes were arranged in ascending order of the date and time received and computed a runs test. A run is defined as a sequence of notifications of the same condition, uninterrupted by notifications of the other condition. For example, the sequence of coin tosses HHTHHHTTTT consists of four runs {HH}{T}{HHH}{TTTT}. Let R be the observed number of runs in a sample of size N. The null hypothesis is that the observed number of runs R is the product random and mutually independent notifications. Note that ρ need not equal $1 - \rho$ as would be the case in fair coin toss. All that is required is that GRYD IR notifications occur at probability ρ independently.

Results for different subsets of the observed data fail to reject the null hypothesis (Table S2). For gang aggravated assaults and homicides combined, the number of observed runs ($R = 852$) is not statistically different from the expected number of runs under random assignment ($E[R] = 840.31$) ($Z = 0.609$, $p = 0.54$). For gang aggravated assaults alone, the observed number of runs ($R = 672$) is not statistically different from the expected number under random assignment ($E[R]$

= 683.99) (Z = 0.73, p = 0.47). Similarly, for gang homicides the observed number of runs (R = 65) is not statistically different from the expected number under random assignment (E[R] = 65.58)(Z = 0.125, p = 0.90). While the percent of crimes that lead to GRYD IR notifications differs by crime type, these notifications are random on a per event basis. Therefore, we conclude that the assignment of events to different conditions approximates and randomized experimental protocol.

Table S2. Results of one sample runs tests for three different categories of gang crimes in South Los Angeles.

	Gang Aggravated Assault + Homicide	Gang Aggravated Assault	Gang Homicide
LAPD + GRYD IR (n1)	622	470	152
LAPD-only (n2)	1290	1249	41
Total (N)	1912	1719	193
Percent LAPD + GRYD IR	32.5%	27.3%	78.7%
Runs	852	672	65
Expected Mean	840.31	683.99	65.58
Expected SD	19.18	16.47	4.62
Z	0.609	0.73	0.125
two-tailed p-value	0.54	0.47	0.90

GRYD IR Notification and Field Deployment. Gang crimes are typically reported to the LAPD first by members of the public. The GRYD IR then receives notification, but only for a fraction of the reported crimes. CIWs are informed following notification of the GRYD Program Office. We include all events where GRYD IR is notified, rather than restricting analysis to events where there is some record of field activity by CIWs. This analytical choice is made due to uncertainty surrounding the recording of field activities of CIWs. Skogan (*9: 5-3*) noted for Chicago's CeaseFire program that it was difficult to assess how frequently and effectively Chicago's violence interrupters mediated conflicts. Record keeping was a cumbersome task and not natural for violence interrupters engaging problems on the street. Moreover, there was considerable worry that formal documentation could incriminate people and therefore was often not collected in the first place. Here, as with the Chicago case, it was difficult to directly evaluate dosage with confidence.

The resulting hypotheses based on notification of GRYD IR are conservative, however. On the one hand, we suppose that CIWs can only have a direct impact on retaliation if they know about a potential triggering event. Thus, notification of GRYD IR is a logical precursor for treatment effects. If notification did not lead to interventions in the field by CIWs then we would expect there to be no difference between the control and treatment conditions. Specifically, even though events were labeled as LAPD + GRYD IR, the absence of field intervention would ensure that such events were no different than LAPD-only one. On the other, if CIWs were able to source information about events on their own and self-deploy, operating outside of the normative channel of communication, then we would expect this contamination to also bias the results towards finding no difference between treatment and control. In other words, some fraction of events labeled as LAPD-only would include the effects of GRYD IR without formal recognition. Thus, LAPD-only interventions would appear similar to LAPD + GRYD because of this hidden activity. Because we are able to document a significant difference between LAPD-only and LAPD + GRYD IR intervention effects (main text Fig. 2C) such confounds are unlikely.

Methods

Multivariate self-exciting point process models. Self-exciting point process models provide a useful mathematical structure for considering the dynamics of retaliatory crime (*18, 30*). The problem we analyze requires that we treat gang retaliation in both space and time as well as events exposed to different experimental conditions. We use a multivariate spatial-temporal self-exciting point process model to capture these conditions:

$$\lambda_u(t,x,y) = \mu_u(x,y) + \sum_{t_i<t} K_{u_i u}\, g(x-x_i, y-y_i, t-t_i).$$

The equation describes the instantaneous rate λ_u at which crimes assigned to condition u. In our specific case, crimes are assigned to two different conditions, those known only to the LAPD and those known to both the LAPD and GRYD IR. Notice that this is a spatial-temporal model. The random background rate at which gang crimes occurs may vary from place to place $\mu(x, y)$, but not in time. The self-exciting triggering function is also spatially dependent with the magnitude of excitation dependent not on how long ago a prior crime occurred ($t - t_i$), but also how nearby in space ($x - x_i$, $y - y_i$). Gang crimes are more likely to trigger a retaliation soon after the intial event and nearby in space. We use particular parametric equations for μ and g:

$$\mu_u(x,y) = \sum_{i=1}^{N} \frac{\beta_{u_i u}}{2\pi\eta^2 T} \times \exp\left(\frac{-\left((x-x_i)^2 + (y-y_i)^2\right)}{2\eta^2} \right), \tag{S1}$$

$$g(x,y,t) = \omega \exp(-\omega t) \times \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2} \right). \tag{S2}$$

The equation for μ treats the stationary background rate of crimes of type u as a sum of Gaussian kernels. The triggering kernel g treats the self-exciting effect of one event as decaying exponentially in time and Gaussian in space. Here β is a weights matrix for the degree to which events assigned to condition u_i contribute the background rate for events assigned to condition u. The parameter T is the total time period represented by the sample of gang crimes, which in this case is 2014-2015. The parameters η and σ represent the spatial scale of influence for background events and retaliatory events, respectively (see *31*). We take $\eta = \sigma$, which simplifies parameter estimation.

Our quantities of interest are estimates of four key relationships encoded in matrix $K_{u_i u}$: (1) k_{11} or the average number of LAPD + GRYD IR gang crimes triggered by a single LAPD + GRYD

IR gang crime; (2) k_{01} the average number of LAPD + GRYD IR gang crime triggered by a single LAPD-only gang crime; (3) k_{10} or the average number of LAPD-only gang crimes triggered by a single LAPD + GRYD IR gang crime; and (4) k_{00} or the average number of LAPD + GRYD IR gang crime triggered by a single LAPD-only gang.

The procedure we use is a type of Maximum Likelihood Estimation (MLE) known as expectation maximization (EM) (*20, 32*). The expectation step of the EM algorithm is used to compute initial probabilities p^b_{ij} and p_{ij} that an event *i* causes event *j* via either the background rate μ or the self-exciting kernel g, respectively. These expectations are then fed to the maximization step where a new set of parameter values (for iteration *k* + 1) are determined by maximizing the expected probability with respect to the observed data. This maximization is done for all parameters taking into consideration whether gang crimes are known only to the LAPD or to both the LAPD and GRYD IR. The algorithm alternates between expectation and maximization until there is no further change in the parameter values.

For completeness, the EM algorithm is structured as below. Note that $n_{\hat{u}}$ is the number of events that belongs to experimental condition type $\hat{u}$.

Complete Data Likelihood Function:

$$Q(\Omega)=\sum_{i=1}^{N}\sum_{j=1}^{N}p_{ij}^{b}log\left(\frac{\beta_{u_iu_j}}{2\pi\eta^2T}e^{\left(-\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2\eta^2}\right)}\right)-\sum_{u=1}^{U}\sum_{i=1}^{N}\beta_{u_iu}$$

$$+\sum_{i<j}p_{ij}log\left(\omega K_{u_iu_j}e^{-\omega(t_j-t_i)}\frac{1}{2\pi\sigma^2}e^{-\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2\sigma^2}}\right)-\sum_{u=1}^{U}\sum_{i=1}^{N}K_{u_iu}(1-e^{-w(T-t_i)})$$

Expectation Step:

$$p_{ij}=K_{u_iu_j}\omega\frac{\exp\left(-\omega(t_j-t_i)\right)}{2\pi\sigma^2}\times\frac{\exp\left(-\frac{(x_j-x_i)^2+(y_j-y_i)^2}{2\sigma^2}\right)}{\lambda_{u_j}(x_j,y_j,t_j)}$$

$$p_{ij}^{b}=\frac{\beta_{u_iu_j}}{2\pi\eta^2T}\frac{\exp\left(-\frac{(x_j-x_i)^2+(y_j-y_i)^2}{2\eta^2}\right)}{\lambda_{u_j}(x_j,y_j,t_j)}$$

Maximization Step:

$$\omega^{(k+1)}=\frac{\sum_{j<i}p_{ij}^{(k)}}{\sum_{i<j}p_{ij}^{(k)}(t_j-t_i)+\sum_{u=1}^{U}\sum_{i=1}^{N}K_{u_iu}(T-t_i)e^{-\omega(T-t_i)}}$$

$$K_{\hat{u}u}^{(k+1)}=\frac{\sum_{i,j\in A_{\hat{u}u}}p_{ij}}{\sum_{l=1}^{n_{\hat{u}}}(1-e^{-w(T-t_{i_l})})}$$

$$\sigma^{2(k+1)}=\frac{\sum_{i<j}p_{ij}^{(k)}\left((x_i-x_j)^2+(y_i-y_j)\right)^2}{2\sum_{i<j}p_{ij}^{(k)}}$$

$$\beta_{\hat{u}u}^{(k+1)}=\frac{\sum_{i,j\in B_{\hat{u}u}}p_{ij}^{b(k)}}{n_{\hat{u}}}$$

$\eta^{2(k+1)} = \frac{\sum_{i,j=1}^{N} p_{ij}^{b(k+1)}\left((x_i - x_j)^2 + (y_i - y_j)^2\right)}{2\sum_{i,j=1}^{N} p_{ij}^{b(k+1)}}$	$A_{\hat{u}u} = \left\{i, j \text{ index of events} \mid t_i < t_j, u_i = \hat{u}, u_j = u\right\}$ $B_{\hat{u}u} = \left\{i, j \text{ index of events} \mid u_i = \hat{u}, u_j = u\right\}$

Non-parametric Model Validation and Voronoi Residuals Analysis

Our choice of an exponential kernel for *g* assumes a certain mechanistic form for the dynamics of gang retaliation. To verify the validity of this choice we use non-parametric methods (*15*). The method makes no assumptions about the shape of the triggering function, and instead provides a data-driven estimate which can be used to help identify possible parameterizations. Below the spatial and temporal response are plotted on log-log scale and appear exponential. In addition, we check the goodness fit for the exponential fit. For the spatial-temporal kernel *g*, the R-square and adjusted R-square are both 1.

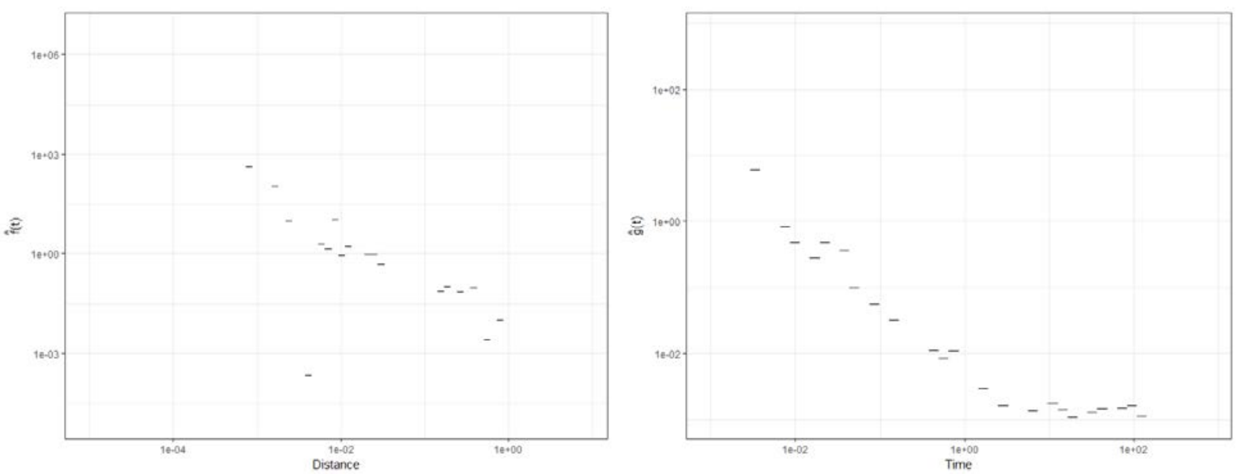


Fig. S3. Spatial and temporal triggering kernels estimated non-parametrically from the data.

Once a model is estimated, a powerful technique for evaluating model performance is Voronoi residuals (*22*). Voronoi residuals allow the examination of differences between the modeled conditional intensity and the observed number of points within spatially adaptive Voronoi cells. Using color scaling allows us to see the spatial locations where a model is over or under estimated. Cells with blue shades indicate an overestimation of the intensity while cells with red indicate an underestimation. In addition, comparison to a color scale defined by a null Poisson model further helps interpret performance (*23*). In Fig. S4, the fitted intensity for the proposed Hawkes model has muted colors as compared to the null model, indicating improved performance. Residuals indicate that the fitted model performed well throughout the spatial window.

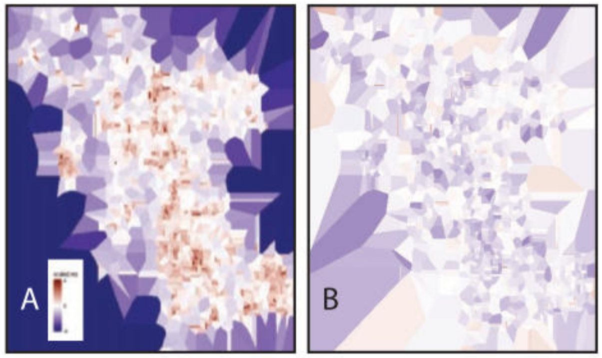


Fig. S4. Voronoi residuals for the Poisson model (A) and a Hawkes model (B).

Stochastic declustering. Gang crimes occurring in a given area represent a mixture of those that are that background events and those that are retaliatory in response to other crimes. We wish to sort events into these two groups to understand how important background and retaliatory processes are for gang violence overall.

Stochastic declustering is a suite of methods developed in the study of earthquake catalogs where the goal is to distinguish between background seismicity and aftershocks (*24*). The same methods can be applied to the study of crime (*16*).

Starting with a self-exciting point process model like the one developed here, stochastic declustering proceeds through a thinning procedure that removes events probabilistically classified as retaliations. The events remaining after thinning represent the background events generated by a spatially non-homogeneous Poisson process $\lambda(t, x, y) = \mu(x, y)$. Specifically, in the univariate case, the probability that an event *j* is a retaliation is given by

$$\rho_j = \frac{\sum_{t_i < t_j} g(t_j - t_i, x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j)}.$$

The probability that an event j is a background event is therefore

$$1 - \rho_j = \frac{\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j)}.$$

For a catalog of *N* total crimes and a point process model fit to those events, the simplest procedure is to generate *N* uniform random variables $U_1, U_2, \ldots, U_N$ in the range [0,1]. An event is classified as a background crime when $U_j < 1 - \rho_j$, otherwise it is removed and classified as a retaliation (*24*).

Note that the assignment of an event to being background or retaliation is a probabilistic classification. On average the relative mixture of background and retaliation events is correct for a given time window and spatial region, but we cannot say with absolute certainty whether any specific event is or is not a retaliation.

Estimating Statistical Significance. Our null hypothesis is that the GRYD IR interventions have no impact. That is GRYD IR does not reduce gang violence. If this null hypothesis is true then the ground truth values of the matrix K should be $k_{11} = k_{01}$ and $k_{10} = k_{00}$. Given estimates k_{11} and k_{01} and standard errors s_{11} and s_{01}, for example, the quantity

$$\frac{k_{11} - k_{01}}{\sqrt{s_{11}^2 + s_{01}^2}}$$

should be t-distributed and for large samples approximately standard normal. The p-value associated with the magnitude of the observed difference between estimates of k_{11} and k_{01} can be computed directly against the cumulative distribution for the standard normal. We reject the null hypothesis based on standard probability criteria.

Results

Gang vs. non-gang crime in South Los Angeles. We adapt the multivariate self-exciting point process model to consider the retaliatory dynamics for gang aggravated assaults and homicides interacting with non-gang aggravated assaults and homicides (Fig. S5). Here we allow for the possibility that retaliation may occur within officially recognized gang crimes as well as between gang crimes and non-gang crimes. We expect gang aggravated assaults and homicides to trigger significantly more retaliations than non-gang crimes of the same type.

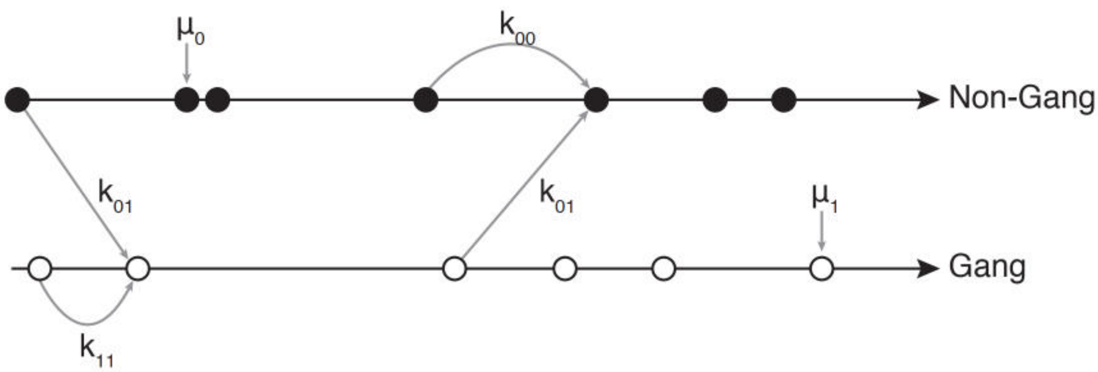


Fig. S5. The conceptual model for the multivariate self-exciting point process applied to non-gang and gang crime as the two test conditions.

Model parameters corresponding to the gang and non-gang crime conditions are as follows:

$$K_{u,u} = \begin{pmatrix} \underset{(0.0271)}{0.2322} & \underset{(0.0034)}{0.0163} \\ \underset{(0.0015)}{0.0090} & \underset{(0.0070)}{0.1014} \end{pmatrix} \qquad \beta_{u,u} = \begin{pmatrix} \underset{(0.0149)}{0.4286} & \underset{(0.0152)}{0.3856} \\ \underset{(0.0135)}{0.1504} & \underset{(0.0044)}{0.7098} \end{pmatrix}$$

where the matrix K gives productivity for the four potential triggering relationships shown in Fig. S5 and matrix β gives the mixture weights for the spatially inhomogeneous background rate. Parameter estimates are shown along with their standard errors in parentheses. Gang crimes trigger significantly more gang retaliations than do non-gang crimes ($k_{11} > k_{01}$ or $0.2322 > 0.0090$, $p < 10^{-6}$). Gang crimes trigger significantly fewer non-gang retaliations than do other non-gang crimes ($k_{10} < k_{00}$ or $0.0163 < 0.1014$ $p < 10^{-6}$). It is notable that the magnitude of within-condition triggering is quite similar ($k_{11} \approx k_{00}$). To the extent that retaliation is a defining feature of gang violence, this may suggest that some retaliatory gang crimes are mistakenly labeled as non-gang crimes.

LAPD-only vs. LAPD + GRYD combined gang and non-gang crime. The apply the multivariate model to the case of aggravated assaults and homicides for non-gang and gang crimes combined where LAPD was notified and where LAPD and GRYD IR were both notified (Fig. S6). This is the first set of test conditions where a treatment effect from GRYD IR notification to may appear.

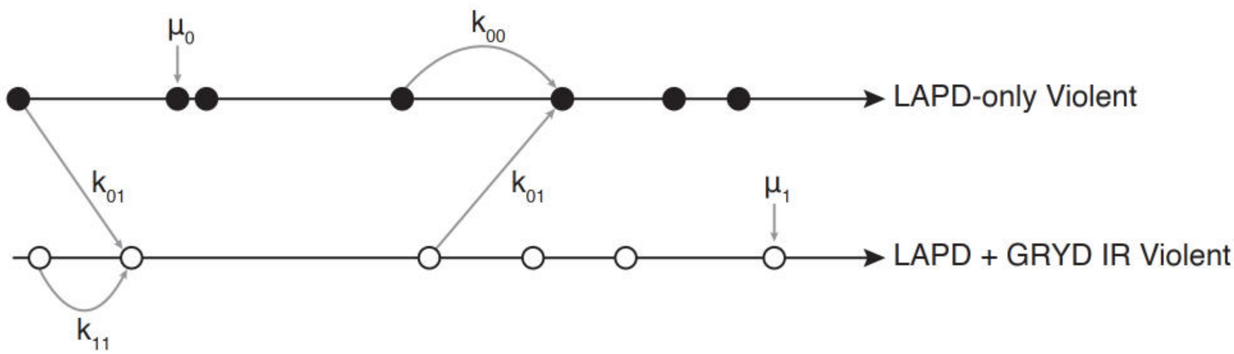


Fig. S6. The conceptual model for the multivariate self-exciting point process applied to gang and non-gang aggravated assaults and homicides known only to the LAPD and to the LAPD plus GRYD IR.

Model parameters corresponding to the gang and non-gang crime conditions are as follows:

$$K_{u_i u} = \begin{pmatrix} \underset{(0.0362)}{0.1401} & \underset{(0.0367)}{0.2841} \\ \underset{(0.0083)}{0.0526} & \underset{(0.0346)}{0.2824} \end{pmatrix} \qquad \beta_{u_i u} = \begin{pmatrix} \underset{(0.0072)}{0.0425} & \underset{(0.0484)}{0.5898} \\ \underset{(0.0026)}{0.0831} & \underset{(0.0197)}{0.6794} \end{pmatrix}$$

The matrix K is shown as Fig. 2B in the main text. Crimes where GRYD IR is notified trigger significantly more retaliations than when GRYD IR is not notified ($k_{11} > k_{01}$ or $0.1401 > 0.0526$, $p < 0.0093$). GRYD IR crimes also trigger more LAPD-only retaliations than other LAPD-only crimes ($k_{10} > k_{00}$ or $0.02841 > 0.2824$, $p = 0.49$), though the difference is not significant. GRYD IR does not reduce retaliation relative to LAPD-only interventions when considering the entire pool of non-gang and gang aggravated assaults and homicides. This result stems from the fact that GRYD IR is notified primarily for gang-related crimes and therefore GRYD IR must confront crimes where the risk of retaliation is much greater from the outset. In other words, GRYD IR may be able to reduce retaliation among those crimes it confronts, but not to levels that characterize the broader pool of crimes that includes many more that are not gang related.

LAPD-only gang crime vs. LAPD + GRYD IR gang crime. We restrict analyses to those aggravated assaults and homicides that are flagged as gang related (Fig. S7). The comparison of the control and treatment condition are on more equal footing for this set of crimes.

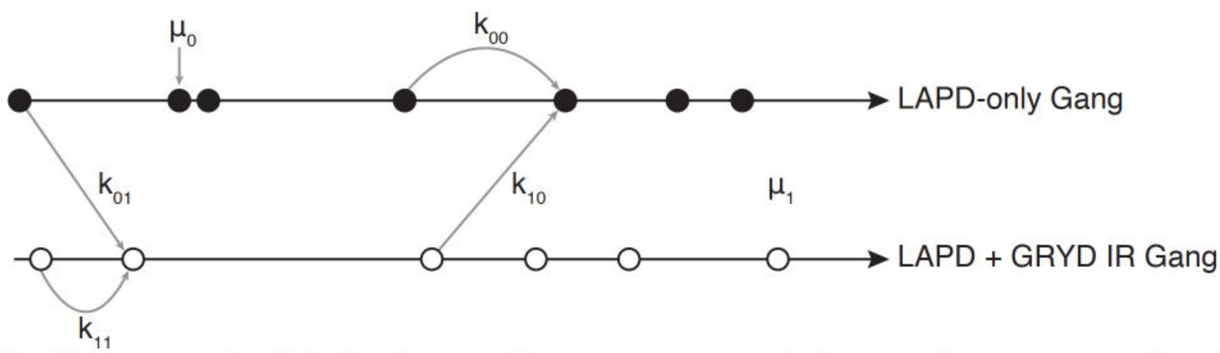


Fig. S7. The conceptual model for the multivariate self-exciting point process applied to gang and non-gang aggravated assaults and homicides known only to the LAPD and to the LAPD plus GRYD IR.

Model parameters corresponding to the gang and non-gang crime conditions are as follows:

$$K_{u_i u} = \begin{pmatrix} \underset{(0.0021)}{0.0015} & \underset{(0.0267)}{0.1483} \\ \underset{(0.0066)}{0.0621} & \underset{(0.0128)}{0.2116} \end{pmatrix} \quad \beta_{u_i u} = \begin{pmatrix} \underset{(0.0282)}{0.3360} & \underset{(0.0168)}{0.4732} \\ \underset{(0.0181)}{0.2574} & \underset{(0.0096)}{0.4889} \end{pmatrix}$$

The matrix K is shown as Fig. 2C in the main text. GRYD IR gang crimes trigger significantly fewer GRYD IR gang retaliations than LAPD-only gang crimes ($k_{11} < k_{01}$ or $0.0015 < 0.0621$, $p < 10^{-6}$). GRYD IR gang crimes also trigger fewer LAPD-only retaliations than other LAPD-only crimes ($k_{10} < k_{00}$ or $0.1483 < 0.0164$, $p = 0.49$).

Estimating prevented crimes.

The parameter values from the spatio-temporal model can be used to estimate the number of crimes prevented by GRYD IR notifications. The sum ($k_{11} + k_{01}$) is the average number of retaliations known to both the LAPD and GRYD IR produced by the two types of triggers, LAPD + GRYD IR and LAPD-only. Similarly, ($k_{10} + k_{00}$) is the average number of retaliations known only to the LAPD produced by the two types of triggers. Note that ($k_{11} + k_{01}$) and ($k_{10} + k_{00}$) are actually measured directly from data and therefore are the observed outcome. We can also define two counterfactual situations. Let ($k_{01} + k_{01}$) be the average number of retaliations that *would have occurred* in the absence of GRYD IR notification for those retaliations known to LAPD + GRYD IR. Here we simply replace k_{11} with a second instance of k_{01}. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. Similarly, let ($k_{00} + k_{00}$) be the average number of retaliations that *would have occurred* in the absence of GRYD IR notification for those retaliations known only to the LAPD. Again we suppose that the LAPD + GRYD IR effect is replaced by the LAPD-only effect in the absence of GRYD IR notification. We then compute the relative effect of GRYD IR notification on the average number of retaliations as:

Relative effect of GRYD IR notification on LAPD + GRYD IR retaliations	Relative effect of GRYD IR notification on LAPD-only retaliations
$\dfrac{(k_{11}+k_{01})-(k_{01}+k_{01})}{(k_{01}+k_{01})}$	$\dfrac{(k_{10}+k_{00})-(k_{00}+k_{00})}{(k_{00}+k_{00})}$

For the South Los Angeles GRYD Zones, GRYD IR notification reduces retaliation among LAPD + GRYD IR events by -48.8% relative to the counterfactual. GRYD IR notification reduces retaliation among LAPD-only events by -15.0% relative to the counterfactual. See Fig. 2C in the main text for corresponding values of *k*.

We use these estimated effects along with the results of stochastic declustering in South Los Angeles to compute numbers of prevented crimes. Stochastic declustering identified a total of 45 LAPD + GRYD IR gang aggravated assaults and homicides in 2014-2015 as retaliatory (see Table 2 in the main text). The remaining 577 gang aggravated assaults and homicides were statistically defined as background events. Similarly, declustering identified a total of 403 LAPD-only gang aggravated assaults and homicides in 2014-2015 as retaliatory (see Table 2 in the main text). The remaining 877 gang aggravated assaults and homicides were statistically identified as background events. The counterfactual conditions suggest that retaliatory gang aggravated assaults and homicides would have been 48.8% and 15.0% higher in the absence of GRYD IR for events recorded, respectively, as LAPD + GRYD IR and LAPD-only. Thus GRYD IR prevented an estimated total 82.2 retaliatory gang aggravated assaults and homicides. An estimate based on city-wide data, suggests that homicides make up on average 5.4% of all gang retaliations. Thus, in South Los Angeles, the prevented crimes are expected to include 4.4 retaliatory homicides and 77.8 retaliatory aggravated assaults.

We refer to McCollister et al. (*28*) for estimates of the costs of crime. In their work the total cost of a single homicide to government, victims and suspects is approximately $8.98 million. The cost of a single aggravated assault is $240,000. We simply multiply these figures by the estimated number of prevent aggravated assaults and homicides, respectively. The estimated number of homicides prevented by GRYD IR may add up to savings between $39.4 million over two years. The savings from prevented gang aggravated assaults in South Los Angeles may amount to an additional $9.5 million over two years. The combined savings per year in South Los Angeles alone may amount to $49.0 million.

Rapid community intervention in response to gang violent crime directly reduces gang retaliation

Authors: P. Jeffrey Brantingham[1]*, Baichuan Yuan[2], Nick Sundback[4], Frederic Paik Schoenberg[3], Andrea Bertozzi[2], Joshua Gordon[3], Jorja Leap[4], Kristine Chan[5], Molly Kraus[5], Sean Malinowski[6], Denise Herz[5].

Affiliations:
[1]Department of Anthropology, University of California Los Angeles, Los Angeles, CA 90095, USA.
[2]Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA.
[3]Department of Statistics, University of California Los Angeles, Los Angeles, CA 90095, USA.
[4]Luskin School of Public Affairs, University of California Los Angeles, Los Angeles, CA 90095, USA.
[5]School of Criminal Justice and Criminalistics, California State University Los Angeles, Los Angeles, CA 90032, USA.
[6]Los Angeles Police Department, Los Angeles, CA 90012, USA.

*Correspondence to: branting@ucla.edu

Abstract:

Retaliation propels gang violence. Spontaneous attacks resulting from chance encounters between rivals, or situational interactions that challenge gang territory or reputation can trigger cycles of tit-for-tat reprisals. Yet it has been difficult to determine if interventions that seek to reduce the likelihood of retaliation translate into lower rates of gang crime. Here we use multivariate spatial-temporal point process models to quantify the magnitude of retaliation arising from gang crimes exposed to different types of interventions. The methods are well-suited to analysis of real-world interventions where there is imperfect separation between experimental conditions. Our analyses of quasi-experimental interventions in Los Angeles indicates that Community Intervention Workers tasked by the Gang Reduction Youth Development program cut gang retaliations by 45.3%, independently of the effects of policing. Efforts to engage impacted families and control rumors reduce the contagious spread of violence if undertaken in the immediate aftermath of gang violent crimes. These findings have important implication for the design, implementation and evaluation of gang violence prevention programs.

One Sentence Summary:

Multivariate spatial-temporal point process models fitted to gang-violent crimes in Los Angeles, CA, show that interventions to reduce the likelihood of retaliation cut the incidence of violent gang crimes nearly in half.

Main Text:

Gang violence is distinctive for its ability to trigger clusters of retaliatory crimes (*1, 2*). Challenges between gangs that threaten geographic territory or gang reputation can easily escalate to a shooting, while a shooting or homicide often demands retribution (*3, 4*) ultimately driving a sequence of tit-for-tat reciprocal attacks (*5-7*). Retaliation adds to the cumulative volume of violent crime and stronger retaliatory forces add more total crime (*8, 9*). It is important therefore to evaluate whether targeted interventions intended to reduce gang violence can do so by disrupting the process of retaliation. Although several studies have examined the aggregate impacts of anti-violence programs, including efforts to interrupt street violence (*10-12*), they produced mixed results (*13-15*). Previous research has not quantified direct impacts of interventions on gang retaliation. Here we show that such impacts can be estimated directly from crime event data using a unique multivariate statistical model.

Recent advances in statistical modeling of point processes reveal the dynamics of randomly occurring events characterized by self-excitation or contagion (*16, 17*){REF}. Such models have conditional intensity (*18*):

$$\lambda(t, x, y) = \mu(x, y) + \sum_{t_i < t} Kg(t - t_i, x - x_i, y - y_i), \tag{1}$$

where λ is the infinitesimal rate at which events accumulate at any point in space-time, given the entire history of the process. The model provides an intuitive characterization of gang violence (*19*). It partitions the cause of crime into background processes $\mu(x, y)$, such as simmering gang rivalries, that generate crimes randomly at a constant, but spatially variable rate, and contagious processes $Kg(t - t_i, x - x_i, y - y_i)$ that locally and briefly amplify the rate at which crime occurs (Fig. 1A). Retaliation is a dominant contagion process (*8, 20*).

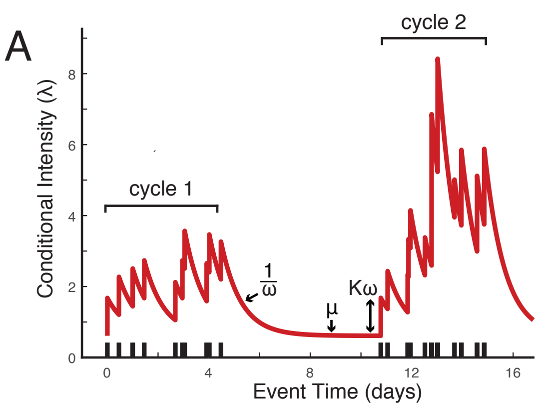


Fig. 1. Self-exciting point process models capture the dynamics of gang violence. (A). A temporal self-exciting point process model $\lambda(t) = \mu + \sum_{t_i<t} Kg(t - t_i)$ with exponential kernel $g(t) = \omega e^{-\omega(t-t_i)}$ fit to a sample of gang aggravated assaults and homicides in South Los Angeles from 2014-2015. Two cycles of gang violence occur within a period of eighteen days. The conditional intensity λ reflects the instantaneous rate of gang crime. The background rate μ is the expected rate of gang crime in the absence of retaliation. A crime causes λ to jump by an amount $K\omega$, increasing the risk of retaliation. When elevated above μ, the risk of retaliation decays at a constant rate ω, with a mean lifetime of $1/\omega$. (B). Gang crimes assigned to two different experimental conditions are modeled as two interacting point processes. Non-retaliatory gang crimes assigned to each condition arise spontaneously at rate μ_j. Retaliations assigned to each experimental condition may be triggered through separate pathways. Pathways k_{11} and k_{10} link previous treatment crimes to treatment and control retaliations, respectively. Pathways k_{00} and k_{01} link previous control crimes to control and treatment retaliations, respectively. If treatment interventions (red events) reduce the risk of gang retaliation, then we expect $k_{11} < k_{01}$ and $k_{10} < k_{00}$.

We extend model (1) to a multivariate framework (*21*) useful for describing real-world interventions. In such quasi-experimental field settings there is often imperfect separation between experimental conditions. To account for interactions we propose the conditional intensity:

$$\lambda_u(t, x, y) = \mu_u(x, y) + \sum_{t_i < t} K_{u_i u}\, g(t - t_i, x - x_i, y - y_i). \quad (2)$$

Here u_i is the type of event i where $u = 0$ represent an event assigned to a non-intervention control condition and $u = 1$ an event assigned to an intervention treatment condition. The model is easily modified to accommodate more than two interacting experimental conditions. The spatially inhomogeneous background rate of events is now partitioned according to condition u (supplemental text). The parameter $K_{u_i u}$ is the expected number of retaliations of type u triggered by an event of type u_i. Thus we have four productivity parameters to estimate $K_{u_i u} = k_{11}$, k_{01}, k_{10}, and k_{00}, representing the four possible interactions between treatment and control conditions (Fig. 1B). If treatment interventions are effective, then estimated parameters should satisfy $k_{11} < k_{01}$ and $k_{10} < k_{00}$. We estimate the model parametrically using an expectation-maximization (EM) algorithm (*22*). A non-parametric model yields substantially similar results. The parametric model fit is evaluated using Voronoi residuals (*23, 24*) (supplemental text).

We analyzed gang and non-gang violence in a unique quasi-experimental setting in Los Angeles where there is approximately random assignment of crimes between two different, but interacting intervention conditions. The random assignment between the two conditions arises naturally out of the crime reporting system (see below). The control condition consists of violent crimes responded to by the Los Angeles Police Department (LAPD-only). The treatment condition consists of crimes responded to by the LAPD, but with additional notification of the City Los Angeles Mayor's Office of Gang Reduction and Youth Development (GRYD) Intervention Incident Response program (LAPD + GRYD IR). Upon receiving notification, GRYD IR tasks Community Intervention Workers with disrupting retaliation through crisis response and rumor control (supplemental text). We focus on crimes occurring in an 87.2 km^2 (33.7 sq miles) area of South Los Angeles during 2014-2015 (Fig. 2D-I, Fig. S1). The ten GRYD IR Zones in South Los Angeles represent only 6.7% of the total land area of Los Angeles (~1,302 km^2) and about 15.5% of the total population (~3.9 million), but accounted for 45.3% of serious gang crimes city-wide in 2014-2015. We limit our consideration to aggravated assaults and criminal homicides, crimes which entail a greater risk of retaliation compared to other crime types (supplementary text).

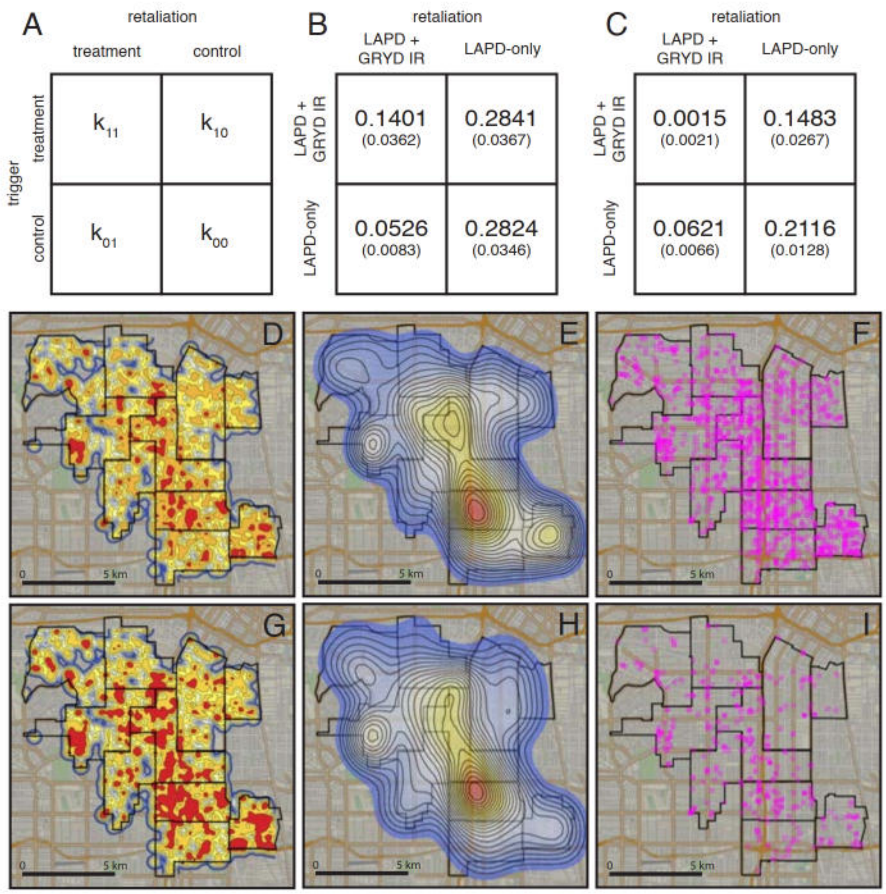


Fig. 2. (A) Matrix representation of productivity K_{ij} with the corresponding triggering pathways noted. Matrix entries are the average number of retaliations assigned to an experimental condition i triggered by an event assigned to condition j. (B). The productivity for combined gang and non-gang aggravated assaults and homicides in South Los Angeles for 2014-15. (C). The productivity for gang aggravated assaults and homicides in South Los Angeles for 2014-15. The control condition includes violent crimes known to the LAPD (LAPD-only). The treatment condition includes crimes known to the LAPD that were also reported to GRYD IR (LAPD + GRYD IR). Standard errors of parameter estimates are shown in parentheses. (D). The log of spatial-temporal background intensity function μ for gang violent crimes mapped over space. (E) Contour plot of the density of background gang aggravated assaults and homicides determined by declustering. (F) Point locations of background gang aggravated assaults and homicides determined by declustering. (G) The log of spatial-temporal triggering kernel g mapped over space. (H) Contour plot of the density of retaliatory gang aggravated assaults and homicides determined by declustering. (I) Point locations of retaliatory gang aggravated assaults and homicides determined by declustering. Boundaries for the ten GRYD IR Zones in South Los Angeles are outlined in black.

In 2014-15, a total of 5,982 aggravated assaults and homicides were reported to the LAPD in the South Los Angeles GRYD Zones, including gang (32.0%) and non-gang crimes (68.0%) (Table 1). GRYD IR was notified in 9.9% of all aggravated assaults, but 71.5% of all homicides. GRYD IR was notified more frequently when the crime was identified as gang related, including in 27.3% of gang aggravated assaults and 78.8% of gang homicides. A key feature facilitating our analysis is that crimes assigned to each of the two test conditions approximated a randomized experimental protocol. To verify we performed runs tests ($Z = 0.609$, $p=0.54$) and two-sample Kolmogorov–Smirnov (KS) tests (K-S=0.069, $p=0.20$ in time and K-S=0.054, $p=0.17$ in space), which found no discernable departures from the null hypothesis that the GRYD IR notifications were random, independent draws from the same population of crimes (supplementary text). Upon LAPD receiving a report of a gang crime, notification of GRYD IR proceeded as if a biased coin was flipped. If the crime was a gang aggravated assault, the coin was biased towards not notifying GRYD IR. If it was a gang homicide, it was biased towards notifying GRYD IR. Therefore, potential treatment effects are not confounded with the process of GRYD IR notification.

Table 1. Gang and non-gang crimes reported only to the LAPD and to the LAPD and GRYD IR in 2014-2015.

	LAPD-only				**LAPD + GRYD IR**				**GRAND TOTAL**	
	Gang N	**Non-Gang N**	**TOTAL N**	**% Gang**	**Gang N**	**Non-Gang N**	**TOTAL N**	**% Gang**	**TOTAL N**	**% Gang**
Aggravated Assault	1,249	3,918	5,167	24.2%	470	99	569	82.6%	5,736	30.0%
Homicide	41	29	70	58.6%	152	24	176	86.4%	246	78.5%
TOTAL	1,290	3,947	5,237	24.6%	622	123	745	83.5%	5,982	32.0%

GRYD IR by design is deployed disproportionately in response to gang crimes, representing 83.5% of aggravated assaults and homicides for which it received notification (Table 1) (supplemental text). The 16.5% of non-gang crimes reported to GRYD IR were likely considered gang-related at the point of initial notification, but were subsequently reclassified. Because GRYD IR confronts a mix of gang and non-gang crimes, we first tested whether GRYD IR has an impact against violent crimes in general. We fit the multivariate model, with an exponential kernel for g, to the full complement of 5,982 crimes (Fig. 2B). The estimate of parameter k_{11} indicates each aggravated assault or homicide exposed to the treatment triggered on average 0.1401 retaliations subsequently known to both LAPD and GRYD IR, whereas k_{01} indicates each control assault or homicide triggered on average 0.0526 retaliatory crimes. The 62.5% higher rate of retaliation for crimes reported to GRYD IR is statistically significant ($p = 0.0092$) (supplemental text). Pathway k_{10} shows that treatment crimes also triggered on average 0.2841 retaliations known only to the LAPD, which is of equivalent magnitude to pathway k_{00} with 0.2824 retaliations known only to the LAPD ($p = 0.486$). In practical terms, every 100 LAPD + GRYD IR (treatment) aggravated assaults and homicides triggered on average 42.4 retaliatory violent crimes ($k_{11} + k_{10}$) compared to an average 33.5 crimes triggered by control crimes ($k_{01}+ k_{00}$). GRYD IR may reduce retaliation among the broader set of gang and non-gang violent crimes, but not to levels characteristic of the LAPD control case with its more generous mix of non-gang crimes and inherently lower risk of retaliation.

We therefore restricted analyses to gang aggravated assaults and homicides to ensure that test conditions were evaluated given events with similar potential for spawning retaliation. Against this set of crimes GRYD IR had a substantial impact (Fig. 2C). Pathway k_{11} triggered an average of 0.0015 retaliations for any one treatment gang crime. By contrast, pathway k_{01} triggered 0.0621 retaliations for any one control gang crime. This represents a 97.6% reduction in retaliation associated with GRYD IR notification ($p < 10^{-6}$). Pathway k_{10} triggered an average of 0.1483 retaliations for any one treatment gang crime. Pathway k_{00} triggered 0.2116 retaliations for any one control gang crime. The 29.9% reduction in retaliations is significant ($p = 0.0163$). Every 100 LAPD + GRYD IR (treatment) gang crimes triggered an average of 15.0 retaliations

(k_{11} + k_{10}) compared to 27.4 retaliations triggered by control crimes (k_{01}+ k_{00}). Overall, the notification of GRYD IR was associated with a 45.3% decrease in retaliatory gang crimes.

To better understand the spatial dynamics of retaliatory gang violence we mapped the background intensity μ and triggering kernel g along with the distributions of background and retaliatory crimes determined via stochastic declustering (*25*) (supplemental text) (Fig. 2D-I). The background risk of gang violence is characterized by numerous compact, but widely distributed hot spots (Fig. 2D), consistent with the observation that the opportunities for violence and strengths of gang rivalries are geographically variable (*8, 26*). The risk of retaliation is concentrated in more continuous bands (Fig. 2G), bridging the discrete areas of background risk. Notably there is a prominent North-South corridor of retaliatory risk that maps to an area locally known as 'death alley' (*27*). The patterns of risk influence the distribution of gang violent crimes (Fig. 2F and 2I). The density of background crimes forms five distinct hot spots (Fig. 2E) suggesting that background crimes are of local, neighborhood origin. The density of retaliatory crimes occupies only two distinct hot spots (Fig. 2H) suggesting that retaliation spreads contagiously beyond immediate neighborhood contexts.

Stochastic declustering (*25*) also allows us to evaluate differences in the frequency of retaliation by crime type across test conditions (Table 2). Background crimes make up 76.6% of all gang aggravated assaults and homicides for both test conditions combined. Retaliatory crimes are proportionally more common among events assigned to the LAPD-only control condition. This imbalance is pronounced for gang aggravated assaults, (46.3% retaliation for LAPD-only vs. 10.3% for LAPD + GRYD IR), but particularly extreme for homicides (24.2% LAPD-only vs. 0.7% for LAPD + GRYD IR), especially considering the baseline bias towards notifying GRYD IR of most gang homicides.

Table 2. Number of retaliatory and background aggravated assaults and homicides in South Los Angeles in 2014-2015 separated by test condition.

	LAPD-only				LAPD + GRYD IR			
	Retaliation	Background	TOTAL	% Retaliation	Retaliation	Background	TOTAL	% Retaliation
Aggravated Assault	395	854	1249	46.3%	44	426	470	10.3%
Homicide	8	33	41	24.2%	1	151	152	0.7%
TOTAL	403	887	1290	45.4%	45	577	622	7.8%

We use the estimated treatment effects along with the results of stochastic declustering to compute numbers of prevented crimes (Table 2). The sum (k_{11} + k_{01}) is the average number of retaliations known to both the LAPD and GRYD IR produced by two types of triggers (Fig. 1B). Similarly, (k_{10} + k_{00}) is the average number of retaliations known only to the LAPD produced by two types of triggers. Note that (k_{11} + k_{01}) and (k_{10} + k_{00}) are measured directly from data and therefore are observed outcomes. We now define two counterfactuals. Let (k_{01} + k_{01}) be the average number of retaliations that would have been triggered in the absence of GRYD IR notification for those events assigned to LAPD + GRYD IR. Let (k_{00} + k_{00}) be the average number of retaliations that would have been triggered in the absence of GRYD IR notification for those events assigned to LAPD-only. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. From stochastic declustering, the observed number of gang retaliations arising from pathways (k_{11} + k_{01}) is 45 and from (k_{10} + k_{00}) is 403 (Table 2). The counterfactual conditions suggest that retaliatory gang crimes would have been 48.8% and 15.0% higher in the absence of GRYD IR for observed pathways (k_{11} + k_{01}) and (k_{10} + k_{00}), respectively (supplemental text). GRYD IR prevented an estimated total 82.2 retaliatory gang crimes, of which 77.8 are expected to have been aggravated assaults and 4.4 are expected to have been homicides. Recent estimates project the overall cost of a single aggravated assault at $240,000 and a single homicide at $8.98 million (*28*). Over the two-year period in 2014-15, the potential savings from GRYD IR in South Los Angeles alone is estimated at $49.0 million.

The multivariate self-exciting point process model presented here makes it possible quantify causal pathways to and precisely identify how GRYD IR disrupts retaliatory gang aggravated assaults and homicides. Instead of determining the causal structure of each event, productivities $K_{u_i u}$ are averaged over the entire time series, which produces very robust and stable estimates. This is particularly important in real-world empirical settings where putative control conditions are rarely, if ever isolated from putative test conditions. Rather control and test conditions are often mixed by structural and practical conditions beyond the control of the observer. The multivariate modeling framework embraces the fact that these mixtures exist and allows for interactions between conditions to proceed as part of the analysis.

References and Notes:

1. S. H. Decker, Collective and normative features of gang violence. *Justice Quarterly* **13**, 243-264 (1996).
2. M. W. Klein, C. L. Maxson, *Street gang patterns and policies*. (Oxford University Press, New York, 2006).
3. L. A. Hughes, J. F. Short, DISPUTES INVOLVING YOUTH STREET GANG MEMBERS: MICRO-SOCIAL CONTEXTS*. *Criminology* **43**, 43-76 (2005).
4. B. A. Jacobs, R. T. Wright, *Street justice: Retaliation in the criminal underworld*. (Cambridge University Press, 2006).
5. B. Bjerregaard, A. J. Lizotte, Gun Ownership and Gang Membership. *The Journal of Criminal Law and Criminology (1973-)* **86**, 37-58 (1995).
6. J. C. Howell, *Gangs in America's communities*. (Sage, 2012).
7. R. Rosenfeld, T. M. Bray, A. Egley, Facilitating violence: A comparison of gang-motivated, gang-affiliated, and nongang youth homicides. *Journal Of Quantitative Criminology* **15**, 495-516 (1999).
8. A. V. Papachristos, Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. *Am. J. Sociol.* **115**, 74-128 (2009).
9. E. A. Stewart, R. L. Simons, RACE, CODE OF THE STREET, AND VIOLENT DELINQUENCY: A MULTILEVEL INVESTIGATION OF NEIGHBORHOOD STREET CULTURE AND INDIVIDUAL NORMS OF VIOLENCE*. *Criminology* **48**, 569-605 (2010).
10. W. G. Skogan, S. M. Hartnett, N. Bump, J. Dubois, *Evaluation of CeaseFire-Chicago*. (U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington, D.C., 2009).

11. D. W. Webster, J. M. Whitehill, J. S. Vernick, F. C. Curriero, Effects of Baltimore's Safe Streets Program on Gun Violence: A Replication of Chicago's CeaseFire Program. *Journal of Urban Health* **90**, 27-40 (2013).
12. J. M. Whitehill, D. W. Webster, S. Frattaroli, E. M. Parker, Interrupting Violence: How the CeaseFire Program Prevents Imminent Gun Violence through Conflict Mediation. *Journal of Urban Health* **91**, 84-95 (2014).
13. J. A. Butts, C. G. Roman, L. Bostwick, J. R. Porter, Cure violence: a public health model to reduce gun violence. *Annual review of public health* **36**, 39-53 (2015).
14. M. W. Klein, Comprehensive gang and violence reduction programs. *Criminology & Public Policy* **10**, 1037-1044 (2011).
15. A. V. Papachristos, Too big to fail. *Criminology & Public Policy* **10**, 1053-1061 (2011).
16. D. Marsan, O. Lengline, Extending earthquakes' reach through cascading. *Science* **319**, 1076-1079 (2008).
17. G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, G. E. Tita, Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106**, 100-108 (2011).
18. A. G. Hawkes, Spectra of Some Self-exciting and Mutually Exciting Point Processes. *Biometrika* **58**, 83-& (1971).
19. M. B. Short, G. O. Mohler, P. J. Brantigham, G. E. Tita, Gang Rivalry Networks Via Coupled Point Process Networks. *Discrete and Continuous Dynamical Systems* **19**, 1459-1477 (2014).
20. B. Green, T. Horel, A. V. Papachristos, Modeling contagion through social networks to explain and predict gunshot violence in chicago, 2006 to 2014. *JAMA Internal Medicine*, (2017).
21. E. Lai *et al.*, Topic Time Series Analysis of Microblogs. *IMA Journal of Applied Mathematics,* **81**, 409-431 (2016).
22. A. Veen, F. P. Schoenberg, Estimation of Space? Time Branching Process Models in Seismology Using an EM? Type Algorithm. *Journal of the American Statistical Association* **103**, 614-624 (2008).
23. A. Bray, K. Wong, C. D. Barr, F. P. Schoenberg, Voronoi cell based residual analysis of spatial point process models with applications to Southern California earthquake forecasts. *Annals of Applied Statistics* **8**, 2247-2267 (2014).
24. J. S. Gordon, R. A. Clements, F. P. Schoenberg, D. Schorlemmer, Voronoi residuals and other residual analyses applied to CSEP earthquake forecasts. *Spatial Statistics* **14, Part B**, 133-150 (2015).
25. J. Zhuang, Y. Ogata, D. Vere-Jones, Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association* **97**, 369-380 (2002).
26. G. Tita, G. Ridgeway, The Impact of Gang Formation on Local Patterns of Crime *Journal of Research in Crime and Delinquency* **44**, 208-237 (2007).
27. N. Santa Cruz, K. Schwencke, in *Los Angeles Times*. (2014), chap. The Homicide Report.
28. K. E. McCollister, M. T. French, H. Fang, The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and Alcohol Dependence* **108**, 98-109 (2010).

Acknowledgments:

This research was made possible by the City of Los Angeles Mayor's Office of Gang Reduction and Youth Development (GRYD) and Los Angeles City contract #C-128086 ("Research Consulting for GRYD"). The authors thank GRYD Director Anne Tremblay, GRYD Office Staff and GRYD Intervention Incident Response Providers. We are also grateful to the Los Angeles Police Department for providing crime data. Thank you to California State University Los Angeles and University of California Los Angeles for their ongoing support. Any opinions, findings, conclusions or recommendations expressed in this study, however, are those of the authors and do not necessarily reflect the views of the GRYD Office.

Supplementary Materials:
Materials and Methods
Figures S1-S8
Tables S1-S2
References (*1-15*)

Supplemental Materials for

Immediate community intervention in response to gang violent crime directly reduces gang retaliation

Authors: P. Jeffrey Brantingham[1]*, Baichuan Yuan[2], Nick Sundback[4], Frederic Paik Schoenberg[3], Andrea Bertozzi[2], Joshua Gordon[3], Jorja Leap[4], Kristine Chan[5], Molly Kraus[5], Denise Herz[5].

Affiliations:
[1]Department of Anthropology, University of California Los Angeles, Los Angeles, CA 90095, USA.
[2]Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA.
[3]Department of Statistics, University of California Los Angeles, Los Angeles, CA 90095, USA.
[4]Luskin School of Public Affairs, University of California Los Angeles, Los Angeles, CA 90095, USA.
[5]School of Criminal Justice and Criminalistics, California State University Los Angeles, Los Angeles, CA 90032, USA.

*Correspondence to: branting@ucla.edu

This PDF File Includes:

Materials and Methods
Figures S1-S8
Tables S1-S2
References (1-15)

Supplementary Materials:

Materials and Methods:

Table of Contents

   a. Multivariate self-exciting point process model
   b. Expectation Maximization
   c. Nonparametric methods and Voronoi residuals
   d. Stochastic declustering
   e. Estimating statistical significance
3. Results
   a. Gang vs. non-gang crime in South Los Angeles
   b. LAPD-only gang and non-gang crime vs. LAPD + GRYD IR gang and non-gang crime
   c. LAPD-only gang crime vs. LAPD + GRYD IR gang crime
   d. Nonparametric Model Estimates and Voronoi Residuals
   e. Estimating prevented crimes

Materials

Data Sources. The analyses presented rely on data collected by the Los Angeles Police Department (LAPD) as well as the City of Los Angeles Mayor's Office of Gang Reduction Youth Development (GRYD). The latter data collection is referred to as the Efforts to Outcomes (ETO) database. Data provided by the LAPD include only officially reported crimes that that have been through the Department's standard process of verification and quality control. Neither calls for service data, nor suspect and arrest data were used. The LAPD data includes records for all reported crime types ranging from public disorder to homicide. Most of these crime types are not directly relevant to understanding the dynamics of gang violence and the impact of GRYD Intervention Incident Response (GRYD IR) (see below).

The GRYD ETO database tracks all instances where the GRYD Office was notified of a crime that could warrant GRYD IR action. The GRYD ETO database includes the following crime types:

- Homicide
- Multiple Victim Shooting
- Single Victim Shooting
- Stabbing
- Shots Fired
- N/A or Unknown
- Other

These crime types do not align perfectly with LAPD (or California Penal Code) crime types. Single victim shooting, multiple victim shootings and stabbings in GRYD ETO data are all classified as aggravated assault (assault with a deadly weapon) in the LAPD data. GRYD crime types N/A or Unknown and Other can align in multiple ways with the LAPD. We adopt the LAPD crime type classification to ensure fair comparison with the entire LAPD dataset.

In addition to these basic pieces of information, the ETO database includes GRYD's own evaluation whether an event is a gang crime. The evaluation is based in part on information provided by CIWs in the field. Such an evaluation is separate from, but likely influenced by the details about the event relayed to GRYD by the LAPD. We take an inclusive approach wherein we treat any event as a gang crime that either LAPD or GRYD labels as such.

Defining Gang Violence. The California Penal Code (CPC § 186.22) provides guidelines for attaching a gang enhancement to any crime. Any crime that is committed for the benefit of a gang is eligible to be labeled as a gang crime, and individuals convicted of committing such crimes may be subject to enhanced sentences. How a gang label is applied in practice by police to individual crimes, however, is not specified by the criminal code. In general, whether a crime is labeled as gang-related much depends upon the circumstances of the crime.

The gang-related label is complicated both by variation in the ways in which gang members and gang activity may be associated with an individual crime events and by potential variation in the criteria used by raters to determine if a crime qualifies as gang related. In general, crimes may be simply gang-affiliated, meaning that the victims and/or perpetrators of the crime are known or suspected gang members, but the crime itself is not a direct results of gang activities (*1*). By contrast, crimes may be gang-motivated, meaning that the crime was committed to further some social or economic goal of the gang, including establishing or maintaining reputation.

A range of criteria might be used by police or intervention workers to determine whether a specific crime is gang-related. In the case of GRYD IR, a determination that an event is gang related is based on basic intelligence about whether: (1) the event was gang motivated; (2) the event occurred in a gang area; (3) the event featured gang involved or affiliated individuals; (4) recent activity occurred between the victim's and/or suspect's gangs, or (5) the event has the potential for retaliation. Different raters may weigh these criteria to different degrees.

Among the crime events known to the Los Angeles Police Department (LAPD) in 2014-2015, 90 different crime types were labeled as gang crimes. A majority of the crime types so labeled may be simply gang-affiliated crimes where the offender happens to also be a gang member. Table S 1 shows the top 15 crime types sorted in descending order by the percentage that bear a gang label. Gang homicides made up 59.9% of all homicides in Los Angeles in 2014-2015. Gang aggravated assaults (also sometimes referred to as assault with a deadly weapon) represented a greater number of events by volume (n =3,805), but constituted only 20.8% of all aggravated assaults recorded during this time frame. By contrast, 171 burglaries (0.6%) were labeled as gang crimes in 2014-2015. Burglaries are thus rarely labeled as gang crime and, in fact, there is little reason to think that even this handful of events had much to do with gang rivalries and the risk of retaliatory gang violence. Rather the suspects or victims in the reported crimes may have been gang-affiliated, which prompted the application of the label. Our analyses therefore focus on gang aggravated assaults and homicides, which are more frequently committed for the benefit of the gang and are particularly prone to initiating retaliation. We adopt an inclusive approach and treat any aggravated assault or homicide as gang-related if the LAPD or GRYD program labels it as such.

Table S 1. Top 15 gang crime types by volume in 2014-2015 for all LAPD data.

Rank†	Crime Type	Gang		Non Gang		TOTAL
		N	%	N	%	N
1	Criminal Homicide	325	59.9%	218	40.1%	543

2	Shots Fired at Inhabited Dwelling	174	35.9%	311	64.1%	485
3	Aggravated Assault	3,805	20.8%	14,500	79.2%	18,305
4	Discharge Firearms/Shots Fired	144	19.2%	605	80.8%	749
5	Attempted Robbery	294	13.8%	1,834	86.2%	2,128
6	Robbery	1,904	12.9%	12,889	87.1%	14,793
7	Violation of Court Order	419	11.1%	3,355	88.9%	3,774
8	Brandish Weapon	177	9.6%	1,659	90.4%	1,836
9	Criminal Threats - No Weapon Displayed	912	8.0%	10,551	92.0%	11,463
10	Other Miscellaneous Crime	171	4.5%	3,591	95.5%	3,762
11	Vandalism - Felony	639	3.0%	20,586	97.0%	21,225
12	Vandalism - Misdemeanor	307	1.7%	17,261	98.3%	17,568
13	Battery - Simple Assault	538	1.5%	35,550	98.5%	36,088
14	Spousal Abuse - Simple Assault	270	1.1%	24,115	98.9%	24,385
15	Burglary	171	0.6%	28,526	99.4%	28,697

† Rank order based on percentage of events labeled as gang-related.

GRYD and GRYD IR. The City of Los Angeles Mayor's Office of Gang Reduction Youth Development (GRYD) deploys a comprehensive strategy aimed at reducing gang involvement and gang violence. It focuses on the provision of gang prevention and intervention services, violence interruption activities and proactive peace-making. Prevention services are aimed at providing alternatives for youth before they join gangs. Intervention services are aimed at provide pathways out of gang life for youth that have already become involved. Violence interruption seeks to disrupt retaliation in the aftermath of gang violent crimes. Proactive peacemaking reflects continuous efforts of Community Intervention Workers (CIWs) to tamp down general community tensions not tied to any one event. Violence interruption within the GRYD program is termed GRYD Intervention Incident Response (GRYD IR) and is the focus of our analysis. The GRYD Office was first established in 2007, community-based service provision began in 2009, and the GRYD comprehensive strategy was created in 2011. The GRYD program design overlaps in part with Chicago's Operation Ceasefire (now Cure Violence) (*2*). GRYD is not a law enforcement strategy and therefore differs from the "pulling levers" model (*3*).

GRYD IR uses a near real-time notification system connecting the LAPD, the GRYD Program Office and CIWs embedded in the communities. Notification and information sharing does not interfere with the principal parties taking independent actions in response to gang crime incidents. The LAPD follows standard law enforcement and investigative protocols in responding to gang crimes. The GRYD Program Office aggregates information and makes allocation decisions for non-law enforcement intervention resources. CIWs engage the community through direct social contacts. CIWs see crisis response in the immediate aftermath of gang violent crimes and rumor control as a critical tool for reducing retaliation.

South Los Angeles gang violence & GRYD IR Zones.

We focus our analysis on violent crimes occurring in an 87.2 km^2 (33.7 sq mi) area of South Los Angeles during 2014-2015 (Fig. S1). As of mid-2015, ten GRYD Zones operated in South Los Angeles forming a contiguous case study area. Prior to mid-2015, only seven of the ten regions shown were formally recognized. Nevertheless, GRYD IR received notifications and responded to crimes over this entire geographic area throughout 2014 and 2015. For example, between July 1, 2014 and December 31, 2014, GRYD IR was notified of 109 events over the South Los Angeles region (Fig. S2). Between July 1, 2015 and December 31, 2015, GRYD IR was notified of 129 events over this region, only an 18% increase in notifications that occurred against a backdrop of increasing violent crime across the region. The 77th 3 GRYD Zone is the only possible exception. This GRYD Zone received few notifications prior to its formal addition. In all other locations, GRYD IR crimes were recorded regardless of whether there was a formal GRYD Zone in place or not.

We therefore treat the South Los Angeles GRYD Zones as a single continuous study region for 2014 and 2015. This region is well bounded, but still expansive enough to understand the spatial dynamics of gang retaliatory violence. There is no official tally of the number of gangs present in the area. The GRYD ETO database notes 54 unique gang names in association with the suspects and victims of gang crimes in the area. This estimate is likely a lower bounds as many events lack information about suspect and victim gang affiliation. Taken at face value, the area hosts 0.6 gangs per km^2. During this period, there were 5,982 aggravated assaults and homicides crimes reported to the LAPD strictly inside the South Los Angeles GRYD Zones (main text Table 1). Of these, 1,912 were flagged as gang related. GRYD IR was notified on 745 of the violent crimes, with

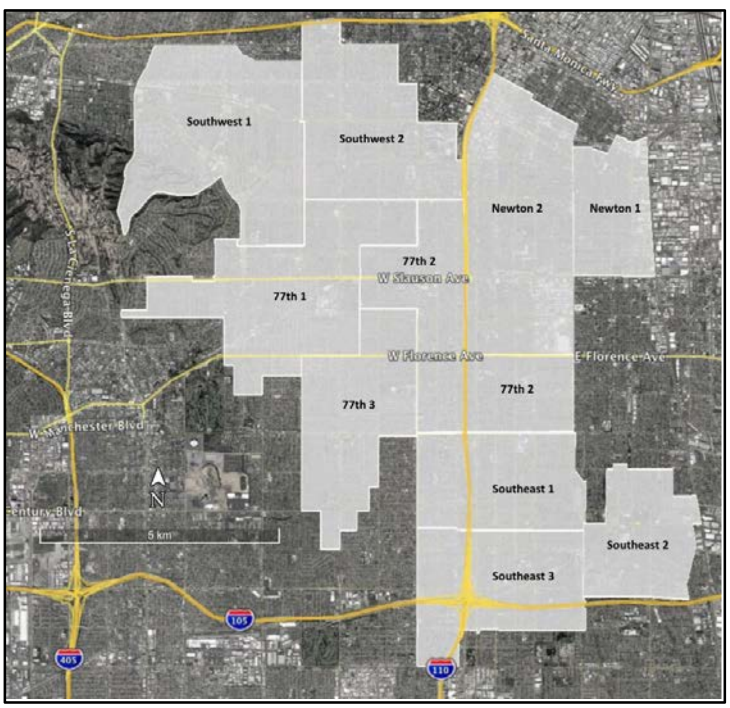


Fig. S1. Google Earth map of South Los Angeles showing the ten GRYD IR Zones in operation as of mid-2015.

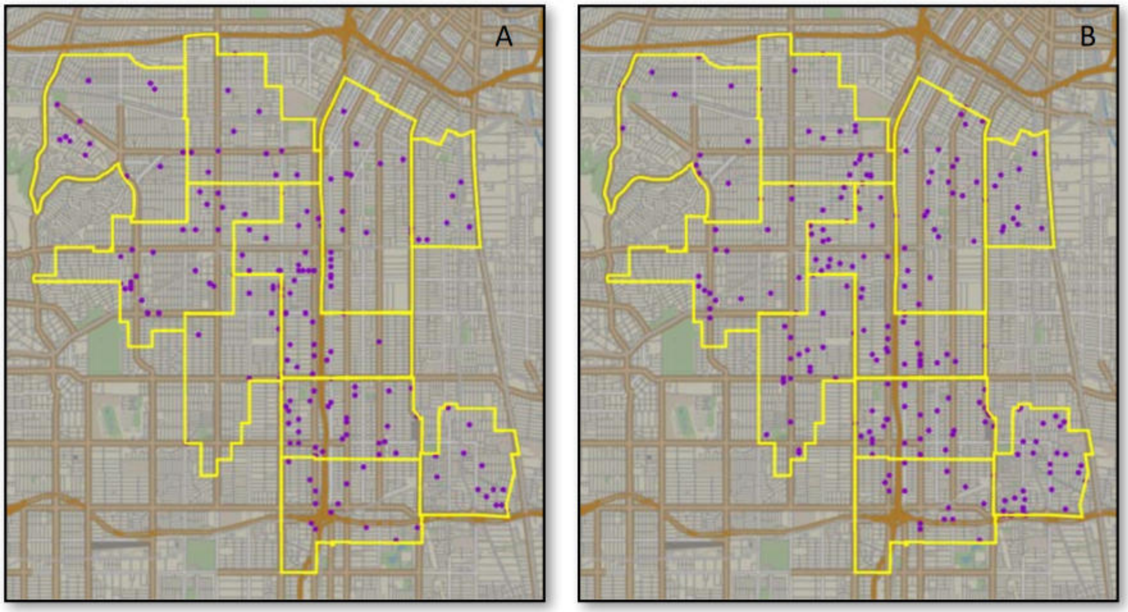


Fig. S2. Locations of GRYD IR Violent Crimes (purple) in a comparable six-month period before (A) and after (B) the July 2015 expansion of GRYD Zones in South Los Angeles. The time periods cover July-December 2014 (A) and July-December 2015 (B).

Runs tests and K-S test. GRYD IR was notified on only a fraction of the total number of gang crimes occurring in the South Los Angeles area. We hypothesized that the notifications were random and independent, establishing test conditions approximating a randomized controlled trial. We performed one-sample runs tests to evaluate whether notifications received by GRYD IR occur at random. We ran two-sample KS tests to test whether GRYD IR and LAPD events have the same spatial-temporal distribution, based on the distance between the empirical distribution functions of the two samples.

Consider a gang crime reported to the LAPD. With probability ρ, GRYD IR is notified of the event. We use LAPD + GRYD IR as a short-hand to identify these events. With probability $1 - \rho$ the crime remains known only to the LAPD. These events are noted as LAPD-only. Over a large sequence of reported gang crimes, we expect a fraction of crimes proportional to ρ to fall into condition LAPD + GRYD IR and a fraction proportional $1 - \rho$ to fall into condition LAPD-only. However, we also expect the assignment of any one event to be random and independent of other assignments. To test this hypothesis, reported gang crimes were arranged in ascending order of the date and time received and computed a runs test. A run is defined as a sequence of notifications of the same condition, uninterrupted by notifications of the other condition. For example, the sequence of coin tosses HHTHHHTTTT consists of four runs {HH}{T}{HHH}{TTTT}. Let R be the observed number of runs in a sample of size N. The null hypothesis is that the observed number of runs R is the product of random and mutually independent notifications. Note that ρ need not equal $1 - \rho$ as would be the case in fair coin toss. All that is required is that each GRYD IR notification occur at probability ρ independently.

Results for different subsets of the observed data fail to reject the null hypothesis (Table S2). For gang aggravated assaults and homicides combined, the number of observed runs (R = 852) is not statistically different from the expected number of runs under random assignment (E[R] = 840.31) (Z = 0.609, p = 0.54). For gang aggravated assaults alone, the observed number of runs

(R = 672) is not statistically different from the expected number under random assignment (E[R] = 683.99) (Z = 0.73, p = 0.47). Similarly, for gang homicides the observed number of runs (R = 65) is not statistically different from the expected number under random assignment (E[R] = 65.58)(Z = 0.125, p = 0.90). While the percent of crimes that lead to GRYD IR notifications differs by crime type, these notifications are random on a per event basis. Furthermore, the results of two-sample KS tests fail to reject the null hypothesis that the data come from the same underlying temporal (KS = 0.069, p = 0.20) and spatial distributions (KS = 0.054, p = 0.17) Therefore, we conclude that the assignment of events to different conditions approximates and randomized experimental protocol.

Table S2. Results of one sample runs tests for three different categories of gang crimes in South Los Angeles.

	Gang Aggravated Assault + Homicide	Gang Aggravated Assault	Gang Homicide
LAPD + GRYD IR (n1)	622	470	152
LAPD-only (n2)	1290	1249	41
Total (N)	1912	1719	193
Percent LAPD + GRYD IR	32.5%	27.3%	78.7%
Runs	852	672	65
Expected Mean	840.31	683.99	65.58
Expected SD	19.18	16.47	4.62
Z	0.609	0.73	0.125
two-tailed p-value	0.54	0.47	0.90

GRYD IR Notification and Field Deployment. Gang crimes are typically reported to the LAPD first by members of the public. GRYD IR receives notification from the LAPD and from CIWs, when they are independently contacted by the community. However, notification of GRYD IR only occurs for a fraction of all reported crimes. We include all events where GRYD IR is notified, rather than restricting analysis to events where there is also some record of field activity by CIWs. This analytical choice is made due to uncertainty surrounding the measurement of dosage associated with CIW activities (see also *2: 5-3*)

The resulting hypotheses based on notification of GRYD IR are conservative. On the one hand, we suppose that CIWs can only have a direct impact on retaliation if they know about a potential triggering event. Thus, notification of GRYD IR is a logical precursor for treatment effects. If notification did not lead to interventions in the field by CIWs then we would expect there to be no difference between the control and treatment conditions. Specifically, even though events were labeled as LAPD + GRYD IR, the absence of field intervention would ensure that such events were no different than LAPD-only one. On the other, if CIWs were able to source information about events on their own and self-deploy, operating outside of the normative channel of communication, then we would expect this contamination to also bias the results towards finding no difference between treatment and control. In other words, some fraction of events labeled as LAPD-only would include the effects of GRYD IR without formal recognition. Thus, LAPD-only interventions would appear similar to LAPD + GRYD because of this hidden activity. Because we are able to document a significant difference between LAPD-only and LAPD + GRYD IR intervention effects (main text Fig. 2C) we conclude such confounds have a minimal impact.

Methods

Multivariate self-exciting point process models. Self-exciting point process models provide a useful mathematical structure for considering the dynamics of retaliatory crime (*4, 5*). The problem we analyze requires that we treat gang retaliation in both space and time as well as events exposed to different experimental conditions. We use a multivariate spatial-temporal self-exciting point process model to capture these conditions:

$$\lambda_u(t,x,y) = \mu_u(x,y) + \sum_{t_i<t} K_{u_i u} g(x-x_i, y-y_i, t-t_i).$$

The equation describes the instantaneous rate λ_u at which crimes assigned to condition u. In our specific case, crimes are assigned to two different conditions, those known only to the LAPD and those known to both the LAPD and GRYD IR. Notice that this is a spatial-temporal model. The random background rate at which gang crimes occurs may vary from place to place $\mu(x, y)$, but not in time. The self-exciting triggering function is also spatially dependent. The magnitude of excitation at time t and location x, y dependens not on how long ago prior crimes $i = 1, 2, \ldots, n$ occurred $(t - t_i)$, but also how nearby in space $(x - x_i, y - y_i)$. Gang crimes are more likely to trigger a retaliation soon after the initial event and nearby in space. We use particular parametric equations for μ and g:

$$\mu_u(x,y) = \sum_{i=1}^{N} \frac{\beta_{u_i u}}{2\pi\eta^2 T} \times \exp\left(\frac{-\left((x-x_i)^2 + (y-y_i)^2\right)}{2\eta^2} \right), \tag{S1}$$

$$g(x,y,t) = \omega \exp(-\omega t) \times \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2} \right). \tag{S2}$$

The equation for μ treats the stationary background rate of crimes of type u as a sum of Gaussian kernels. The triggering kernel g treats the self-exciting effect of one event as decaying exponentially in time and Gaussian in space. Here β is a weights matrix for the degree to which events assigned to condition u_i contribute the background rate for events assigned to condition u. The parameter T is the total time period represented by the sample of gang crimes, which in this case is 2014-2015. The parameters η and σ represent the spatial scale of influence for background events and retaliatory events, respectively (see *6*). We take $\eta = \sigma$, which simplifies parameter estimation.

Our quantities of interest are estimates of four key relationships encoded in matrix $K_{u_i u}$: (1) k_{11} or the average number of LAPD + GRYD IR gang crimes triggered by a single LAPD + GRYD

IR gang crime; (2) k_{01} the average number of LAPD + GRYD IR gang crime triggered by a single LAPD-only gang crime; (3) k_{10} or the average number of LAPD-only gang crimes triggered by a single LAPD + GRYD IR gang crime; and (4) k_{00} or the average number of LAPD + GRYD IR gang crime triggered by a single LAPD-only gang. The comparisons of interest are between pathways that trigger the same type of retaliation. Thus we are interested in the difference between the two pathways that trigger LAPD + GRYD IR retaliations (Fig. S3A). Separately, we are interested in the difference between the two pathways that trigger LAPD-only retaliations (Fig. S3B). Fig. S3 flags in red the events known to GRYD IR and therefore the pathways that have the potential to be impacted by GRYD IR interventions. The events on the LAPD-only timeline, flagged in black, are not known to GRYD IR and therefore cannot experience a direct impact of GRYD IR interventions. If GRYD IR interventions have an impact, then we expect $k_{11} < k_{01}$ and $k_{10} < k_{00}$. In other words, fewer retaliations will follow GRYD IR interventions, whether or not GRYD IR knows about those retaliations.

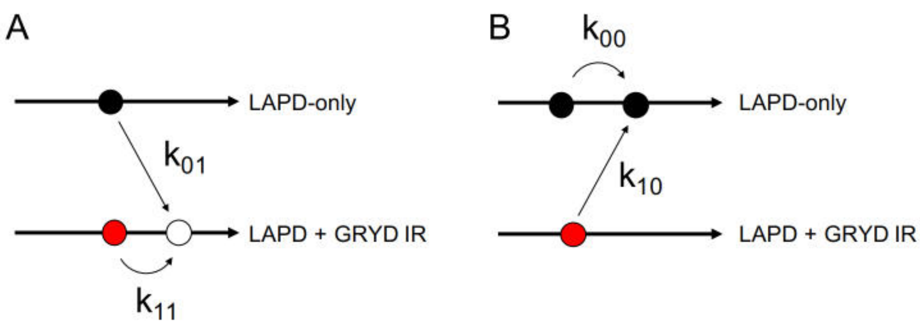


Fig. S3. Triggering pathways separated by the experimental condition for resulting retaliations. (A) Retaliation is known to both LAPD and GRYD IR. (B) Retaliation is known only to the LAPD. Crimes where the treatment intervention effect can have an impact are flagged in red.

The estimation procedure we use is a type of Maximum Likelihood Estimation (MLE) known as expectation maximization (EM) (*7, 8*). The expectation step of the EM algorithm is used to compute initial probabilities p^b_{ij} and p_{ij} that an event *i* causes event *j* via either the background rate μ or the self-exciting kernel g, respectively. These expectations are then fed to the maximization step where a new set of parameter values (for iteration $k + 1$) are determined by maximizing the expected probability with respect to the observed data. This maximization is done for all parameters taking into consideration whether gang crimes are known only to the LAPD or to both the LAPD and GRYD IR. The algorithm alternates between expectation and maximization until there is no further change in the parameter values.

For completeness, the EM algorithm is structured as below. Note that $n_{\hat{u}}$ is the number of events that belongs to experimental condition type $\hat{u}$:

Complete Data Likelihood Function:
$Q(\Omega)=\sum_{i=1}^{N}\sum_{j=1}^{N}p_{ij}^{b}log(\frac{\beta_{u_iu_j}}{2\pi\eta^2 T}e^{\left(-\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2\eta^2}\right)})-\sum_{u=1}^{U}\sum_{i=1}^{N}\beta_{u_iu}$ $+\sum_{i<j}p_{ij}log(\omega K_{u_iu_j}e^{-\omega(t_j-t_i)}\frac{1}{2\pi\sigma^2}e^{-\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2\sigma^2}})-\sum_{u=1}^{U}\sum_{i=1}^{N}K_{u_iu}(1-e^{-w(T-t_i)})$

Expectation Step:	
$p_{ij} = K_{u_i u_j}\omega \frac{\exp\left(-\omega(t_j - t_i)\right)}{2\pi\sigma^2} \times \frac{\exp\left(-\frac{(x_j - x_i)^2 + (y_j - y_i)^2}{2\sigma^2}\right)}{\lambda_{u_j}(x_j, y_j, t_j)}$ $p_{ij}^b = \frac{\beta_{u_i u_j}}{2\pi\eta^2 T} \frac{\exp\left(-\frac{(x_j - x_i)^2 + (y_j - y_i)^2}{2\eta^2}\right)}{\lambda_{u_j}(x_j, y_j, t_j)}$	
Maximization Step:	
$\omega^{(k+1)} = \frac{\sum_{j<i} p_{ij}^{(k)}}{\sum_{i<j} p_{ij}^{(k)}(t_j - t_i) + \sum_{u=1}^{U}\sum_{i=1}^{N} K_{u_i u}(T - t_i)e^{-\omega(T - t_i)}}$ $\sigma^{2(k+1)} = \frac{\sum_{i<j} p_{ij}^{(k)}\left((x_i - x_j)^2 + (y_i - y_j)\right)^2}{2\sum_{i<j} p_{ij}^{(k)}}$ $\eta^{2(k+1)} = \frac{\sum_{i,j=1}^{N} p_{ij}^{b(k+1)}\left((x_i - x_j)^2 + (y_i - y_j)^2\right)}{2\sum_{i,j=1}^{N} p_{ij}^{b(k+1)}}$	$K_{\hat{u}u}^{(k+1)} = \frac{\sum_{i,j\in A_{\hat{u}u}} p_{ij}}{\sum_{l=1}^{n_{\hat{u}}}(1 - e^{-w(T - t_{i_l})})}$ $\beta_{\hat{u}u}^{(k+1)} = \frac{\sum_{i,j\in B_{\hat{u}u}} p_{ij}^{b(k)}}{n_{\hat{u}}}$ $A_{\hat{u}u} = \left\{i, j \text{ index of events} \mid t_i < t_j, u_i = \hat{u}, u_j = u\right\}$ $B_{\hat{u}u} = \left\{i, j \text{ index of events} \mid u_i = \hat{u}, u_j = u\right\}$

Non-parametric Model Fitting and Voronoi Residuals Analysis

Our parametric model choices were based on previous research on crime patterns indicating that exponential kernels provide a good description of the data (*4*). We extended fully non-parametric model estimation methods in (*9*) and (*10*) to the multivariate case. The non-parametric model is similar to the parametric form:

$$\lambda_u(x,y) = \mu_k(x,y) + \sum_{t_i<t} K_{u_i u} v(t - t_i, x - x_i, y - y_i).$$

But here

$$\mu_u(x,y) = \gamma_u \tau(x,y) = \frac{\gamma_u}{T}\sum_{i=1}^{N}\frac{p_{ii}}{2\pi d_i^2} exp\left(-\frac{(x - x_i)^2 + (y - y_i)^2}{2d_i^2}\right)$$

and we assume that $v(x, y, t) = g(t)f(x, y)$, which will be estimated non-parametrically. The term d_i us computed by finding the radius of the smallest disk centered at (x_i, y_i) that contains at least n_p other events, and is greater than some small value ϵ representing the location error. In (*11*) they suggest taking n_p between 15-100 and $\epsilon = 0.02$ degrees.

The log-likelihood function is:

$$l = \sum_{u=1}^{U} \left(\sum_{i=1}^{N} log(\lambda_u(t_i, x_i, y_i)) - \int_0^T \iint_S \lambda_u(t, x, y) dsdt \right).$$

From EM we determine the nonparametric algorithm. We define p_{ij} as the probability that event i triggers j for $t_i < t_j$ and p_{ii} as the probability that i is from background and $p_{ij} = 0$ for $t_i > t_j$. We define n_t^{bins} as the number of bins in time and n_r^{bins} as the number of bins in space. C_k is the set of events pairs (i, j) such that t_j - t_i belongs to the k^{th} bin. D_k is the set of events pairs (i, j) such that r_{ij} is the distance between i and j belongs to the k^{th} bin. N_α is the number of events with type u. Finally, δ_t is the size of k^{th} bin in time and δ_r is the size of the k^{th} bin in space. Further discussion of parameters can be found in (*10*).

The algorithm for our nonparametric method is:

- **Step 1:** Initialize the $P^{(0)} = (p_{ij})$ matrix randomly, index $v = 0$.

- **Step 2:** Update

  $$\gamma_u^{(v)} = \frac{\sum_{u_i = u} p_{ii}^{(v)}}{Z^{(v)}}$$

  where u_i is the type of event i and

  $$\frac{1}{Z^{(v)}} \int_0^T \iint_S \tau(x, y) dsdt = 1.$$

- **Step 3:** Update

  $$K_{\alpha\beta}^{(v)} = \frac{\sum_{u_i=\alpha} \sum_{u_j=\beta} p_{ij}}{N_\alpha}, \quad g_k^{(v)} = \frac{\sum_{i,j \in C_k} p_{ij}^{(v)}}{\delta t_k \sum_{i,j} p_{ij}^{(v)}}, \quad \text{and } h_k^{(v)} = \frac{\sum_{i,j \in D_k} p_{ij}^{(v)}}{\delta r_k \sum_{i,j} p_{ij}^{(v)}},$$

  where $k = 1, \ldots, n_t^{bins}$ and $k = 1, \ldots, n_r^{bins}$ for $g_k^{(v)}$ and $h_k^{(v)}$, respectively.

- **Step 4:** Update

$$p_{ij}^{(v+1)} = K_{u_i u_j}^{(v)} g^{(v)}(t_j - t_i) f^{(v)}(r_{ij}) \quad \text{for } t_i < t_j,$$

and

$$p_{jj}^{(v+1)} = \mu_{u_j}(x_j, y_j).$$

Then normalize such that for any j $\sum_{i=1}^{N} p_{ij} = 1$. Here $2\pi r f^{(v)} = h(r)$.

- **Step 5:** If $max_{ij} \left\| p_{ij}^{(v+1)} - p_{ij}^{(v)} \right\| < \epsilon$, then the algorithm has converged. In practice we take $\epsilon = 10^{-3}$. Otherwise, set $v \leftarrow v + 1$ and repeat Steps 2–5 until convergence.

Once a model is estimated, a powerful technique for evaluating model performance is Voronoi residuals (*12*). Voronoi residuals allow the examination of differences between the modeled conditional intensity and the observed number of points within spatially adaptive Voronoi cells. Using color scaling allows us to see the spatial locations where a model is over or under estimated. Cells with blue shades indicate an overestimation of the intensity while cells with red indicate an underestimation. In addition, comparison to a color scale defined by a null Poisson model further helps interpret performance (*13*).

Stochastic declustering. Gang crimes occurring in a given area represent a mixture of those that are background events and those that are retaliatory in response to other crimes. We wish to sort events into these two groups to understand how important background and retaliatory processes are for gang violence overall.

Stochastic declustering is a suite of methods developed in the study of earthquake catalogs where the goal is to distinguish between background seismicity and aftershocks (*11*). The same methods can be applied to the study of crime (*14*).

Starting with a self-exciting point process model like the one developed here, stochastic declustering proceeds through a thinning procedure that removes events probabilistically classified as retaliations. The events remaining after thinning represent the background events generated by a spatially non-homogeneous Poisson process $\lambda(t, x, y) = \mu(x, y)$. Specifically, in the univariate case, the probability that an event j is a retaliation is given by

$$\rho_j = \frac{\sum_{t_i < t_j} g(t_j - t_i, x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j)}.$$

The probability that an event j is a background event is therefore

$$1-\rho_j=\frac{\mu(x_j,y_j)}{\lambda(t_j,x_j,y_j)}.$$

For a catalog of N total crimes and a point process model fit to those events, the simplest procedure is to generate N uniform random variables $U_1, U_2, \ldots, U_N$ in the range [0,1]. An event is classified as a background crime when $U_j < 1 - \rho_j$, otherwise it is removed and classified as a retaliation (*11*).

Note that the assignment of an event to being background or retaliation is a probabilistic classification. On average the relative mixture of background and retaliation events is correct for a given time window and spatial region, but we cannot say with absolute certainty whether any specific event is or is not a retaliation.

Estimating Statistical Significance. Our null hypothesis is that the GRYD IR interventions have no impact. That is GRYD IR does not reduce gang violence. If this null hypothesis is true then the ground truth values of the matrix K should be $k_{11} = k_{01}$ and $k_{10} = k_{00}$. Given estimates k_{11} and k_{01} and standard errors s_{11} and s_{01}, for example, the quantity

$$\frac{k_{11}-k_{01}}{\sqrt{s_{11}^2+s_{01}^2}}$$

should be t-distributed and for large samples approximately standard normal. The p-value associated with the magnitude of the observed difference between estimates of k_{11} and k_{01} can be computed directly against the cumulative distribution for the standard normal. We reject the null hypothesis based on standard probability criteria.

Results

Gang vs. non-gang crime in South Los Angeles. We adapt the multivariate self-exciting point process model to consider the retaliatory dynamics for gang aggravated assaults and homicides interacting with non-gang aggravated assaults and homicides (Fig. S5). Here we allow for the possibility that retaliation may occur within officially recognized gang crimes as well as between gang crimes and non-gang crimes. We expect gang aggravated assaults and homicides to trigger significantly more retaliations than non-gang crimes of the same type.

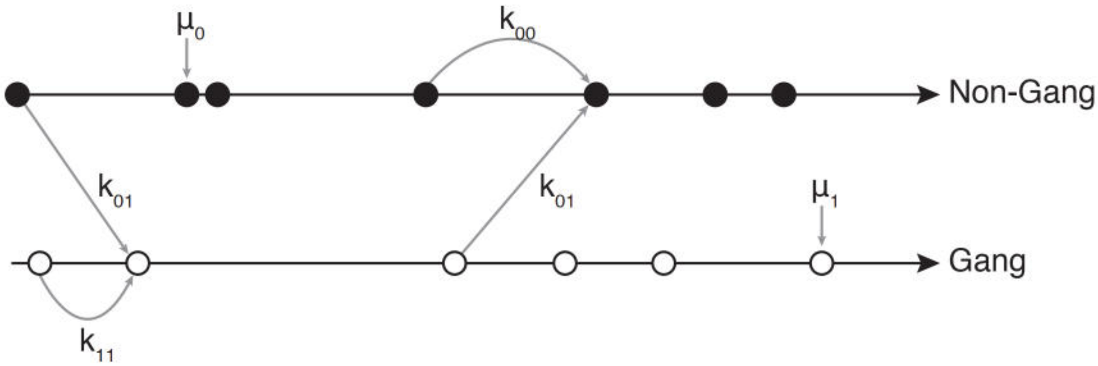


Fig. S4. The conceptual model for the multivariate self-exciting point process applied to non-gang and gang crime as the two test conditions.

Model parameters corresponding to the gang and non-gang crime conditions are as follows:

$$K_{u_i u} = \begin{pmatrix} \underset{(0.0271)}{0.2322} & \underset{(0.0034)}{0.0163} \\ \underset{(0.0015)}{0.0090} & \underset{(0.0070)}{0.1014} \end{pmatrix} \quad \beta_{u_i u} = \begin{pmatrix} \underset{(0.0149)}{0.4286} & \underset{(0.0152)}{0.3856} \\ \underset{(0.0135)}{0.1504} & \underset{(0.0044)}{0.7098} \end{pmatrix}$$

where the matrix K gives productivity for the four potential triggering relationships shown in Fig. S5 and matrix β gives the mixture weights for the spatially inhomogeneous background rate. Parameter estimates are shown along with their standard errors in parentheses. Gang crimes trigger significantly more gang retaliations than do non-gang crimes ($k_{11} > k_{01}$ or $0.2322 > 0.0090$, $p < 10^{-6}$). Gang crimes trigger significantly fewer non-gang retaliations than do other non-gang crimes ($k_{10} < k_{00}$ or $0.0163 < 0.1014$ $p < 10^{-6}$). The magnitude of within-condition triggering for gang crimes is more than double that for non-gang crimes ($k_{11} > k_{00}$). To the extent that retaliation is a defining feature of gang violence, this may suggest that some retaliatory gang crimes are mistakenly labeled as non-gang crimes. Alternatively, non-gang crimes may produce more somewhat more retaliation than is often recognized.

LAPD-only vs. LAPD + GRYD combined gang and non-gang crime. We apply the multivariate model to the case of aggravated assaults and homicides for non-gang and gang crimes combined where LAPD was notified and where LAPD and GRYD IR were both notified (Fig. S6). This is the first set of test conditions where a treatment effect from GRYD IR notification may appear.

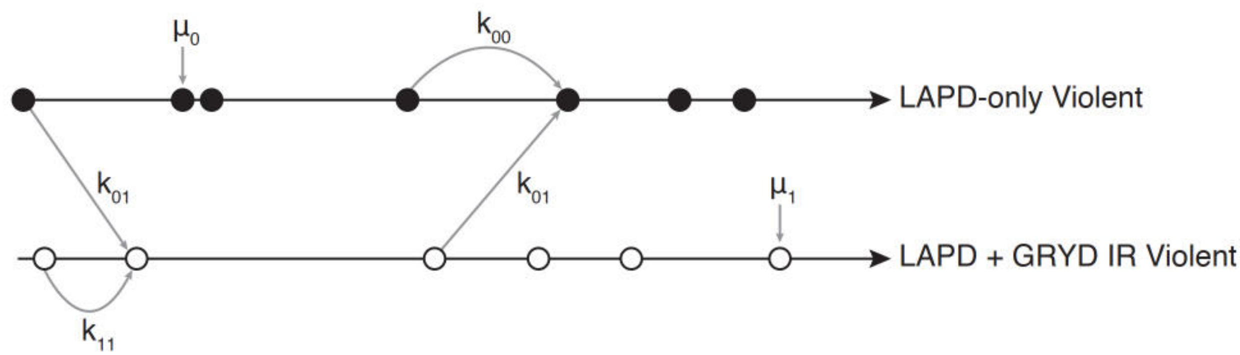


Fig. S5. The conceptual model for the multivariate self-exciting point process applied to gang and non-gang aggravated assaults and homicides known only to the LAPD and to the LAPD and GRYD IR.

Model parameters corresponding to the combined gang and non-gang crime conditions are as follows:

$$K_{u_i u} = \begin{pmatrix} \underset{(0.0362)}{0.1401} & \underset{(0.0367)}{0.2841} \\ \underset{(0.0083)}{0.0526} & \underset{(0.0346)}{0.2824} \end{pmatrix} \qquad \beta_{u_i u} = \begin{pmatrix} \underset{(0.0072)}{0.0425} & \underset{(0.0484)}{0.5898} \\ \underset{(0.0026)}{0.0831} & \underset{(0.0197)}{0.6794} \end{pmatrix}$$

The matrix K is shown as Fig. 2B in the main text. Crimes where GRYD IR is notified trigger significantly more retaliations than when GRYD IR is not notified ($k_{11} > k_{01}$ or $0.1401 > 0.0526$, $p < 0.0093$). GRYD IR crimes also trigger more LAPD-only retaliations than other LAPD-only crimes ($k_{10} > k_{00}$ or $0.02841 > 0.2824$, $p = 0.49$), though the difference is not significant. GRYD IR does not reduce retaliation relative to LAPD-only interventions when considering the entire pool of non-gang and gang aggravated assaults and homicides. This result stems from the fact that GRYD IR is notified primarily for gang-related crimes and therefore GRYD IR must confront crimes where the risk of retaliation is much greater from the outset. In other words, GRYD IR may be able to reduce retaliation among those crimes it confronts, but not to levels that characterize the broader pool of crimes that includes many more that are not gang related.

LAPD-only gang crime vs. LAPD + GRYD IR gang crime. We restrict analyses to those aggravated assaults and homicides that are flagged as gang-related (Fig. S7). The comparison of the control and treatment condition is on more equal footing for this set of crimes.

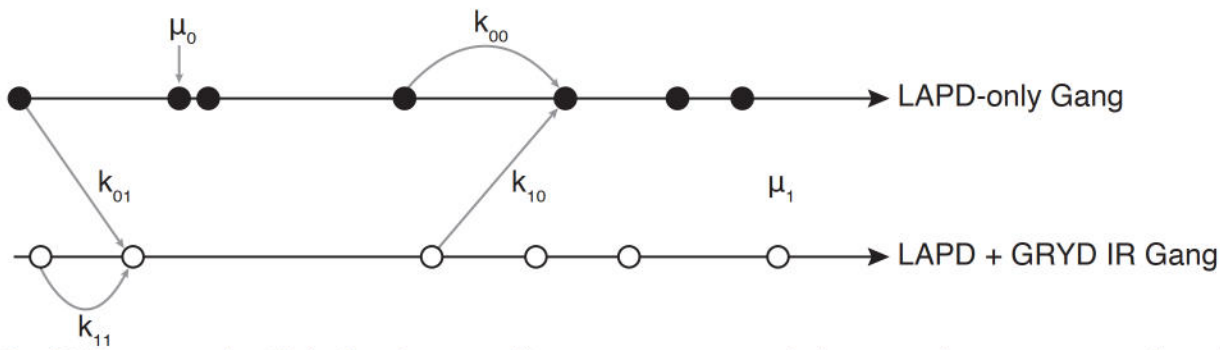


Fig. S6. The conceptual model for the multivariate self-exciting point process applied to gang and non-gang aggravated assaults and homicides known only to the LAPD and to the LAPD plus GRYD IR.

Model parameters corresponding to the gang and non-gang crime conditions are as follows:

$$K_{u_i u} = \begin{pmatrix} \underset{(0.0021)}{0.0015} & \underset{(0.0267)}{0.1483} \\ \underset{(0.0066)}{0.0621} & \underset{(0.0128)}{0.2116} \end{pmatrix} \quad \beta_{u_i u} = \begin{pmatrix} \underset{(0.0282)}{0.3360} & \underset{(0.0168)}{0.4732} \\ \underset{(0.0181)}{0.2574} & \underset{(0.0096)}{0.4889} \end{pmatrix}$$

The matrix K is shown as Fig. 2C in the main text. GRYD IR gang crimes trigger significantly fewer GRYD IR gang retaliations than LAPD-only gang crimes ($k_{11} < k_{01}$ or $0.0015 < 0.0621$, $p < 10^{-6}$). GRYD IR gang crimes also trigger fewer LAPD-only retaliations than other LAPD-only crimes ($k_{10} < k_{00}$ or $0.1483 < 0.0164$, $p = 0.49$).

Nonparametric Model Estimates and Voronoi Residuals. The nonparametric estimates for the triggering pathways for each of our cases are substantially similar to the parametric model. However, it is not possible to estimate standard errors for the nonparametric model.

Non-gang vs. Gang Crime	LAPD + GRYD IR vs. LAPD-only violent crime (gang and non-gang)	LAPD + GRYD IR vs. LAPD-only gang crime
$K_{u_i u} = \begin{pmatrix} 0.3851 & 0.1688 \\ 0.0757 & 0.3458 \end{pmatrix}$	$K_{u_i u} = \begin{pmatrix} 0.0674 & 0.2320 \\ 0.0550 & 0.3039 \end{pmatrix}$	$K_{u_i u} = \begin{pmatrix} 0.0640 & 0.2320 \\ 0.1187 & 0.3039 \end{pmatrix}$

In Fig. S6, Voronoi residuals from the fitted intensity for the proposed Hawkes model has muted colors as compared to the null model, indicating improved performance. Residuals indicate that the fitted model performed well throughout the spatial window.

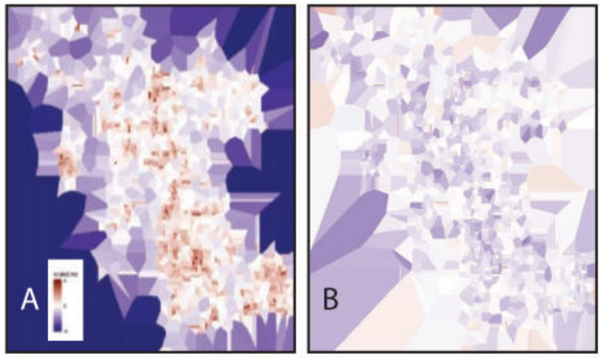


Fig. S7. Voronoi residuals for the Poisson model (A) and the Hawkes model (B).

Estimating prevented crimes. The parameter values from the spatio-temporal model can be used to estimate the number of crimes prevented by GRYD IR notifications. The sum (k_{11} + k_{01}) is the average number of retaliations known to both the LAPD and GRYD IR produced by the two types of triggers, LAPD + GRYD IR and LAPD-only. Similarly, (k_{10} + k_{00}) is the average number of retaliations known only to the LAPD produced by the two types of triggers. Note that (k_{11} + k_{01}) and (k_{10} + k_{00}) are actually measured directly from data and therefore are the observed outcome.

We can also define two counterfactual situations. Let (k_{01} + k_{01}) be the average number of retaliations that *would have occurred* in the absence of GRYD IR notification for those retaliations known to LAPD + GRYD IR. Here we simply replace k_{11} with a second instance of k_{01}. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. Similarly, let (k_{00} + k_{00}) be the average number of retaliations that *would have occurred* in the absence of GRYD IR notification for those retaliations known only to the LAPD. Again we suppose that the LAPD + GRYD IR effect is replaced by the LAPD-only effect in the absence of GRYD IR notification. We then compute the relative effect of GRYD IR notification on the average number of retaliations as:

Relative effect of GRYD IR notification on LAPD + GRYD IR retaliations	Relative effect of GRYD IR notification on LAPD-only retaliations
$\dfrac{(k_{11}+k_{01})-(k_{01}+k_{01})}{(k_{01}+k_{01})}$	$\dfrac{(k_{10}+k_{00})-(k_{00}+k_{00})}{(k_{00}+k_{00})}$

For the South Los Angeles GRYD Zones, GRYD IR notification reduces retaliation among LAPD + GRYD IR events by -48.8% relative to the counterfactual. GRYD IR notification reduces retaliation among LAPD-only events by -15.0% relative to the counterfactual. See Fig. 2C in the main text for corresponding values of k.

We use these estimated effects along with the results of stochastic declustering in South Los Angeles to compute numbers of prevented crimes. Stochastic declustering identified a total of 45 LAPD + GRYD IR gang aggravated assaults and homicides in 2014-2015 as retaliatory (see Table 2 in the main text). The remaining 577 gang aggravated assaults and homicides were statistically defined as background events. Similarly, declustering identified a total of 403 LAPD-only gang aggravated assaults and homicides in 2014-2015 as retaliatory (see Table 2 in the main text). The remaining 877 gang aggravated assaults and homicides were statistically identified as background events. The counterfactual conditions suggest that retaliatory gang aggravated assaults and homicides would have been 48.8% and 15.0% higher in the absence of GRYD IR for events recorded, respectively, as LAPD + GRYD IR and LAPD-only. Thus GRYD IR prevented an estimated total 82.2 retaliatory gang aggravated assaults and homicides. An estimate based on city-wide data, suggests that homicides make up on average 5.4% of all gang retaliations. Thus, in South Los Angeles, the prevented crimes are expected to include 4.4 retaliatory homicides and 77.8 retaliatory aggravated assaults.

We refer to McCollister et al. (*15*) for estimates of the costs of crime. In their work the total cost of a single homicide to government, victims and suspects is approximately \$8.98 million. The cost of a single aggravated assault is \$240,000. We simply multiply these figures by the estimated number of prevent aggravated assaults and homicides, respectively. The estimated number of homicides prevented by GRYD IR may add up to savings between \$39.4 million over two years. The savings from prevented gang aggravated assaults in South Los Angeles may amount to an additional \$9.5 million over two years. The combined savings per year in South Los Angeles alone may amount to \$49.0 million.

References.

1. R. Rosenfeld, T. M. Bray, A. Egley, Facilitating violence: A comparison of gang-motivated, gang-affiliated, and nongang youth homicides. *Journal Of Quantitative Criminology* **15**, 495-516 (1999).
2. W. G. Skogan, S. M. Hartnett, N. Bump, J. Dubois, *Evaluation of CeaseFire-Chicago*. (U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington, D.C., 2009).
3. A. A. Braga, D. L. Weisburd, The Effects of 'Pulling Levers' Focused Deterrence Strategies on Crime. *Campbell Systematic Reviews* **6**, (2012).
4. G. Mohler, M. Short, P. J. Brantingham, F. Schoenberg, G. Tita, Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106**, 100-108 (2011).
5. M. B. Short, G. O. Mohler, P. J. Brantigham, G. E. Tita, Gang Rivalry Networks Via Coupled Point Process Networks. *Discrete and Continuous Dynamical Systems* **19**, 1459-1477 (2014).
6. M. B. Short, P. J. Brantingham, A. L. Bertozzi, G. E. Tita, Dissipation and displacement of hotspots in reaction-diffusion models of crime. *Proceedings of the National Academy of Sciences* **107**, 3961 (2010).
7. A. Veen, F. P. Schoenberg, Estimation of Space? Time Branching Process Models in Seismology Using an EM? Type Algorithm. *Journal of the American Statistical Association* **103**, 614-624 (2008).
8. G. Mohler, Marked point process hotspot maps for homicide and gun crime prediction in Chicago. *International Journal of Forecasting* **30**, 491-497 (2014).
9. D. Marsan, O. Lengline, Extending earthquakes' reach through cascading. *Science* **319**, 1076-1079 (2008).

10. E. W. Fox, F. P. Schoenberg, J. S. Gordon, Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric Hawkes point process models of earthquake occurrences. *The Annals of Applied Statistics* **10**, 1725-1756 (2016).
11. J. Zhuang, Y. Ogata, D. Vere-Jones, Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association* **97**, 369-380 (2002).
12. A. Bray, K. Wong, C. D. Barr, F. P. Schoenberg, Voronoi cell based residual analysis of spatial point process models with applications to Southern California earthquake forecasts. *Annals of Applied Statistics* **8**, 2247-2267 (2014).
13. J. S. Gordon, R. A. Clements, F. P. Schoenberg, D. Schorlemmer, Voronoi residuals and other residual analyses applied to CSEP earthquake forecasts. *Spatial Statistics* **14, Part B**, 133-150 (2015).
14. G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, G. E. Tita, Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106**, 100-108 (2011).
15. K. E. McCollister, M. T. French, H. Fang, The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and Alcohol Dependence* **108**, 98-109 (2010).


ARC OF SCIENCE
RESEARCH TO RESULTS
EXHIBITOR

Andrea Bertozzi
University of California, Los Angeles
The Mathematics of Crime
Predictive Policing with LAPD and SCPD
PredPol
Field Interview Cards Hollenbeck LAPD
Fighting Crime
Los Angeles Police Department
Randomized Controlled Field Trials of Predictive Policing
LOS ANGELES POLICE
NBC LEARN
THE UNIVERSITY OF CALIFORNIA
LET THERE BE LIGHT
UCLA
University of California, Los Angeles
Department of Mathematics

Captain Jorge R Rodriguez
Commanding Officer LAPD, Newton Area
c/o Honore G. Rausch, Secretary
Los Angeles Police Department
Newton Community Police Station
3400 South Central Avenue
Los Angeles, California 90011

Date 17 February 2017

Our Ref. 2017/052/I/IGCI/IW2017/NN/ms

Contact Margaret Samuel
m.samuel@interpol.int

Subject **INTERPOL World 2017 – Invitation to Speak at the Congress**

Dear Captain Rodriguez,

The security landscape is evolving with the advancement of technologies. Criminals are taking advantage of technology, ease of international travel and the anonymous world of virtual business to disrupt public security and commercial stability.

INTERPOL established its Global Complex for Innovation to respond to increasing challenges in the operational landscape. The goal is to enable police worldwide to stay one step ahead of criminals by providing law enforcement capabilities and possible solutions to confront increasingly ingenious and sophisticated challenges.

As such, the INTERPOL World event was launched in 2015 to be a strategic platform for exchanges between actors *confronted* with security challenges and the actors *developing* innovative solutions for such challenges. The inaugural event saw many international manufacturers and providers showcasing innovative technologies and solutions from both public and private sectors.

INTERPOL World is supported by Singapore's Ministry of Home Affairs, World Economic Forum and Singapore Exhibition and Convention Bureau.

INTERPOL World 2017
With a vision of fostering innovation for future security challenges, the second edition of INTERPOL World will take place at Suntec Singapore Convention and Exhibition Centre. More details can be found at www.interpol-world.com.

The three day INTERPOL World Congress is a knowledge exchange platform that engages in a multi-stakeholder approach to discuss about strategic challenges and solutions to future crime. The themes of the Congress are:

- **4 July 2017, Tuesday**
  Shedding light on the "Dark side"– Cyberspace and the future of policing; managing cyber threats to society from the "hidden" Internet.
- **5 July 2017, Wednesday**
  Prevention – Getting smarter, faster and more precise. Preparing policing strategies, approach and tactics for managing urban centers and global cities of the future.

- **6 July 2017, Thursday**
  Identity management and detection in a borderless world. Law enforcement, migration and border management in an age of globalization.

Each day, the Congress will take on a similar format as follows:

- **Segment 1 : INTERPOL World Dialogue**
  A moderated session with expert panelists to discuss and shade new light on ways to overcome emerging security and public safety issues affecting the international community.
- **Segment 2 : Strategic Perspective**
  Presentations from academia, researchers on in-depth analysis of the situation, trends, future trajectory and recommended policy and/or technological solutions.
- **Segment 3 : Operational Perspective and Case Studies**
  This session will involve solution providers or manufacturers to showcase futuristic R&D and/or new solutions to tackle the problem statement.
  Law enforcement agencies will also be invited to share their successful implementation of innovations and technologies in their field of work.

INTERPOL would like to invite you to share a case study on Los Angeles Police Department's adoption of a cloud-based crime prediction software (PredPol) and how application of analytical techniques, particularly statistical techniques have helped your department to identify promising targets for police intervention and ultimately help to prevent or solve crime. Your presentation should be about 20min and will be scheduled on **5 July 2017, Segment 3.**

If you accept this invitation:

- The event, INTERPOL World, will cover your air travel, accommodation expenses and land transfers during your stay in Singapore in July 2017. For your air travel to Singapore, we will arrange your ticket as per your organization's practice and regulation.
- Your speaker badge will allow you to access the INTERPOL World Exhibition, Congress, VIP lounges and all networking events during the duration of the event. The three day INTERPOL World Exhibition, taking place from 5-7 July 2017, will showcase solutions from 300 international manufacturers and solution providers.

Please feel free to contact my following colleagues should you need more information:

- Ms Margaret Samuel
  Directorate Secretary, Innovation Centre, INTERPOL Global Complex for Innovation
  m.samuel@interpol.int

We would be delighted to benefit from your recognized expertise and sincerely hope to meet you at INTERPOL World 2017. It is our strong opinion that your contribution to the Congress would be greatly beneficial for all participants.

We look forward to your reply by 28 February 2017.

Yours sincerely,

中谷昇

Noboru Nakatani
Executive Director
INTERPOL Global Complex for Innovation

IPAM 2017 RIPS-LAPD Project

Conversational Turn-Taking in Police Body-Worn Video

Industry Sponsor: Deputy Chief Sean Malinowski (LAPD Chief of Staff); Sgt. Dan Gomez, Sgt Rogelio Nunez, Ofcr Bill Coleman, Mr. Arnold Suzukamo (LAPD-IT Bureau).

Academic Mentor:

Academic Supervisors: Jeff Brantingham, UCLA Anthropology; Dr. Craig Uchida, Justice & Security Strategies

Introduction

Body-worn video (BWV) or on-body cameras provide a novel means to collect very fine-information about police-public interactions. The general use model requires officers to initiate recording of video whenever there is an encounter with a member of the public. During such interactions, BWV is recorded in real-time. Recording is terminated at the officer's discretion. BWV is not streamed or reviewed in real-time, but rather is uploaded to a secure cloud storage system at the end of an officer's shift.

BWV is designed to provide another line of evidence for the actions of individuals and the outcomes of interactions between police and members of the public. BWV is therefore evidence relevant to legal proceedings like any other form of evidence collected by police. In a limited number studies, BWV has been shown to reduce the likelihood that situations escalate to a point requiring use of force.

There are considerable challenges facing wide-spread use of BWV. Even small scale deployments are expected to lead to massive volumes of video data that will quickly outstrip the ability of law enforcement agencies to analyze. The resulting fallback position will be to review BWV footage only when it corresponds to adverse outcomes (e.g., use of force). Most video will go unused. Many of the potential benefits of BWV may therefore go unrealized.

The 2017 LAPD-RIPS Project

The 2017 RIPS-LAPD team will work to develop methods for the automatic discrimination and labeling of audio-video segments into the following categories: (1) the focal police officer speaking; (2) other actors speaking; and (3) overlapping speech involving the focal officer and others. The focal police officer is defined as the officer wearing the camera. The goal is not speech content recognition, or transcription. Rather we wish to identify when police officers exclusively are speaking relative to one or more other actors in a video scene and when the officer and others are trying to override one another with speech. Measures of conversational turn taking may then be computed. Conversational turn taking may provide evidence of when interactions are escalating or de-escalating without specific knowledge of the content of speech. Understanding when interactions escalate and de-escalate can be of tremendous value in helping to minimize the risk of adverse outcomes in police-public interactions.

The project will rely on a range of data types BWV metadata (e.g., time stamps), BWV audio, and the video images themselves. Computations may be done in Matlab, Mathematica, C, C++, R, Java, or other appropriate computational language.

Key Milestones:

1. Statistical assessment of LAPD BWV and other associated data.
2. Develop speech segmentation methods.
3. Measuring conversational turn taking.
4. Testing of efficacy of methods.
5. Present to LAPD.

References

Ariel, Barak, WilliamA Farrar, and Alex Sutherland. 2014. The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology*:1-27.

Kim, Samuel, Sree Harsha Yella, and Fabio Valente. 2012. "Automatic detection of conflict escalation in spoken conversations." Interspeech September, 1167-1170.

Merkurjev, Ekaterina, Justin Sunu, and Andrea L. Bertozzi. "Graph MBO method for multiclass segmentation of hyperspectral stand-off detection video." Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014.

Jason Xu

Department of Biomathematics
The University of California, Los Angeles

Phone: (520) 247-4646
Email: jqxu@ucla.edu

Education

Ph.D. Statistics, The University of Washington, August 2016.

Thesis: Likelihood-Based Inference for Partially Observed Multi-Type Markov Branching Processes

Committee: Vladimir Minin (advisor), Peter Guttorp, Jon Wakefield

B.S. Mathematics, *Summa Cum Laude*, The University of Arizona, June 2012.

Highest Honors, Budapest Semesters in Mathematics, Spring 2011.

Class of 2008 Valedictorian, Ironwood Ridge High School, Tucson, Arizona.

Selected Honors

Carol M. Newton Award, UCLA Department of Biomathematics, 2016

National Science Foundation Mathematical Sciences Postdoctoral Research Fellowship (NSF MSPRF), 2016

‡Z. W. Birnbaum Award, University of Washington, 2015

David P. Byar Young Investigator Competition Travel Award, American Statistical Association, 2015

National Defense Science and Engineering Graduate (NDSEG) Fellowship, 2012

NSF Graduate Research Fellowship Honorable Mention, 2012

‡College of Science Excellence in Undergraduate Research Award, University of Arizona, 2012

‡Department of Mathematics Outstanding Senior Award, University of Arizona, 2012

Churchill Scholarship Nominee, University of Arizona, 2012

Pillars of Excellence Award, University of Arizona, 2012

SIGMAA award for Research in Probability and Statistics, Joint Mathematics Meetings, 2011

Phi Beta Kappa Honorary Society, 2011-present

Flinn Foundation Scholarship, 2008

National Merit Scholar, 2008

‡ denotes university awards conferred to single recipient annually

Research and Teaching Positions

NSF Mathematical Sciences Postdoctoral Research Fellow, 2016—present.

Statistical analysis of partially observed interacting particle systems, mentored by Kenneth Lange.

Instructor, BIOMATH 210, UCLA (graduate): Optimization Methods, Fall Quarter 2016.

Research Assistant, Center for Inference and Dynamics of Infectious Diseases (CIDID) grant, August 2015—August 2016.

Efficient exact likelihood-based inference for fitting stochastic SIR models to massive disease data, advised by Jon Wakefield and Vladimir Minin

Instructor, STAT 341, UW (undergraduate): Intro to Probability and Mathematical Statistics II, Winter Quarter 2016.

NDSEG Fellow, 2012—2015.

Stochastic modeling and inference for continuous-time Markov branching processes and hidden Markov models with applications to epidemiology, biology, genomics, advised by Vladimir Minin.

Researcher, UCLA Institute of Pure and Applied Mathematics, June—August 2012.

Optimal pointing problems for satellite-to-region coverage, advised by Nam Lee in RIPS program. Novel minimax solution and software used internally by California-based nonprofit The Aerospace Corporation.

Research Assistant, VIGRE grant, University of Arizona Department of Mathematics, June 2011—May 2012.

Momentum-based and non-reversible Markov chain Monte Carlo methods to improve mixing by avoiding diffusive "backtracking" behavior, advised by Kevin Lin.

NSF REU, Claremont Colleges, June—August 2010.

Analysis and simulation of spatial point processes, advised by Mark Huber. Developed new analytical upper bound on the phase transition of continuous-time repulsive point processes and approximated bounds via Monte Carlo perfect simulation.

NSF REU, Willamette Valley Mathematics Research Consortium, June—August 2009.

Research in combinatorial game theory, advised by Hans Erik Nordstrom. Proved existence and solution of winning strategy to combinatorial NIM-type game *Odd Wins*.

Preprints

1. **Xu, J.**, Chi, E., Yang, M., Lange, K. An MM Algorithm for Split Feasibility Problems. [`arXiv Link`]

2. **Xu, J.**, Koelle, S., Wu, C., Guttorp, P., Dunbar, C. E., Abkowitz, J. L. and Minin, V. N. Statistical Inference in Partially Observed Stochastic Compartmental Models, with Application to Cell Lineage Tracking of *in vivo* Hematopoiesis. [`arXiv Link`]

3. Ho, L., **Xu, J.**, Crawford, F. W., Minin, V. N., and Suchard, M. A. Birth/birth-death Processes and their Computable Transition Probabilities with Statistical Applications. [`arXiv Link`]

4. Hardin, W., Li, R., **Xu, J.**, Shelton, A., Minin, V. N., Paredez, A. R. Flagella force generation drives *Giardia lamblia* cytokinesis. *Under review*

5. **Xu, J.**, Wakefield, J, Minin, V. N. Scalable Exact Posterior Inference for Stochastic SIR Models via a Two-type Branching Process Model. *In preparation*

6. Gustafson, A., Mohammed, I., Narayanan, H., and **Xu, J.** A Probably-Approximately-Correct Algorithm for Learning $C^{1,1}(\mathbb{R}^d)$ Functions from Noisy Samples. *In preparation*

7. **Xu, J.**, Guttorp, P., and Abkowitz, J. L. Simulation and Visualization of Hematopoiesis as a Stochastic Process. *In preparation*

Publications

8. Koelle, S., Espinoza, D., Wu, C., **Xu, J.**, Lu, R., Li, B., Donahue, R. E., and Dunbar, C. E. (2017) Quantitative Long-Term Stability of Hematopoietic Stem and Progenitor Cell Clonal Output in Transplanted Rhesus Macaques. *In press, Blood.*

9. **Xu, J.**, Koelle, S., Wu, C., Guttorp, P., Dunbar, C. E., Abkowitz, J. L. and Minin, V. N. (2016) Stochastic Modeling of Hematopoietic Stem and Progenitor Cell Barcoding Data from Rhesus Macaques Challenges the Classic Model of Hematopoiesis. *Blood.* [`bloodjournal.org Link`]

10. **Xu, J.** and Minin, V. N. (2015). Efficient Transition Probability Computation for Continuous-Time Branching Processes via Compressed Sensing. *Uncertainty in Artificial Intelligence (UAI).* [`auai.org Link`]

11. **Xu, J.**, Guttorp, P., Kato-Maeda, M. and Minin, V. N. (2015). Likelihood-Based Inference for Discretely Observed Birth-Death-Shift Processes, with Applications to Evolution of Mobile Genetic Elements. *Biometrics.* [`arXiv Link`] [`Biometrics Link`]

12. Foti, N.†, **Xu, J.**†, Laird, D., and Fox, E. B. (2014). Stochastic Variational Inference for Hidden Markov Models. *Neural Information Processing Systems (NIPS).* [`nips.cc Link`]

13. Huber, M. L., McCall, E., Rozenfeld, D. and **Xu, J.** (2012). Bounds on the Artificial Phase Transition for Perfect Simulation of Repulsive Point Processes. *Involve*, Vol. 5, No. 3, pp. 247-255. [`arXiv Link`]

† denotes joint first authorship.

Selected Presentations and Travel Awards

1. 12/2016, "Stochastic Modeling of Hematopoietic Stem and Progenitor Cell Barcoding Data from Rhesus Macaques Challenges the Classic Model of Hematopoiesis", American Society of Hematology (ASH) 2016 Annual Meeting, San Diego, CA.

   ASH Abstract Achievement Award

2. 11/2016, "Inference for Partially Observed Multi-type Branching Processes", Frontiers in Systems and Integrative Biology Seminar, UCLA, Los Angeles, CA.

3. 10/2016, "Stochastic Modeling and Inference with Multi-type Branching Processes", Invited Talk, Claremont McKenna College Applied Mathematics Colloquium, Claremont, CA.

4. 07/2016, "Statistical inference in partially observed stochastic compartmental models with application to cell lineage tracking of *in vivo* hematopoiesis", Oral Presentation, XXVIIIth International Biometric Conference, Victoria, BC.

5. 10/2015, "Statistically Reconstructing Dynamics of Blood Cell Differentiation from High Throughput Genetic Barcoding Experiments", Probabilistic Modeling in Genomics, Cold Spring Harbor, NY.

6. 08/2015, "Likelihood-Based Inference for Discretely Observed Birth-Death-Shift Processes, with Applications to Evolution of Mobile Genetic Elements", Topic Contributed Talk, Joint Statistical Meetings, Seattle, WA.

   ASA Biometrics Section Travel Award

7. 07/2015, "Efficient Transition Probability Computation for Continuous-Time Branching Processes via Compressed Sensing", Uncertainty in Artificial Intelligence, Amsterdam, Netherlands.

   UAI Travel Scholarship Award

8. 04/2015, "Continuous-time Branching Process Transition Probability Computation via Compressed Sensing", Invited Talk, Fred Hutchinson Cancer Research Center, Seattle, WA.

9. 12/2014, "Stochastic Variational Inference for Hidden Markov Models", Neural Information Processing Systems, Montreal, Quebec, Canada.

   CSSS Research Presentation and Training Grant

   NIPS Travel Award

10. 07/2014, "Fitting Multi-type Branching Models to Panel Data", Joint Statistical Meetings, Boston, MA.

    GSFEI Travel Award

11. 01/2013, "Regional Coverage Optimization with Steerable Satellite Sensors", Oral Presentation, Joint Mathematics Meetings, San Diego, CA.

12. 03/2012, “Odd Wins: Parity in Combinatorial Game Theory", Oral Presentation, SUnMaRC Conference, Tucson, AZ.
13. 01/2011, “Bounds on the Artificial Phase Transition for Perfect Simulation of Repulsive Point Processes", Joint Mathematics Meetings, New Orleans, LA.

    JMM Outstanding Poster Presentation award

Other Service

Non-profit Consultant, UW Statistical Consulting Center and STATCOM (Statistics in the Community), 2014—2016.

Tutor, UW Statistics Tutor and Study Center, 2013—2015.

Reviewer, Journal of Computational and Graphical Statistics, Bayesian Analysis, Neural Information Processing Systems

2017 RIPS-LAPD Project:

Conversational Turn-Taking in Police Body-Worn Video

Industry Sponsor: Deputy Chief Sean Malinowski (LAPD Chief of Staff); Sgt. Dan Gomez, Mr. Arnold Suzukamo (LAPD-IT Bureau).

Academic Mentor: Jason Xu (UCLA)

Academic Supervisors: Jeff Brantingham, UCLA Anthropology; Dr. Craig Uchida, Justice & Security Strategies

Introduction

Body-worn video (BWV) or on-body cameras provide a novel means to collect very fine-information about police-public interactions. The general use model requires officers to initiate recording of video whenever there is an encounter with a member of the public. During such interactions, BWV is recorded in real-time. Recording is terminated at the officer's discretion. BWV is not streamed or reviewed in real-time, but rather is uploaded to a secure cloud storage system at the end of an officer's shift.

BWV is designed to provide another line of evidence for the actions of individuals and the outcomes of interactions between police and members of the public. BWV is therefore evidence relevant to legal proceedings like any other form of evidence collected by police. In a limited number studies, BWV has been shown to reduce the likelihood that situations escalate to a point requiring use of force.

There are considerable challenges facing wide-spread use of BWV. Even small scale deployments are expected to lead to massive volumes of video data that will quickly outstrip the ability of law enforcement agencies to analyze. The resulting fallback position will be to review BWV footage only when it corresponds to adverse outcomes (e.g., use of force). Most video will go unused. Many of the potential benefits of BWV may therefore go unrealized.

The 2017 LAPD-RIPS Project

The 2017 RIPS-LAPD team will work to develop methods for the automatic discrimination and labeling of audio-video segments into the following categories: (1) the focal police officer speaking; (2) other actors speaking; and (3) overlapping speech involving the focal officer and others. The focal police officer is defined as the officer wearing the camera. The goal is not speech content recognition, or transcription. Rather we wish to identify when police officers exclusively are speaking relative to one or more other actors in a video scene and when the officer and others are trying to override one another with speech. Measures of conversational turn taking may then be computed. Conversational turn taking may provide evidence of when interactions are escalating or de-escalating without specific knowledge of the content of speech.

Understanding when interactions escalate and de-escalate can be of tremendous value in helping to minimize the risk of adverse outcomes in police-public interactions.

The project will rely on a range of data types BWV metadata (e.g., time stamps), BWV audio, and the video images themselves. Computations may be done in Matlab, Mathematica, C, C++, R, Java, or other appropriate computational language.

Key Milestones:

1. Statistical assessment of LAPD BWV and other associated data.
2. Develop speech segmentation methods.
3. Measuring conversational turn taking.
4. Testing of efficacy of methods.
5. Present to LAPD.

References

Ariel, Barak, WilliamA Farrar, and Alex Sutherland. 2014. The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology*:1-27.

Kim, Samuel, Sree Harsha Yella, and Fabio Valente. 2012. "Automatic detection of conflict escalation in spoken conversations." Interspeech September, 1167-1170.

Merkurjev, Ekaterina, Justin Sunu, and Andrea L. Bertozzi. "Graph MBO method for multiclass segmentation of hyperspectral stand-off detection video." Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014.

2017 RIPS-LAPD Project:

Conversational Turn-Taking in Police Body-Worn Video

Industry Sponsor: Deputy Chief Sean Malinowski (LAPD Chief of Staff); Sgt. Dan Gomez, Mr. Arnold Suzukamo (LAPD-IT Bureau).

Academic Mentor: Jason Xu (UCLA)

Academic Supervisors: Jeff Brantingham, UCLA Anthropology; Dr. Craig Uchida, Justice & Security Strategies

Introduction

Body-worn video (BWV) or on-body cameras provide a novel means to collect very fine-information about police-public interactions. The general use model requires officers to initiate recording of video whenever there is an encounter with a member of the public. During such interactions, BWV is recorded in real-time. Recording is terminated at the officer's discretion. BWV is not streamed or reviewed in real-time, but rather is uploaded to a secure cloud storage system at the end of an officer's shift.

BWV is designed to provide another line of evidence for the actions of individuals and the outcomes of interactions between police and members of the public. BWV is therefore evidence relevant to legal proceedings like any other form of evidence collected by police. In a limited number studies, BWV has been shown to reduce the likelihood that situations escalate to a point requiring use of force.

There are considerable challenges facing wide-spread use of BWV. Even small scale deployments are expected to lead to massive volumes of video data that will quickly outstrip the ability of law enforcement agencies to analyze. The resulting fallback position will be to review BWV footage only when it corresponds to adverse outcomes (e.g., use of force). Most video will go unused. Many of the potential benefits of BWV may therefore go unrealized.

The 2017 LAPD-RIPS Project

The 2017 RIPS-LAPD team will work to develop methods for the automatic discrimination and labeling of audio-video segments into the following categories: (1) the focal police officer speaking; (2) other actors speaking; and (3) overlapping speech involving the focal officer and others. The focal police officer is defined as the officer wearing the camera. The goal is not speech content recognition, or transcription. Rather we wish to identify when police officers exclusively are speaking relative to one or more other actors in a video scene and when the officer and others are trying to override one another with speech. Measures of conversational turn taking may then be computed. Conversational turn taking may provide evidence of when interactions are escalating or de-escalating without specific knowledge of the content of speech.

Understanding when interactions escalate and de-escalate can be of tremendous value in helping to minimize the risk of adverse outcomes in police-public interactions.

The project will rely on a range of data types BWV metadata (e.g., time stamps), BWV audio, and the video images themselves. Computations may be done in Matlab, Mathematica, C, C++, R, Java, or other appropriate computational language.

Key Milestones:

1. Statistical assessment of LAPD BWV and other associated data.
2. Develop speech segmentation methods.
3. Measuring conversational turn taking.
4. Testing of efficacy of methods.
5. Present to LAPD.

References

Ariel, Barak, WilliamA Farrar, and Alex Sutherland. 2014. The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology*:1-27.

Kim, Samuel, Sree Harsha Yella, and Fabio Valente. 2012. "Automatic detection of conflict escalation in spoken conversations." Interspeech September, 1167-1170.

Merkurjev, Ekaterina, Justin Sunu, and Andrea L. Bertozzi. "Graph MBO method for multiclass segmentation of hyperspectral stand-off detection video." Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014.

LAPD Trial Overview

Overview

The Los Angeles Police Department, alongside Justice & Security Strategies Inc., UCLA, and the Los Angeles Police Foundation, is conducting a three-year study of the use of body worn cameras by law enforcement. Under a grant from the National Institute of Justice and the United States Department of Justice, the four organizations are evaluating the use of cameras and examining the way in which footage from cameras is analyzed by third party machine learning vendors. LAPD plans to partner with the vendor(s) that demonstrates a solution that enhances LAPD's existing body-worn video workflows by producing high quality data.

Trial Requirements

EVALUATION CRITERIA

There are two key components that will be evaluated:

1. **Accuracy**: LAPD will provide a set of 400 videos. Each video is associated with one of three categories: vehicle pursuit, pedestrian stop, traffic stop. Vendors will be tested on their ability to train a model that correctly matches a video to its proper category.
2. **Performance**: Performance is a vendor's ability to go beyond identification and categorization of the video footage to add value to LAPD's body-worn camera workflows. There are multiple factors, including speed of analysis and reducing the overall amount of time it takes for the department to review footage.

EVALUATION PROCESS

LAPD will give each vendor 30 days to demonstrate Accuracy and Performance based on approximately 400 videos of uncategorized footage (this follows a 60-day period where vendors had access to 300 categorized videos to train their models). The vendors will prepare their results for a demo day, an evaluation that will take place at LAPD Headquarters.*

The following members will participate in vendor evaluation and selection: Sgt. Dan Gomez, LAPD; Arnold Suzukamo, LAPD; Sgt. Jose Macias, LAPD; Dr. Craig Uchida, JSS; Todd Maxwell, Department of Justice; Dr. Jeff Brantingham, UCLA.

**Following demo day, the evaluation team will provide each vendor with a summary of findings from the evaluation process.*

IMPORTANT DATES

Start date - *The date when vendors receive and download all videos in the training set.*

March 20, 2017

****Evaluation set date** - *Approximately 60 days after the Start date vendors will receive the test video set.*

~May 19, 2017

Presentation date - *Approximately 90 days after the Start date. Demo day presentations are approximately 60-90 minutes.*

~June 19, 2017

****To improve the validity of the test, we suggest you consider shortening the evaluation period from 30 days to 7 days. The results produced in a 7-day period will be more illustrative of the accuracy and performance of the machine learning models, and will give each vendor a better opportunity to demonstrate the speed of the automated analysis. In addition, 30 days is a long enough period to introduce methods that might alter the automated analysis, like manually generating results or tailoring the machine learning models to the test data set.*

DATA

All LAPD provided videos will be associated with one of the following three categories:

1. **Vehicle Pursuit**: An event involving one or more law enforcement officers in a patrol vehicle attempting to apprehend a suspect operating a motor vehicle while the suspect is attempting to avoid arrest (or detention) by using high speed driving or other evasive tactics. Vehicle Pursuit also involves events resulting of the aftermath of the pursuit such as a standoff, an area search or a foot chase.
2. **Traffic Stop**: An event involving one or more law enforcement officers in a patrol vehicle temporarily stopping a motor vehicle to investigate a possible crime or traffic violation, interacting with the driver of the stopped motor vehicle, and issuing a ticket and/or warning to the driver of the stopped motor vehicle.
3. **Pedestrian Stop**: A pedestrian stop is one or more law enforcement officers on foot, on a bicycle, or in a patrol vehicle temporarily stopping a pedestrian that is on foot or on a bicycle to investigate a possible crime or traffic violation, interacting with the pedestrian, and issuing a ticket and/or warning to the pedestrian that was stopped.

All vendors will receive the same two sets of un-redacted video footage recorded by LAPD officers wearing body-worn cameras in the field:

1. **Training data:** 300 videos (~100 videos per category)
2. **Evaluation data**: Uncategorized sample of 400 videos (this sample is limited to videos in the three categories)

The vendors are encouraged to spend time with the LAPD team to better understand the department's workflows and processes, and use this information and insight as an input to development. Beyond working directly with the LAPD team, the vendors are not to solicit feedback from the evaluation team.

The vendors may supplement the training and test data provided by LAPD with third party data. However, the vendors are prohibited from using third party solutions with the exception of: 1) a CJIS compliant workforce and 2) Microsoft Azure Services.

BUDGET

There are no financial obligations for LAPD and its partners during the trial. All vendors must cover the full cost of participation. The LAPD has no obligation to buy or purchase after the trial.

Selection

The evaluation team will select the vendor that demonstrates the best Accuracy and Performance at the end of the 90-day trial. Multiple vendors may be selected. For the vendor(s) that is selected, LAPD will work closely with the vendor to implement the technology solution within the department. Terms and pricing will be negotiated after the selection has been made.

UNIVERSITY OF CALIFORNIA, LOS ANGELES UCLA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

INSTITUTE FOR PURE AND APPLIED MATHEMATICS
BOX 957121
LOS ANGELES, CA 90095-7121

06/29/2017
Los Angeles Police Department

Dear Sgt. Dan Gomez, Sgt. Rogelio Nunez and Ofcr Bill Coleman,

Enclosed are two copies of the Statement of Work (SOW) and two copies of this cover letter. The SOW outlines our team's current understanding of the problem and addresses our planned approach to a solution.

Please show your approval of the SOW by signing both copies of the cover letter in the space provided on this page, or by indicating your proposed changes, and returning one copy of each (SOW and signed cover letter) to me by Friday, July 8. Otherwise, after that date, we will assume LAPD's tacit approval.

Sincerely,

Alistair Letcher
RIPS Project Manager

Institute for Pure and Applied Mathematics (IPAM)
Attn RIPS: Alistair Letcher

Enclosure: LAPD RIPS Work Statement

Cc: Susana Serna, RIPS Program Director
Stacey Beggs, IPAM Assistant Director

Accepted this________day of July 2017

By: ______________________________

Project Work Statement for RIPS 2017

Sponsor: Los Angeles Police Department
Title: Conversational Turn-Taking in Police Body-Worn Video

Student Participants: Collin Cademartori, Xi Chen, Alistair Letcher, Jelena Trisovic.

Academic Mentor: Jason Xu (UCLA).

Academic Supervisors: Jeff Brantingham, UCLA Anthropology; Dr. Craig Uchida, Justice & Security Strategies.

Industry Sponsor: Deputy Chief Sean Malinowski (LAPD Chief of Staff); Sgt. Dan Gomez, Mr. Arnold Suzukamo (LAPD-IT Bureau).

Date: 06/29/2017

Abstract

The Los Angeles Police Department (LAPD) is the law enforcement agency for the city of Los Angeles, California.

Body-worn video (BWV) or on-body cameras provide a novel means to collect very fine-information about police-public interactions. The general use model requires officers to initiate recording of video whenever there is an encounter with a member of the public. During such interactions, BWV is recorded in real-time. Recording is terminated at the officer's discretion. BWV is not streamed or reviewed in real-time, but rather is uploaded to a secure cloud storage system at the end of an officer's shift.

BWV is designed to provide another line of evidence for the actions of individuals and the outcomes of interactions between police and members of the public. BWV is therefore evidence relevant to legal proceedings like any other form of evidence collected by police. In a limited number studies, BWV has been shown to reduce the likelihood that situations escalate to a point requiring use of force.

There are considerable challenges facing widespread use of BWV. Even small scale deployments are expected to lead to massive volumes of video data that will quickly outstrip the ability of law enforcement agencies to analyze. The resulting fallback position will be to review BWV footage only when it corresponds to adverse outcomes (e.g., use of force). Most video will go unused. Many of the potential benefits of BWV may therefore go unrealized.

Project Description

The 2017 RIPS-LAPD team will work to develop methods for the automatic discrimination and labeling of BWV audio segments into the following categories: (1) the focal police officer speaking; (2) other actors speaking; and (3) overlapping speech involving the focal officer and others. The focal police officer is defined as the officer wearing the camera. The goal is not speech content recognition, or transcription. Rather we wish to identify when police officers exclusively are speaking relative to one or more other actors in a video scene and when the officer and others are trying to override one another with speech. Measures of conversational turn taking may then be computed. Conversational turn taking may provide evidence of when interactions are escalating or de-escalating without specific knowledge of the content of speech.

Understanding when interactions escalate and de-escalate can be of tremendous value in helping to minimize the risk of adverse outcomes in police-public interactions.

The project will rely on a range of data types: BWV metadata (e.g., time stamps), BWV audio and the video images themselves, although audio data is expected to take central stage.

Objectives

Our goal is to design a system that divides the body worn audio (BWA) data into three categories: 1) the police officer speaking 2) other actors speaking 3) overlapping speech involving the officer and others. This task falls under the broad class of problems called speaker diarization. Within this framework, we hope to design a statistical learning algorithm whose output will help measure turn taking statistics. These outputs can then be used to automatically recognize interpretable events of interest, such as whether the situation is escalating or de-escalating.

Performing this task using a supervised learning algorithm requires producing a ground truth of labeled data. This would involve labeling speakers after manually listening to a large portion of the BWA data (which we would like to avoid). This may turn out to be necessary, but we aim to develop an unsupervised method based on the following building blocks. We begin by data pre-processing, removing evident noise from the signal. The next step is feature selection and extraction, which entails selecting informative representations from the audio files as inputs to the learning algorithm. We will begin by selecting interpretable or standard features in audio analysis, perhaps then applying dimension reduction tools such as Principal Component Analysis to optimise our choice of variables. Once selected, these features will be used to distinguish human vs non-human noise (possibly including police radio, traffic, etc), distinguish between different speakers, and identify the police officer among them. This task answers the question of 'who spoke when', and is more informative than audio segmentation.

If time permits, we aim to use prosodic features (intonation, volume etc) and conversational features (duration of turns, individual speaking time, amount of overlap, number of participants etc) to automatically detect whether the conflict is escalating or not. We hope to conclude by evaluating the statistical accuracy/effectiveness of our approach.

Mathematical Background

We will explore a variety of mathematical techniques from the fields of digital signal processing and machine learning when developing and testing algorithms for conflict escalation detection. In particular, we may draw from the areas of digital filtering, Fourier analysis, and wavelet theory for the purpose of removing noise from the audio data and attempting to isolate the human voices. We will also explore various mathematical representations of the sound data to try to capture and emphasize features of the human voice which can be used to differentiate between speakers. Machine learning and statistics techniques will be applied to the resulting representation of the sound data. To discriminate speech vs nonspeech and different speakers, we will begin by applying classical unsupervised techniques such as clustering, at least as a first pass. If the need for supervised learning becomes apparent, we will label some BWA data manually and turn to techniques such as support vector machines or neural networks.

Time permitting, we will investigate Gaussian mixture models and dynamic models such as hidden Markov models, which make better use of temporal information and whose outputs may prove more fruitful in detecting overlapping speech.

Computational Background

We may use both MathWorks MATLAB and Python for writing software implementations of the techniques described above. MATLAB will be employed for exploring the efficacy and performance of various algorithms on the data and for creating visualizations of the results. We will also do some testing and try to implement software in Python, for purposes of execution speed and software portability.

Deliverables

From Team to Sponsor

- Mid-term presentation and progress report.
- Projects day (final) presentation.
- Final written report.
- Code used for the processing of BWV data and statistical analysis of output (if any), along with documentation. Any code sent to LAPD from our project will be in accordance with the Software Disclaimer for RIPS (see attached file).

From Sponsor to Team

- Dataset to be received by 06/26/2017.
- Timely response to email questions and conference calls if needed.
- Site visit.
- A helicopter ride at LAPD's convenience, if possible!

Schedule

June (Weeks 1-2)

- Receive and begin analysis of data front body worn cameras
- Research possible methods for extracting speaker data from audio

July (Weeks 3-6)

- Implement and test various methods for extracting information from audio data
- Deliver mid-term presentation on the 20th of July
- Possible site visit

August (Weeks 7-9)

- Finalize research, develop and document all software for delivery
- Possible site visit
- Deliver final written report and give projects day presentation on 17th of August

References

Ariel, Barak, William A Farrar, and Alex Sutherland. 2014. The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology*:1-27.

Kim, Samuel, Sree Harsha Yella, and Fabio Valente. 2012. "Automatic detection of conflict escalation in spoken conversations." *Interspeech September*, 1167-1170.

Merkurjev, Ekaterina, Justin Sunu, and Andrea L. Bertozzi. "Graph MBO method for multiclass segmentation of hyperspectral stand-off detection video." *Image Processing (ICIP)*, 2014 IEEE International Conference on. IEEE, 2014.

S. E. Tranter and D. A. Reynolds, "An overview of automatic speaker diarization systems," in *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 5, pp. 1557-1565, Sept. 2006.

X. Anguera, S. Bozonnet, N. Evans, C. Fredouille, G. Friedland and O. Vinyals, "Speaker Diarization: A Review of Recent Research," in *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 2, pp. 356-370, Feb. 2012.

IPAM Software Disclaimer for RIPS Sponsors

June 14, 2017

We want our RIPS sponsors to be aware of the nature of software developed by RIPS project teams. IPAM does not regard RIPS software as anything more than a prototype developed as a proof-of-concept only, and it is never developed for commercial use nor is it warranted by IPAM in any way. Here are some points to remember:

1. Software developed by a RIPS project team that appears to have been created wholly by a project team, may in fact contain proprietary codes borrowed from other sources; the sponsor must assume all risk for using such software.

2. IPAM makes every effort to discourage misuse of proprietary software by RIPS project participants; IPAM cannot be held responsible for such misuse.

3. As participants in an academic program, RIPS students will at times be permitted to use software that cannot be used by sponsors without a license.

4. Any restriction required by the sponsor on the use of special software, or platform needed to run the software, should be declared by the sponsor at the time of negotiating the project Work Statement. Otherwise the project team is free to choose software solutions as they see fit.

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2017 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______
Emily Loughran, Director of Licensing
Date: 6.2.17

By: Matt Haberland
Matt Habeland NSF REU Academic Mentor
Date: 6/19/2017

By: ______
Hao Li, NSF REU Academic Mentor
Date: 6/19/17

By: ______
Osman Akar, NSF REU Student
Date: June 19, 2017

By: ______
Adam Lemuel Dhillon, NSF REU Student
Date: 6/19/17

By: 陈泓林
Honglin Chen, NSF REU Student
Date: 06/19/2017

By: Alejandro E
Alexander Insuk Song, NSF REU Student
Date: June 19, 2017

<u>**CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT**</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2017 ("UCLA IPAM RIPS") (Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ______________________
Emily W. Loughran, Sr. Director of Licensing
Date: 6·2·17

By: ______________________
Jason Xu, RIPS Academic Mentor
Date: 6-19-17

By: ______________________
Collin Cademartori, RIPS Student
Date: 6-19-17

By: Xi Chen
Xi Chen, RIPS Student
Date: 6-19-17

By: Alistair Letcher
Alistair Letcher, RIPS Student
Date: 6-19-17

By: Jelena Trisovic
Jelena Trisovic, RIPS Student
Date: 6-19-17

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from video using machine learning algorithms. Researchers will examine video footage from LAPD to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

MAGGIE GOODRICH, Chief Information Officer
Commanding Officer
Information Technology Bureau

Date: 7-6-16

INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ______
Emily Loughran, Director of Licensing
Date: 6.16.16

By: ______ Giang Tran
First Last Name, RIPS Academic Mentor
Date: 06/21/2016

By: ______
First Last Name, RIPS Student
Date: 21 June, 2016
David Madras

By: ______
First Last Name, RIPS Student
Date: 21 June 2016
Stephanie Allen

By: ______
First Last Name, RIPS Student
Date: 21 June 2016
Greg Zanotti

By: ______
First Last Name, RIPS Student
Date: 21 June 2016
Ye Ye

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2017 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: _______________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: _______________________
Emily W. Loughran, Sr. Director of Licensing
Date: 6·2·17

By: _______________________
Jason Xu, RIPS Academic Mentor
Date: _______________________

By: _______________________
Collin Cademartori, RIPS Student
Date: _______________________

By: _______________________
Xi Chen, RIPS Student
Date: _______________________

By: _______________________
Alistair Letcher, RIPS Student
Date: _______________________

By: _______________________
Jalena Trisovic, RIPS Student
Date: _______________________

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2017 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: _______________________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: _______________________________
Emily Loughran, Director of Licensing
Date: 6.2.17

By: _______________________________
Matt Habeland NSF REU Academic Mentor
Date: _______________________________

By: _______________________________
Hao Li, NSF REU Academic Mentor
Date: _______________________________

By: _______________________________
Osman Akar, NSF REU Student
Date: _______________________________

By: _______________________________
Adam Lemuel Dhillon, NSF REU Student
Date: _______________________________

By: _______________________________
Honglin Chen, NSF REU Student
Date: _______________________________

By: _______________________________
Alexander Insuk Song, NSF REU Student
Date: _______________________________

By: ______________________________

Honglin Chen, NSF REU Student

Date: ____________________________

By: ______________________________

Tiankuang Zhou, NSF REU Student

Date: ____________________________

Does violence interruption work?

P. Jeffrey Brantingham[1*], Baichuan Yuan[2], Nick Sundback[3], Frederic Paik Schoenberg[4], Andrea L. Bertozzi[2], Joshua Gordon[4], Jorja Leap[3] Kristine Chan[5] Molly Kraus[5] Sean Malinowski[6] Denise Herz[5]

1 Department of Anthropology, University of California Los Angeles, Los Angeles, CA 90095, USA.
2 Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA.
3 Luskin School of Public Affairs, University of California Los Angeles, Los Angeles, CA 90095
4 Department of Statistics, University of California Los Angeles, Los Angeles, CA 90095, USA.
5 School of Criminal Justice and Criminalistics, California State University Los Angeles, Los Angeles, CA 90032, USA.
6 Los Angeles Police Department, Los Angeles, CA 90012, USA.

* branting@ucla.edu

Abstract

Retaliation propels gang violence. Spontaneous attacks resulting from chance encounters between rivals, or situational interactions that challenge gang territory or reputation can trigger cycles of tit-for-tat reprisals. Yet it has been difficult to determine if interventions that seek to reduce the likelihood of retaliation translate into lower rates of gang crime. Here we use multivariate spatial-temporal point process models to quantify the magnitude of retaliation arising from gang crimes given two different types of post-event interventions. The methods are well-suited to analysis of real-world interventions where there is imperfect separation between test conditions. Our analyses of quasi-experimental interventions in Los Angeles indicates that civilian Community Intervention Workers, tasked by the Gang Reduction Youth Development program, cut gang retaliations by 45.3%, independently of the effects of policing. Efforts to engage impacted families and control rumors undertaken in the immediate aftermath of gang violent crimes reduced the contagious spread of violence. These findings have important implication for the design, implementation and evaluation of gang violence prevention programs.

Introduction

Gang violence is distinctive for its ability to trigger clusters of retaliatory crimes [1, 2]. Interactions between gangs that threaten geographic territory or gang reputation can easily escalate to a shooting, while a shooting or homicide often demands retribution in kind [3, 4] ultimately driving a sequence of tit-for-tat reciprocal attacks [5–7]. Retaliation adds to the cumulative volume of violent crime and stronger retaliatory forces add more total crime [8–10]. It is important therefore to evaluate whether targeted interventions intended to reduce gang violence can do so by seeking to disrupt the process of retaliation [11]. The core premise of violence interruption is that street-level conflict mediation and rumor control in the immediate aftermath of a gang violent crime not only dissuades at-risk individuals from seeking quick payback [12], but also reduces tensions in the community that drive a broadly perceived need for street justice [4]. Prompt, community-focused deescalation is therefore thought to reduce the likelihood of retaliation. Although several studies have examined the aggregate impacts of comprehensive anti-violence programs, including as one part efforts to interrupt street violence [11, 13–15], these produced mixed results [16–18]. Previous research has not quantified the direct impacts of violence interruption on gang retaliation. Here we show that such impacts can be estimated directly from crime event data using a unique multivariate point process model.

0.0.1 Defining gangs and gang violence

Gangs and gang crime are inherently difficult to define [19, 20]. In Los Angeles, the focus of the empirical case here, the California Penal Code (CPC) provides a common and consistent starting point for law enforcement and violence interruption efforts. A "criminal street gang" is defined as an "ongoing organization, association, or group of three or more persons, whether formal or informal, having as one of its primary activities the commission of one or more of the criminal acts enumerated [in the CPC], having a common name or common identifying sign or symbol, and whose members individually or collectively engage in or have engaged in a pattern of criminal gang activity" (CPC §186.22(f)). A gang crime is therefore a "felony committed for the benefit of, at the direction of, or in association with any criminal street gang, with the specific intent to promote, further, or assist in any criminal conduct by gang members" (CPC §186.22(b)). The penal code definition may be criticized for being overly narrow, failing to recognize the diversity of social relationships and activity patterns attributable to gangs and, perhaps more importantly, the gradations of individual gang involvement or embeddedness [21]. Crime events might similarly reflect degrees of connection to gangs and gang activity. In general, a crime may be gang motivated, meaning that it was the result of activity in support of the gang, or simply gang affiliated, meaning that it was committed by one or more individuals embedded in a gang, but was otherwise unconnected to gang activity [7]. Whereas individual self-reporting may provide a suitable approach for characterizing the diversity ways in which individuals might be involved in gangs, there is no equivalent method for labeling gang crimes. We therefore rely on law enforcement and interventionist determinations of whether a crime is gang-related. These determinations are based on a range context-dependent criteria including whether the crime occurred in a known gang area, featured gang-involved or gang-affiliated individuals, appears connected to recent activity between nearby gangs, appears to be gang motivated, or is deemed to have the potential for retaliation. The more criteria associated with a given crime event, the more likely it is to be labeled a gang crime. Nevertheless, we recognize that gang member involvement in a crime may carry more weight than other criteria when labeling an event as gang-related, especially for law enforcement. We focus only on aggravated assaults and homicides, violent crimes where gang-motivated retaliation plays a central role [22, 23].

Retaliatory violence is similarly difficult to define. Qualitative criteria play a dominant role in determining whether a given crime is a retaliation, or has the potential to trigger a retaliation [4, 8]. The views of victims, witnesses, general knowledge about gang rivalries, and the long-term and recent history of gang interactions, including violence crimes, all weigh in making a such determinations. Whether or not a crime is deemed gang-related is also clearly connected to whether or not it is deemed to be a retaliation or have a potential for retaliation. While satisfying the desire to treat each crime as having a complex social and situational cause, such qualitative labels are difficult to verify on their own. We take an alternative approach that uses the statistical dependence between events as the basis for identifying retaliation (see below).

0.1 Modeling Gang Violence

Recent advances in statistical modeling of point processes reveal the dynamics of randomly occurring events characterized by self-excitation or contagion [24–27]. We consider Hawkes process models [28] with conditional intensity:

$$\lambda(t,x,y) = \mu(x,y) + \sum_{t_i<t} Kg(t-t_i, x-x_i, y-y_i), \tag{1}$$

where λ is the infinitesimal rate at which events accumulate at any point in space-time, given the entire history of the process. The model provides an intuitive characterization of gang violent crime events [29](Fig. 1A). It partitions the causes of crime into background processes $\mu(x,y)$, such as simmering gang rivalries, that generate crimes at a constant, but spatially variable rate, and contagion processes $Kg(t-t_i, x-x_i, y-y_i)$ that locally and briefly amplify the rate at which crime occurs.

The Hawkes model presents a novel probabilistic conceptualization of retaliation. To wit, background crimes are statistically independent of previous recorded events. Background crimes therefore cannot be retaliatory since retaliation by definition requires that a crime was committed in response to one or more prior crimes. Conversely, crimes triggered via self-excitation are statistically dependent on one or more prior events. Statistical dependence between crimes is robust criterion for identifying retaliatory crimes, though such crimes should not be equated exclusively with tit-for-tat violence between a dyad of rival gangs [8, 29]. Two crime events may be statistically dependent through a range of contagion-like processes occurring both within and between groups [1, 22, 30, 31]. For example, a gang may successfully attack one rival, which encourages them to quickly mount another attack against against a completely different rival [8]. Alternatively, conflict between individuals within a single gang may lead to retaliations that do not escape that local social network [32]. As discussed below, treatment interventions are mounted *as if* retaliation is the dominant contagion process. We suggest also that our probabilistic measure of retaliation is potentially more robust, and certainly more broadly applicable than methods that elicit a qualitative judgment from the street about whether one crime is a retaliation for some other crime. People are notoriously bad a evaluating the delinquency of peers [33, 34] and, in the absence of direct involvement, there is little reason to believe that they are any better at judging whether one event is a retaliation for another. Indeed, the central importance of rumor control in the civilian gang interventions discussed below proves as much. Even if such street-level judgments were reliable, the probabilistic measure suggested here is tractable for the large number of incidents where there is no qualitative information to be had.

We extend the above model to a multivariate framework [35] useful for describing real-world interventions at the scale of individual events. In such quasi-experimental field settings there is often imperfect separation between experimental conditions. To account for interactions we propose the conditional intensity:

$$\lambda_u(t,x,y) = \mu_u(x,y) + \sum_{t_i<t} K_{u_i u} g(t-t_i, x-x_i, y-y_i), \tag{2}$$

Here u_i is the type of event i where $u_i = 0$ represent an event assigned to a non-intervention control condition and $u_i = 1$ an event assigned to an intervention treatment condition. The model is easily modified to accommodate more than two interacting experimental conditions. The spatially inhomogeneous background rate in model [2] is now partitioned by condition u. The parameter $K_{u_i u}$ is now the expected number of retaliations of type u directly triggered by an event of type u_i. Thus we have four productivity parameters to estimate, $K_{u_i u} = k_{11}, k_{01}, k_{10}$, and k_{00}, representing the four possible interactions between treatment and control conditions (Fig. 1B). If treatment interventions are effective, then we expect fewer retaliations on average following events exposed to the treatment. Estimated productivity parameters should therefore satisfy $k_{11} < k_{01}$ and $k_{10} < k_{00}$. We estimate the model parametrically using an expectation-maximization (EM) algorithm [36] and evaluate the model using simulation, non-parametric models and Voronoi residuals [37, 38] (see Supplementary Information).

Experimental Setting

We analyzed gang violence in a unique quasi-experimental setting in Los Angeles. Each gang crime reported to the police prompts a field intervention either by police, or by the police and civilian gang intervention workers. The allocation of intervention conditions following each reported crime is approximately random, providing a measure of control over exogenous confounds (see below).

The control condition consists of field interventions by the Los Angeles Police Department (LAPD-only). LAPD follows standard law enforcement and investigative procedures when responding to a violent crime. Within the first minutes and hours after a reported gang crime, LAPD is focused on securing the scene, collecting physical evidence, and interviewing victims, suspects and other witnesses in an attempt to establish the physical and social circumstances of the crime. The LAPD is undoubtedly concerned with mitigating the potential for retaliation, their primary goal is to identify and hold accountable those responsible for the reported crime.

The treatment condition consists of field interventions by the police, but with additional notification of the City of Los Angeles Mayor's Office of Gang Reduction and Youth Development Incident Response (LAPD + GRYD IR). The LAPD field response is identical to that for the control condition. GRYD IR interventions take a different course. Upon receiving notification, regional program coordinators in the main GRYD Office task civilian Community Intervention Workers (CIWs) with responding to the scene. The protocol is for response to occur within 30 minutes of receiving notification. CIWs are not concerned with solving the current crime, but rather use non-law enforcement methods in an attempt to disrupt future retaliation. Many situationally-dependent actions, calibrated to micro-level processes of the gang [22], may be involved in disrupting retaliation. These include, but are not limited to engaging known actors who, based on the circumstances of the crime, may be at specific risk of seeking retribution, eliciting the support of community leaders, or facilitating access to services for a victim's family. However, rumor control by CIWs is perceived to be the most important of all the actions taken. Rumor control focuses on stopping the spread of misinformation about whether the crime was gang-related, who was responsible for the crime and who is likely to seek retribution. Importantly, CIWs vigorously protect their independence from the police, which they see as critical to their "license to operate" in the community. Tasking by the GRYD regional program coordinator creates an information buffer between CIWs and the LAPD, while the CIWs call to "never cross the yellow tape" provides a physical buffer. We therefore hypothesize that the additional efforts of civilian CIWs reduces the risk of retaliation above and beyond the effects of the police.

It is important to be clear that other anti-violence efforts form a common backdrop to both LAPD and GRYD IR interventions. In addition to incident response, the GRYD program provides comprehensive prevention and intervention services in gang-impacted communities. These seek to divert at-risk youth and help individuals leave the gang life, respectively. The police, city and independent non-profit organizations also support

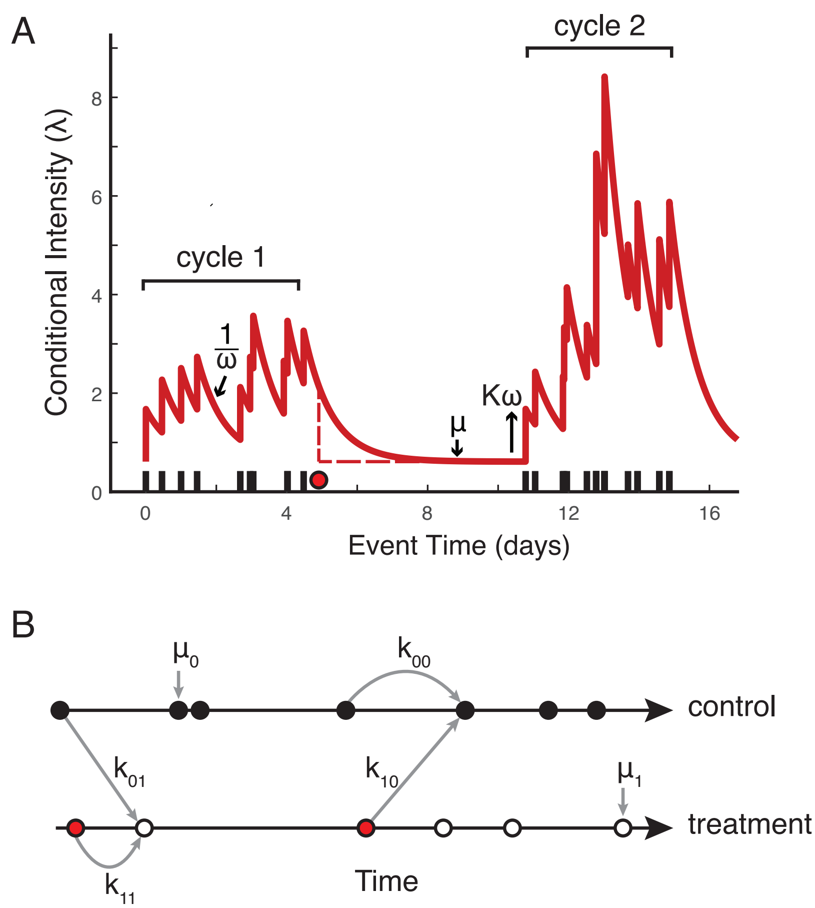


Fig 1. Self-exciting point process models capture the dynamics of gang violent crime events. (A). A temporal self-exciting point process model $\lambda(t) = \mu + \sum_{t_i<t} Kg(t - t_i)$ with exponential kernel $g(t) = \omega e^{-\omega t}$ fit to a sample of gang aggravated assaults and homicides in South Los Angeles from 2014-2015. Two cycles of gang violent crimes occur within a period of eighteen days. The conditional intensity λ reflects the instantaneous rate of gang crime. The background rate μ is the expected rate of gang crime in the absence of retaliation. A crime causes λ to jump by an amount $K\omega$, increasing the risk of retaliation. The risk of retaliation following a single crime decays exponentially with a rate ω and mean lifetime of $1/\omega$. If an intervention (red event) reduces the risk of retaliation, then we expect the conditional intensity to fall (dashed line) and future crimes to be less likely to occur than in the absence of intervention (B). Gang crimes assigned to two different experimental conditions are modeled as two interacting point processes. Non-retaliatory gang crimes assigned to each condition arise spontaneously at rate μ_j. Retaliations assigned to each experimental condition may be triggered through separate pathways. Parameter k_{ij} is an estimate of the average number of retaliations of type j triggered by a single crime of type i. Pathways k_{11} and k_{10} link previous treatment crimes to treatment and control retaliations, respectively. Pathways k_{00} and k_{01} link previous control crimes to control and treatment retaliations, respectively. If treatment interventions (red events) reduce the risk of gang retaliation, then we expect $k_{11} < k_{01}$ and $k_{10} < k_{00}$.

youth recreation programs, job training, tattoo removal, addiction counseling, and gun buyback events, all of which may contribute to a broader anti-violence effects. Custody, probation and parole are also at play in the background. These potential confounds are mediated by the random notification process that prompts interventions (see below).

In sum, the experimental units under consideration are the places and persons exposed to interventions in the immediate aftermath of a reported gang violent crime. However, the outcome of interest is the frequency and pattern of gang crimes subsequent to intervention exposure. We do not know the specific actions taken by police or civilian gang intervention workers during any one intervention. However, that intervention dosage is not directly measured does not preclude identifying treatment effects. Indeed, the setting here is very similar to most hot spot policing experiments where the experimental units are geographic locations and the local populations in those hot spots that receive police attention, while the measured outcome is crime volume or calls for service. The content of any one hot spot intervention—what was done by police and with whom police interacted—remains largely unobserved [39–41]. But the lack of such detailed information has not precluded measuring the effects of hot spot interventions. In the present case we know that a crime event prompts one of two types of interventions (LAPD-only or LAPD + GRYD IR) and we seek to measure the effects of those interventions without knowing exactly what happened in each intervention case.

1 Data

We analyze crimes occurring in an 87.2 km^2 area of South Los Angeles during 2014-2015 (Fig. 1 in SI Appendix). The area of interest is covered by ten GRYD Zones, formal areas designated by the city as eligible for GRYD services. The area represents 6.7% of the total land area of Los Angeles (~1,302 km2) and about 15.5% of the total population (~3.9 million), but accounted for 45.3% of serious gang crimes city-wide in 2014-2015. We limit our consideration to aggravated assaults and criminal homicides, crimes which entail a greater risk of retaliation compared to other crime types. Aggravated assaults involve the use of a deadly weapon, which in the context of gang violence is almost always a gun. In 2014-15, a total of 5,982 aggravated assaults and homicides were reported to the LAPD in the South Los Angeles GRYD Zones, including both gang (32.0%) and non-gang crimes (68.0%) (Table 1). GRYD IR was notified in 9.9% of all aggravated assaults, but 71.5% of all homicides. GRYD IR was notified more frequently when the crime was identified as gang related, including in 27.3% of gang aggravated assaults and 78.8% of gang homicides.

Table 1. Gang and non-gang crimes reported only to the LAPD and to the LAPD and GRYD IR in 2014-2015.

	LAPD-only			
	Gang N	Non-Gang N	TOTAL N	% Gang
Agg. Assault	1,249	3,918	5,167	24.2%
Homicide	41	29	70	58.6%
TOTAL	1,290	3,947	5,237	24.6%

	LAPD + GRYD IR				Grand Total	
	Gang N	Non-Gang N	TOTAL N	% Gang	TOTAL N	% Gang
Agg. Assault	470	99	569	82.6%	5,736	30.0%
Homicide	152	24	176	86.4%	246	78.5%
TOTAL	622	123	745	83.5%	5,982	32.0%

A key feature of the data facilitating our analysis is that the allocation of the two test conditions approximated a randomized experimental protocol. This *as if* random assignment arises naturally out of the crime reporting system. Upon receiving a report of a gang crime, LAPD's notification of GRYD IR proceeded as if a biased coin was flipped. If the crime was a gang aggravated assault, the coin was biased towards not notifying GRYD IR. If it was a gang homicide, it was biased towards notifying GRYD IR. To verify we performed runs tests ($Z = 0.609$, $p = 0.54$) and two-sample Kolmogorov–Smirnov (KS) tests (K-S $= 0.069$, $p = 0.20$ in time and K-S $= 0.054$, $p = 0.17$ in space), which found no discernible departures from the null hypothesis that the GRYD IR notifications were random, independent Bernoulli trials (SI Appendix). We hypothesize that the source of randomness in notification of GRYD IR is not intentional, but rather stems from inherent stochasticity in moment-to-moment demand on police time and attention. Sometimes this demand interferes with notification of GRYD IR, but the interference is random with respect to the event itself. As a result, the potential range of fixed and random factors that precipitated the crime are also randomized across intervention conditions. As a result, measured differences can be more confidently assigned to the effects of treatment rather than exogenous confounds.

The*as if* random assignment of conditions does not block interactions between crimes, however. As illustrated in Figure 1B, crimes randomly assigned to the control condition (LAPD-only) may trigger events subsequently assigned to the same (k_{00}), or the alternate intervention condition (k_{01}). Similarly, crimes randomly assigned to the treatment condition (LAPD + GRYD IR) may trigger events assigned to the same (k_{11}), or alternate intervention condition (k_{10}). These interactions between crimes mean that a stable unit treatment value assumption (SUTVA) is not met [42]. As a result, any observed treatment effect only holds for the empirically observed pattern of assignments between treatment and control conditions. If $\mathbf{d}$ is an indicator vector describing the observed sequence of assignments of N crimes between control and treatment conditions, then the event-level treatment effect may only be written as $\delta_i(\mathbf{d}) = y_i^1(\mathbf{d}) - y_i^0(\mathbf{d})$ [43], where y_i^0 is the effect for events assigned to control and y_i^1 the effect for events assigned to treatment. The treatment effect therefore might be different for each and every unique sequence of assignments $\mathbf{d}$. For example, the measured treatment effect for three events with an assignment pattern $\mathbf{d} = (1, 0, 1)$, might be different from an alternative random assignment $\mathbf{d} = (0, 1, 1)$ due to the interactions between events. Our approach is to estimate the interactions and evaluate whether these are in the direction expected if GRYD IR is able to disrupt retaliation. We also evaluate how treatment effects differ based on non-overlapping sequences of assignments.

Results

GRYD IR was deployed disproportionately in response to gang crimes, representing 83.5% of aggravated assaults and homicides for which it received notification (Table 1). The 16.5% of non-gang crimes reported to GRYD IR were likely considered gang-related at the point of initial notification, but were subsequently reclassified. Nevertheless, because GRYD IR does confront a mix of gang and non-gang crimes, we first tested whether GRYD IR had an impact against violent crimes in general. We fit the spatial-temporal multivariate model (Equation 2), with an exponential kernel for g (SI Appendix), to the full complement of 5,982 crimes (Fig. 2B). The estimate of parameter k_{11} indicates each aggravated assault or homicide exposed to the treatment on average triggered 0.1401 retaliations subsequently known to both LAPD and GRYD IR. Parameter k_{01} indicates each control crimes triggered on average 0.0526 retaliations. The estimated 62.5% higher rate of retaliation stemming from events exposed to the treatment condition is statistically significant ($p = 0.0092$) [44](SI Appendix). Pathway k_{10} shows that treatment crimes also triggered on average 0.2841 retaliations known only to the LAPD, which is of equivalent magnitude to pathway k_{00} with 0.2824 retaliations ($p = 0.486$). In practical terms, the fitted model suggests that every 100 LAPD + GRYD IR (treatment) aggravated assaults and homicides triggered on average 42.4 retaliatory violent crimes (k_{11} + k_{10}), compared to an average 33.5 retaliations triggered by LAPD-only (control) crimes ($k_{01} + k_{00}$). GRYD IR may reduce retaliation among the broader set of gang and non-gang violent crimes, but not to levels characteristic of the LAPD control case, with its greater mix of non-gang crimes and inherently lower risk of retaliation.

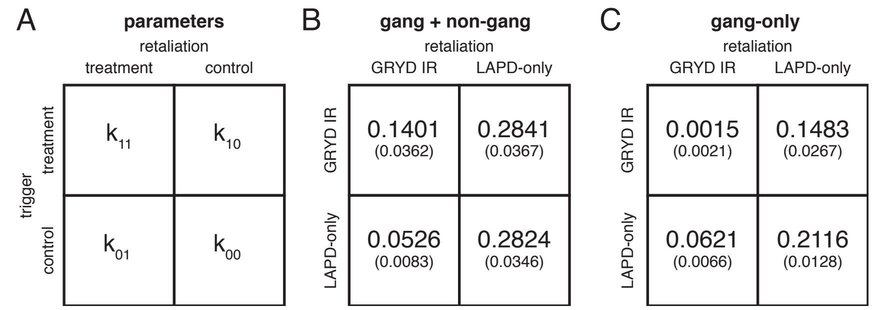


Fig 2. (A) Matrix representation of productivity K_{ij} with the corresponding triggering pathways noted. Matrix entries are the average number of retaliations assigned to an experimental condition i triggered by an event assigned to condition j. (B). The productivity for combined gang and non-gang aggravated assaults and homicides in South Los Angeles for 2014-15. (C). The productivity for gang-only aggravated assaults and homicides in South Los Angeles for 2014-15. The control condition includes violent crimes known to the LAPD (LAPD-only). The treatment condition includes crimes known to the LAPD that were also reported to GRYD IR (LAPD + GRYD IR). Standard errors of parameter estimates are shown in parentheses.

We therefore restricted analyses to gang aggravated assaults and homicides to ensure that test conditions were evaluated for events with similar potential for spawning retaliation. Against this set of crimes GRYD IR had a substantial impact (Fig. 2C). Pathway k_{11} triggered an average of 0.0015 retaliations for any one treatment gang crime. By contrast, pathway k_{01} triggered 0.0621 retaliations for any one control gang crime. This represents a 97.6% reduction in retaliation associated with GRYD IR notification ($p < 10^{-6}$). Pathway k_{10} triggered an average of 0.1483 retaliations for any one treatment gang crime. Pathway k_{00} triggered 0.2116 retaliations for any one control gang crime. The 29.9% reduction in retaliations is also significant ($p = 0.0163$). Every 100 LAPD + GRYD IR (treatment) gang crimes triggered an average of 15.0 retaliations ($k_{11} + k_{10}$) compared to 27.4 retaliations triggered by LAPD-only (control) crimes ($k_{01} + k_{00}$). Overall, the notification of GRYD IR was associated with a 45.3% decrease in retaliatory gang crimes.

Given the interactions between events, it is possible that the observed effects are unique to the exact sequence of observed assignments **d** between control and treatment conditions [43]. We evaluated this constraint by dividing the 2014-2015 data into eight non-overlapping blocks of 90 days each. The sequence of assignments in each non-overlapping block is independent and represents a different realized value of **d**. Table 2 shows that the parameter estimates for each of the triggering pathways do indeed differ in each non-overlapping block, consistent with the notion that treatment effects are dependent on the exact sequence of assignments **d**. However, in each block it is also the case that the direction of the treatment effect is as hypothesized. That is, the notification of GRYD IR results in fewer gang retaliations compared with the control case where only LAPD responds (i.e., $k_{11} < k_{01}$ and $k_{10} < k_{00}$) in each block. This encourages a more generous conclusion that GRYD IR has an impact across different possible sequences of interactions between crimes, though whether this holds under all possible assignments of **d** is impossible to say.

Table 2. Model estimation for gang crimes occurring in non-overlapping three month periods.

block	gang crimes (N)	k11	k01	k10	k00
1	162	1.01E-79	0.1194	0.1256	0.1815
2	211	0.0157	0.0617	0.1988	0.2246
3	231	0.0341	0.0698	0.1300	0.2610
4	245	6.46E-64	0.0655	0.1507	0.2059
5	259	7.48E-24	0.1034	0.1840	0.2040
6	255	2.33E-17	0.0610	0.2017	0.2222
7	315	0.0191	0.0962	0.2167	0.2867
8	212	0.0155	0.0362	0.1268	0.2626

To better understand the spatial dynamics of retaliatory gang violence we mapped the background intensity and self-excitation parts of the conditional intensity along with the distributions of background and retaliatory crimes determined via stochastic declustering [45] (Supplemental Information) (Fig. 3A-F). The background risk of gang violence is characterized by numerous compact, but widely distributed hot spots (Fig. 3A), consistent with the observation that the opportunity for violence and strength of gang rivalries is geographically variable [8, 46]. The risk of retaliation is concentrated in more continuous bands (Fig. 3D), bridging the discrete areas of background risk. Notably there is a prominent North-South corridor of retaliatory risk that maps to an area locally known as 'death alley' [47]. The patterns of risk influence the distribution of gang violent crimes (Fig. 3C and F). The density of background crimes forms five distinct hot spots (Fig. 3B) suggesting that background crimes are of local, neighborhood origin. The density of retaliatory crimes occupies two distinct hot spots (Fig. 3E), suggesting that retaliation spreads contagiously beyond immediate neighborhood contexts.

Stochastic declustering also allows us to evaluate differences in the frequency of retaliation by crime type across test conditions (Table 3). Background crimes make up 76.6% of all gang aggravated assaults and homicides for both test conditions combined. Retaliatory crimes are proportionally more common among events assigned to the LAPD-only control condition. This imbalance is pronounced for gang aggravated assaults (46.3% retaliation for LAPD-only vs. 10.3% for LAPD + GRYD IR), but particularly extreme for homicides (24.2% LAPD-only vs. 0.7% for LAPD + GRYD IR), especially considering the baseline bias towards notifying GRYD IR of most gang homicides.

Table 3. Number of retaliatory and background aggravated assaults and homicides in South Los Angeles in 2014-2015 separated by test condition.

	LAPD-only			
	Retaliation	Background	TOTAL	% Retaliation
Agg. Assault	395	854	1,249	46.3%
Homicide	8	33	41	24.2%
TOTAL	403	887	1,290	45.4%

We use the estimated treatment effects along with the results of stochastic declustering to compute numbers of prevented crimes (Table 3)(Supplementary Information). The sum $(k_{11} + k_{01})$ is the average number of retaliations known to both the LAPD and GRYD IR produced by any one event (Fig. 1B). Similarly, $(k_{10} + k_{00})$ is

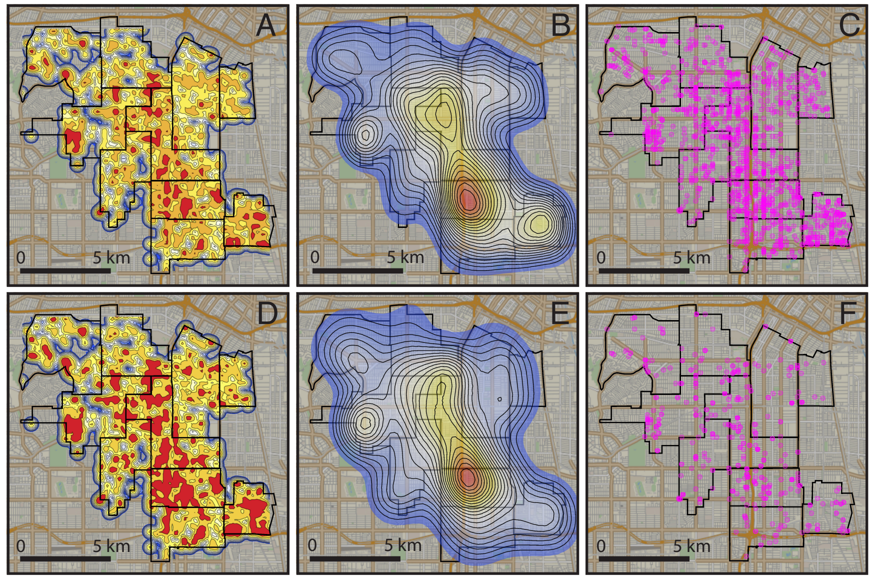


Fig 3. Stochastic declustering of gang crimes in South Los Angeles. (A). The log of spatial-temporal background intensity function μ for gang violent crimes mapped over space. (B) Contour plot of the density of background gang aggravated assaults and homicides determined by declustering. (C) Point locations of background gang aggravated assaults and homicides determined by declustering. (D) The log of spatial-temporal self-excitation of retaliation $\lambda - \mu$ mapped over space. (E) Contour plot of the density of retaliatory gang aggravated assaults and homicides determined by declustering. (F) Point locations of retaliatory gang aggravated assaults and homicides determined by declustering. Boundaries for the ten GRYD IR Zones in South Los Angeles are outlined in black.

	LAPD + GRYD IR			
	Retaliation	Background	TOTAL	% Retaliation
Agg. Assault	44	426	470	10.3%
Homicide	1	151	152	0.7%
TOTAL	45	577	622	7.8%

the average number of retaliations known only to the LAPD produced by any one event. Note that $(k_{11} + k_{01})$ and $(k_{10} + k_{00})$ are measured directly from data and therefore are observed outcomes. We now define two counterfactuals. Let $(k_{01} + k_{01})$ be the average number of retaliations that would have been triggered in the absence of GRYD IR for those events assigned to LAPD + GRYD IR. Let $(k_{00} + k_{00})$ be the average number of retaliations that would have been triggered in the absence of GRYD IR for those events assigned to LAPD-only. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. From stochastic declustering, the observed number of gang retaliations arising from pathways $(k_{11} + k_{01})$ is 45 and from $(k_{10} + k_{00})$ is 403 (Table 3). The counterfactual conditions suggest that retaliatory gang crimes would have been 48.8% and 15.0% higher in the absence of GRYD IR for observed pathways $(k_{11} + k_{01})$ and $(k_{10} + k_{00})$, respectively. GRYD IR prevented an estimated total 82.2 retaliatory gang crimes, of which 77.8 are expected to have been aggravated assaults and 4.4 are expected to have been homicides.

Discussion

The multivariate self-exciting point process model introduced here allows us to quantify how GRYD IR disrupts retaliatory gang crimes. Instead of determining the causal structure of each event, productivities $K_{u_i u}$ are averaged over the entire time series, which yields very robust and stable estimates. This is particularly important in real-world empirical settings where control and test conditions are often mixed by structural and practical conditions beyond the control of the observer. The multivariate modeling framework embraces this limitation and allows for interactions between conditions to proceed as part of the analysis. In the present case, we have the added benefit that the process by which GRYD IR was notified of gang violent crimes generated *as if random* assignment of events between the two test conditions. We were thus able to more completely disentangle the effects of intervention from other exogenous confounds that might lead some events to be preferentially treated only by the police and others by both the police and GRYD IR. We find that GRYD IR reduced gang retaliation by 45.3%, which corresponds to 82.2 fewer gang aggravated assaults and homicides over a two year period. Recent estimates place the overall cost of a single aggravated assault at $240,000 and a single homicide at $8.98 million [48]. Over the two-year period in 2014-15, the potential savings from gang aggravated assaults and homicides prevented by GRYD IR is estimated at $49.0 million in South Los Angeles alone.

These results have important implications for the design, implementation and evaluation of gang violence intervention programs. CIWs provide tangible benefits in reducing gang violence that complement, rather than compete with the effects of policing. The impact likely traces to the intimate connection CIWs share with the community they serve. CIWs are usually individuals who grew up in the community and may themselves have had some prior involvement with gangs. This high degree of social embedding and their prior experience gives CIWs a so-called *license to operate* in the community and from which they derive authority to mediate conflicts at the street level. Civilian intervention workers operating at the street level are not a new phenomenon [49]. But it has been difficult to establish their effectiveness independent of other violence-prevention efforts and the actions of police. The results presented here indicate that finding, cultivating and retaining such individuals would seem to be an essential part of designing gang violence interruption programs. It is also important to emphasize what CIWs do in the field. Of the many actions CIWs might take, rumor control in the immediate aftermath of a gang crime is deemed by them to the be the most important. Thus, the GRYD IR approach is unlike Boston's Operation Ceasefire or the "pulling levers" focused deterrence model [50, 51], which focuses on gathering intelligence about individual offending and communicating the costs of further offending to those individuals. It also differs from the Chicago's Ceasefire [11], where violence interruption appears to emphasize a more broad-based program of conflict mediation. In Los Angeles, the GRYD IR interventions are prompted by specific events and emphasize controlling information about those events as the main route to reducing retaliations.

Our approach is distinctive from studies that examine the effects of anti-violence programs before and after the onset of treatment, or between geographically non-overlapping control and treatment regions [11, 18]. Our study leverages noise in the gang crime reporting system, combined with a unique model for estimating the interactions between events, to detect and quantify the effects of violence interruption. The impact is significant amounting to a 45.3% reduction in retaliations or as many as 88.2 fewer aggravated assaults and homicides over a two-year period. However, these findings are also limited in several ways. First, they pertain only to the effects of violence interruption on the occurrence of retaliatory crimes. Stochastic declustering suggests that the majority of gang crimes are actually background events (Table 3).

Event-based violence interruption efforts to reduce retaliation may have little impact on the independent processes that generate these background events. Second, the observed impact of GRYD IR on retaliation may be lower than it might otherwise be because of noise in the notification system. In the ideal case, GRYD IR would receive notification for each and every gang crime occurring within an area of operation. In practice, however, notification is made only for a fraction of all events. This random process of assignment may be serendipitous for science, but it more critically represents missed opportunities to disrupt retaliation. Specifically, the results presented here suggest that eligible events that did not receive GRYD IR attention, likely produced more retaliations than would otherwise have been the case. Implementation of violence interruption programs modeled after GRYD IR would benefit from improvement in notification procedures that ensures more events are brought to the attention of CIWs.

Methods

Prior research on self-exciting point process models of crime suggest particular parametric equations for μ and g [40, 52]:

$$\mu_u(x,y) = \sum_{i=1}^{N} \frac{\beta_{u_i u}}{2\pi\sigma^2 T} \times \exp\left(-\frac{(x-x_i)^2+(y-y_i)^2}{2\sigma^2}\right).$$

$$g(x,y,t) = \omega \exp\left(-\omega t\right) \times \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

where β is a weights matrix for the degree to which events assigned to condition u_i contribute to the background rate for events assigned to condition u, T is the total time period represented by the sample of gang crimes, and σ is the spatial scale of influence for background events and retaliatory events. The kernel g is transformed to polar coordinates. The time scale of self-excitation is governed by ω (see Fig. 1A).

We use expectation maximization (EM) to estimate model [2] [36, 53]. The expectation step is used to compute initial probabilities that an event i causes event j via either the background rate μ or the self-exciting kernel g. These expectations are then fed to the maximization step where a new set of parameter values (for iteration $k+1$) are determined by maximizing the expected probability with respect to the observed data. This maximization is done for all parameters taking into consideration the condition u to which a gang crime is assigned. The algorithm alternates between expectation and maximization until there is no further change in the parameter values. The full EM algorithm is presented in the SI Appendix.

Once a model is estimated, Voronoi residuals provide a powerful technique for evaluating model performance [37] (Fig. 4). Voronoi residuals measure the differences between the modeled conditional intensity and the observed number of points within spatially adaptive Voronoi cells. Comparison to a color scale defined by a null Poisson model further helps interpret performance [38]. The null model corresponds to a case where all gang crimes are statistically independent background events. Voronoi residuals from the fitted intensity for the multivariate Hawkes model (Fig. 4B) has muted colors as compared to the null model (Fig. 4A), indicating improved performance. The fitted Hawkes process model performed well throughout the spatial window. The Poisson model underestimates the conditional intensity in areas such as the so-called 'death alley' where the Hawkes model identifies the greatest amount of retaliation (see Fig. 3D).

To address the question of whether our conclusion is dependent on the choice of parametric model, we constructed a model-independent multivariate self-exciting point process following the methods of [27, 35]. We estimated this point process on the same

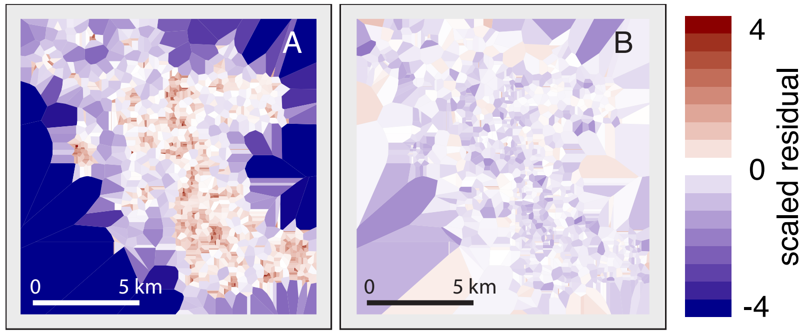


Fig 4. Voronoi residuals for spatial-temporal Poisson (A) and Hawkes process models (B). Color scaling identifies the spatial locations where a model is over or under estimated. Cells with blue shades (negative residuals) indicate an overestimation of the intensity, while cells with red shades (positive residuals) indicate an underestimation.

dataset non-parametrically. The results presented in the SI Appendix reinforce the conclusion that GRYD IR reduces gang retaliations.

We tested the parametric model on synthetic datasets corresponding to different intervention scenarios. We extended the methods in [54] to allow simulation of space-time multivariate Hawkes processes with different triggering matrices K. We chose triggering matrices corresponding to hypothetical positive, neutral and negative effects of GRYD IR intervention relative to the control. The results presented in the SI Appendix show that our methods are effectively able to recover ground-truth model parameters from simulated events and statistically distinguish between different intervention effects.

Acknowledgments

This research was made possible by the City of Los Angeles Mayor's Office of Gang Reduction and Youth Development (GRYD) and Los Angeles City contract #C-128086 ("Research Consulting for GRYD"). The authors thank GRYD Director Anne Tremblay, GRYD Office Staff and GRYD Intervention Incident Response Providers. We are also grateful to the Los Angeles Police Department for providing crime data. Thank you to California State University Los Angeles and University of California Los Angeles for their ongoing support. Any opinions, findings, conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of the GRYD Office.

References

1. Decker SH. Collective and normative features of gang violence. Justice Quarterly. 1996;13(2):243–264.

2. Klein MW, Maxson CL. Street gang patterns and policies. New York: Oxford University Press; 2006.

3. Hughes LA, Short JF. Disputes involving youth street gang members: Micro-social contexts. Criminology. 2005;43(1):43–76.

4. Jacobs BA, Wright R. Street justice: Retaliation in the criminal underworld. Cambridge University Press; 2006.

5. Bjerregaard B, Lizotte AJ. Gun ownership and gang membership. The Journal of Criminal Law and Criminology (1973-). 1995;86(1):37–58.
6. Howell JC. Gangs in America's communities. Sage; 2011.
7. Rosenfeld R, Bray TM, Egley A. Facilitating violence: A comparison of gang-motivated, gang-affiliated, and nongang youth homicides. Journal of Quantitative Criminology. 1999;15(4):495–516.
8. Papachristos AV. Murder by structure: Dominance relations and the social structure of gang homicide 1. American Journal of Sociology. 2009;115(1):74–128.
9. Stewart EA, Simons RL. Race, code of the street, and violent delinquency: A multilevel investigation of neighborhood street culture and individual norms of violence. Criminology. 2010;48(2):569–605.
10. Zeoli AM, Grady S, Pizarro JM, Melde C. Modeling the Movement of Homicide by Type to Inform Public Health Prevention Efforts. American Journal of Public Health. 2015;105(10):2035–2041.
11. Skogan WG, Hartnett SM, Bump N, Dubois J. Evaluation of CeaseFire-Chicago. Washington, D.C.: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice; 2009.
12. Dymnicki AB, Henry D, Quintana E, Wisnieski E, Kane C. Outreach Workers' Perceptions of Positive and Negative Critical Incidents: Characteristics Associated with Successful and Unsuccessful Vioence Interruption. Journal of Community Psychology. 2013;41(2):200–217. doi:10.1002/jcop.21523.
13. Webster DW, Whitehill JM, Vernick JS, Curriero FC. Effects of Baltimore's Safe Streets Program on gun violence: A replication of Chicago's CeaseFire Program. Journal of Urban Health. 2013;90(1):27–40.
14. Whitehill JM, Webster DW, Frattaroli S, Parker EM. Interrupting violence: how the CeaseFire Program prevents imminent gun violence through conflict mediation. Journal of Urban Health. 2014;91(1):84–95.
15. Wilson JM, Chermak S. Community-driven violence reduction programs. Criminology & Public Policy. 2011;10(4):993–1027.
16. Butts JA, Roman CG, Bostwick L, Porter JR. Cure violence: a public health model to reduce gun violence. Annual review of public health. 2015;36:39–53.
17. Klein MW. Comprehensive gang and violence reduction programs. Criminology & Public Policy. 2011;10(4):1037–1044.
18. Papachristos AV. Too big to fail. Criminology & Public Policy. 2011;10(4):1053–1061.
19. Ball RA, Curry GD. The Logic of Definition in Criminology: Purposes and Methods for Defining "Gangs". Criminology. 1995;33(2):225–245.
20. Curry GD. In: Decker SH, Pyrooz DC, editors. The Logic of Defining Gangs Revisited. John Wiley & Sons, Inc; 2015. p. 7–27.
21. Decker SH, Pyrooz DC, Sweeten G, Moule RK. Validating Self-Nomination in Gang Research: Assessing Differences in Gang Embeddedness Across Non-, Current, and Former Gang Members. Journal of Quantitative Criminology. 2014;30(4):577–598.

22. McGloin JM, Collins ME. In: Decker SH, Pyrooz DC, editors. Micro-Level Processes of the Gang. John Wiley & Sons, Inc; 2015. p. 276–293.

23. Sierra-Arevalo M, Papachristos AV. In: Decker SH, Pyrooz DC, editors. Social Network Analysis and Gangs. John Wiley & Sons, Inc; 2015. p. 157–177.

24. Cressie N, Wikle CK. Statistics for spatio-temporal data. John Wiley & Sons; 2015.

25. Delay D, Vere-Jones D. An introduction to the theory of point processes. Vol. I; 2003.

26. Diggle PJ. Statistical analysis of spatial and spatio-temporal point patterns. CRC Press; 2013.

27. Marsan D, Lengline O. Extending earthquakes' reach through cascading. Science. 2008;319(5866):1076–1079.

28. Hawkes AG. Spectra of some self-exciting and mutually exciting point processes. Biometrika. 1971; p. 83–90.

29. Short M, Mohler G, Brantingham PJ, Tita G. Gang Rivalry Dynamics Via Coupled Point Process Networks. Discrete & Continuous Dynamical Systems-Series B. 2014;19(5).

30. Green B, Horel T, Papachristos AV. Modeling contagion through social networks to explain and predict gunshot violence in Chicago, 2006 to 2014. JAMA internal medicine. 2017;.

31. Wiley SA, Levy MZ, Branas CC. In: Letzter G, Lauter K, Chambers E, Flournoy N, Grigsby JE, Martin C, et al., editors. The Impact of Violence Interruption on the Diffusion of Violence: A Mathematical Modeling Approach. Cham: Springer International Publishing; 2016. p. 225–249. Available from: `http://dx.doi.org/10.1007/978-3-319-34139-2_10`.

32. Decker SH, Curry GD. Gangs, gang homicides, and gang loyalty: Organized crimes or disorganized criminals. Journal of Criminal Justice. 2002;30(4):343–352.

33. John H Boman I, Young JTN, Baldwin JM, Meldrum RC. Specifying the Sources of Misperceptions of Peer Deviance. Criminal Justice and Behavior. 2014;41(1):91–113.

34. Young JTN, Barnes JC, Meldrum RC, Weerman FM. Assessing and Explaining Misperceptions of Peer Delinquency. Criminology. 2011;49(2):599–630. doi:10.1111/j.1745-9125.2011.00232.x.

35. Fox EW, Schoenberg FP, Gordon JS. Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric Hawkes point process models of earthquake occurrences. The Annals of Applied Statistics. 2016;10(3):1725–1756.

36. Veen A, Schoenberg FP. Estimation of space–time branching process models in seismology using an em–type algorithm. Journal of the American Statistical Association. 2008;103(482):614–624.

37. Bray A, Wong K, Barr CD, Schoenberg FP, et al. Voronoi residual analysis of spatial point process models with applications to California earthquake forecasts. The Annals of Applied Statistics. 2014;8(4):2247–2267.

38. Gordon JS, Clements RA, Schoenberg FP, Schorlemmer D. Voronoi residuals and other residual analyses applied to CSEP earthquake forecasts. Spatial Statistics. 2015;14:133–150.

39. Braga AA, Papachristos AV, Hureau DM. The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-Analysis. Justice Quarterly. 2014;31(4):633–663.

40. Mohler G, Short MB, Malinowski S, Johnson M, Tita GE, Bertozzi AL, et al. Randomized Controlled Field Trials of Predictive Policing. Journal of the American Statistical Association. 2015;110(512):1399–1411.

41. Sherman LW, Weisburd D. General Deterrent Effects of Police Patrol in Crime "Hot Spots": A Randomized, Controlled Trial. Justice Quarterly. 1995;12:625–648.

42. Rubin DB. Comment. Journal of the American Statistical Association. 1980;75(371):591–593.

43. Morgan SL, Winship C. Counterfactuals and causal inference. 2nd ed. Cambridge, England: Cambridge University Press; 2015.

44. Ogata Y. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. Annals of the Institute of Statistical Mathematics. 1978;30(1):243–261. doi:10.1007/bf02480216.

45. Zhuang J, Ogata Y, Vere-Jones D. Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Association. 2002;97(458):369–380.

46. Tita G, Ridgeway G. The impact of gang formation on local patterns of crime. Journal of Research in Crime and Delinquency. 2007;44(2):208–237.

47. Santa Cruz N, Schwenke K. type [; 2014].

48. McCollister KE, French MT, Fang H. The cost of crime to society: New crime-specific estimates for policy and program evaluation. Drug and alcohol dependence. 2010;108(1):98–109.

49. Klein MW. Street Gangs and Street Workers. Englewood Cliffs: Prentice Hall; 1971.

50. Braga AA, Kennedy DM, Waring EJ, Piehl AM. Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire. Journal of research in crime and delinquency. 2001;38(3):195–225.

51. Braga AA, Weisburd DL. The Effects of 'Pulling Levers' Focused Deterrence Strategies on Crime. Campbell Systematic Reviews. 2012;6.

52. Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE. Self-exciting point process modeling of crime. Journal of the American Statistical Association. 2011;106(493):100–108.

53. Mohler G. Marked point process hotspot maps for homicide and gun crime prediction in Chicago. International Journal of Forecasting. 2014;30(3):491–497.

54. Lewis PAW, Shedler GS. Simulation of nonhomogeneous poisson processes by thinning. Naval Research Logistics Quarterly. 1979;26(3):403–413. doi:10.1002/nav.3800260304.

Does Violence Interruption Work?

P. Jeffrey Brantingham[a,1], Baichuan Yuan[b], Nick Sundback[c], Frederick Paik Schoenberg[d], Andrea L. Bertozzi[b], Joshua Gordon[d], Jorja Leap[c], Kristine Chan[e], Molly Kraus[e], Sean Malinowski[f], and Denise Herz[e]

[a]Department of Anthropology, University of California Los Angeles, Los Angeles, CA 90095, USA.; [b]Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA.; [c]Luskin School of Public Affairs, University of California Los Angeles, Los Angeles, CA 90095, USA.; [d]Department of Statistics, University of California Los Angeles, Los Angeles, CA 90095, USA.; [d]Department of Statistics, University of California Los Angeles, Los Angeles, CA 90095, USA.; [e]School of Criminal Justice and Criminalistics, California State University Los Angeles, Los Angeles, CA 90032, USA.; [f]Los Angeles Police Department, Los Angeles, CA 90012, USA.

This manuscript was compiled on June 7, 2017

Retaliation propels gang violence. Spontaneous attacks resulting from chance encounters between rivals, or situational interactions that challenge gang territory or reputation can trigger cycles of tit-for-tat reprisals. Yet it has been difficult to determine if interventions that seek to reduce the likelihood of retaliation translate into lower rates of gang crime. Here we use multivariate spatial-temporal point process models to quantify the magnitude of retaliation arising from gang crimes exposed to different types of interventions on a per event basis. The methods are well-suited to analysis of real-world event-based interventions where there is imperfect separation between test conditions. Our analyses of quasi-experimental interventions in Los Angeles indicates that civilian Community Intervention Workers, tasked by the Gang Reduction Youth Development program, cut gang retaliations by 45.3%, independently of the effects of policing. Efforts to engage impacted families and control rumors reduced the contagious spread of violence when undertaken in the immediate aftermath of gang violent crimes. These findings have important implication for the design, implementation and evaluation of gang violence prevention programs.

crime | violence | criminal street gangs | self-exciting point process | quasi-experiment

Gang violence is distinctive for its ability to trigger clusters of retaliatory crimes (1, 2). Challenges between gangs that threaten geographic territory or gang reputation can easily escalate to a shooting, while a shooting or homicide often demands retribution in kind (3, 4) ultimately driving a sequence of tit-for-tat reciprocal attacks (5–7). Retaliation adds to the cumulative volume of violent crime and stronger retaliatory forces add more total crime (8, 9). It is important therefore to evaluate whether targeted interventions intended to reduce gang violence can do so by disrupting the process of retaliation. The core premise of violence interruption is that street-level conflict mediation and rumor control in the immediate aftermath of a gang violent crime not only dissuades known at-risk individuals from seeking quick payback (10), but also reduces tensions in the community that drive a broadly perceived need for street justice. Prompt, community-focused deescalation is therefore thought to reduce the likelihood of retaliation. Although several studies have examined the aggregate impacts of anti-violence programs, including efforts to interrupt street violence (11–14), these produced mixed results (15–17). Previous research has not quantified direct impacts of interventions on gang retaliation. Here we show that such impacts can be estimated directly from crime event data using a unique multivariate point process model.

Models

Recent advances in statistical modeling of point processes reveal the dynamics of randomly occurring events characterized by self-excitation or contagion (18–21). We consider Hawkes process models (22) with conditional intensity:

$$\lambda(t, x, y) = \mu(x, y) + \sum_{t_i<t} Kg(t - t_i, x - x_i, y - y_i), \quad [1]$$

where λ is the infinitesimal rate at which events accumulate at any point in space-time, given the entire history of the process. The model provides an intuitive characterization of gang violence (23)(Fig. 1A). It partitions the causes of crime into background processes $\mu(x, y)$, such as simmering gang rivalries, that generate crimes at a constant, but spatially variable rate, and contagion processes $Kg(t - t_i, x - x_i, y - y_i)$ that locally and briefly amplify the rate at which crime occurs. Background crimes are statistically independent of previous recorded events and therefore by definition cannot be retaliatory. Crimes triggered by self-excitation or contagion are statistically dependent on one or more prior events. Retaliation is a key contagion process creating statistical dependence between crimes (8, 24, 25).

We extend the above model to a multivariate framework (26) useful for describing real-world interventions at the scale of individual events. In such quasi-experimental field settings there is often imperfect separation between experimental conditions. To account for interactions we propose the conditional intensity:

$$\lambda_u(t, x, y) = \mu_u(x, y) + \sum_{t_i<t} K_{u_i u} g(t - t_i, x - x_i, y - y_i), \quad [2]$$

Significance Statement

Gang violence is a major challenge for many large and mid-sized US cities. In some cities, gang homicides can account for more than half of the total homicides in any year. One approach to reducing gang violence is to task civilian community interventions workers to mediate conflict and control rumors following a reported gang crime. The premise is that efforts to quickly deescalate community tensions disrupts the process of retaliation. We use a unique multivariate spatial-temporal point process model fitted to gang-violent crimes in South Los Angeles, CA, to show that violence interruption cuts the incidence of violent gang retaliation nearly in half.

P.J.B. and B.Y. designed the research; P.J.B., B.Y., F.P.S., N.S., A.L.B., J.L., K.C., M.K., S.M. and D.H. performed research; P.J.B., B.Y., N.S., J.G., K.C., and M.K. analyzed data; P.J.B., B.Y., F.P.S., and A.L.B. wrote the paper.

[1]To whom correspondence should be addressed. E-mail: brantingucla.edu

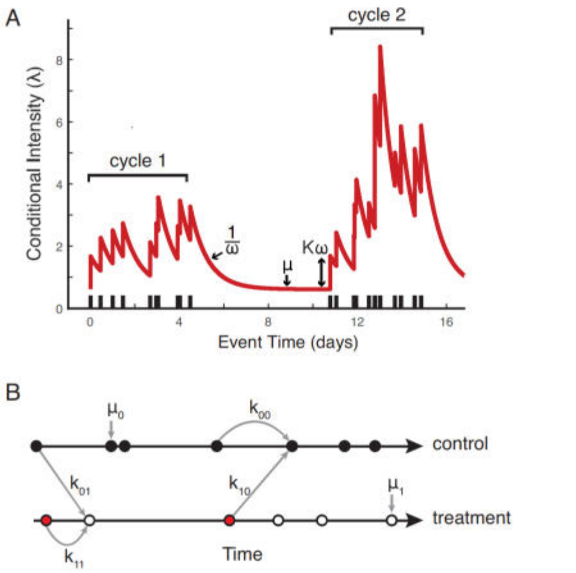


Fig. 1. Self-exciting point process models capture the dynamics of gang violence. (A). A temporal self-exciting point process model $\lambda(t) = \mu + \sum_{t_i<t} Kg(t - t_i)$ with exponential kernel $g(t) = \omega e^{-\omega t}$ fit to a sample of gang aggravated assaults and homicides in South Los Angeles from 2014-2015. Two cycles of gang violence occur within a period of eighteen days. The conditional intensity λ reflects the instantaneous rate of gang crime. The background rate μ is the expected rate of gang crime in the absence of retaliation. A crime causes λ to jump by an amount $K\omega$, increasing the risk of retaliation. The risk of retaliation following a single crime decays exponentially with a rate ω and mean lifetime of $1/\omega$. (B). Gang crimes assigned to two different experimental conditions are modeled as two interacting point processes. Non-retaliatory gang crimes assigned to each condition arise spontaneously at rate μ_j. Retaliations assigned to each experimental condition may be triggered through separate pathways. Parameter k_{ij} is an estimate of the average number of retaliations of type j triggered by a single crime of type i. Pathways k_{11} and k_{10} link previous treatment crimes to treatment and control retaliations, respectively. Pathways k_{00} and k_{01} link previous control crimes to control and treatment retaliations, respectively. If treatment interventions (red events) reduce the risk of gang retaliation, then we expect $k_{11} < k_{01}$ and $k_{10} < k_{00}$.

Here u_i is the type of event i where $u_i = 0$ represent an event assigned to a non-intervention control condition and $u_i = 1$ an event assigned to an intervention treatment condition. The model is easily modified to accommodate more than two interacting experimental conditions. The spatially inhomogeneous background rate in model [2] is now partitioned by condition u. The parameter $K_{u_i u}$ is now the expected number of retaliations of type u directly triggered by an event of type u_i. Thus we have four productivity parameters to estimate, $K_{u_i u} = k_{11}, k_{01}, k_{10}$, and k_{00}, representing the four possible interactions between treatment and control conditions (Fig. 1B). If treatment interventions are effective, then we expect fewer retaliations on average following events exposed to the treatment. Estimated productivity parameters should therefore satisfy $k_{11} < k_{01}$ and $k_{10} < k_{00}$. We estimate the model parametrically using an expectation-maximization (EM) algorithm (27) and evaluate the model using simulation, non-parametric models and Voronoi residuals (28, 29) (See Materials and Methods and SI Appendix).

Experimental Setting and Data

We analyzed both non-gang and gang violence in a unique quasi-experimental setting in Los Angeles where there is approximately random assignment of crimes between two different, but interacting intervention conditions (see below). The control condition consists of violent crimes reported to the Los Angeles Police Department (LAPD-only). LAPD follows standard enforcement and investigative procedures for all reported gang crimes. The treatment condition consists of violent crimes reported to the LAPD, but with additional notification of the City Los Angeles Mayor's Office of Gang Reduction and Youth Development Incident Response (LAPD + GRYD IR). Upon receiving notification, GRYD IR tasks civilian Community Intervention Workers (CIWs) with disrupting retaliation through community-based crisis response and rumor control. We hypothesize that the additional effort of civilian CIWs reduces the risk of retaliation above and beyond the effects of police.

We focus on crimes occurring in an 87.2 km^2 area of South Los Angeles during 2014-2015 (Fig. S1). The ten GRYD IR Zones in South Los Angeles represent 6.7% of the total land area of Los Angeles (~1,302 km2) and about 15.5% of the total population (~3.9 million), but accounted for 45.3% of serious gang crimes city-wide in 2014-2015. We limit our consideration to aggravated assaults and criminal homicides, crimes which entail a greater risk of retaliation compared to other crime types. Aggravated assaults are defined by the use of a deadly weapon, which in the context of gang violence is almost always a gun. In 2014-15, a total of 5,982 aggravated assaults and homicides were reported to the LAPD in the South Los Angeles GRYD Zones, including gang (32.0%) and non-gang crimes (68.0%) (Table 1). GRYD IR was notified in 9.9% of all aggravated assaults, but 71.5% of all homicides. GRYD IR was notified more frequently when the crime was identified as gang related, including in 27.3% of gang aggravated assaults and 78.8% of gang homicides.

A key feature facilitating our analysis is that crimes assigned to each of the two test conditions approximated a randomized experimental protocol. The random assignment arises naturally out of the crime reporting system. Upon LAPD receiving a report of a gang crime, notification of GRYD IR proceeded as if a biased coin was flipped. If the crime was a gang aggravated assault, the coin was biased towards not notifying GRYD IR. If it was a gang homicide, it was biased towards notifying GRYD IR. To verify we performed runs tests ($Z = 0.609$, $p = 0.54$) and two-sample Kolmogorov–Smirnov (KS) tests (K-S $= 0.069$, $p = 0.20$ in time and K-S $= 0.054$, $p = 0.17$ in space), which found no discernible departures from the null hypothesis that the GRYD IR notifications were random, independent draws from the same population of crimes (SI Appendix). This fact helps control for potential confounding between treatment effects and the process of notification.

Results

GRYD IR was deployed disproportionately in response to gang crimes, representing 83.5% of aggravated assaults and homicides for which it received notification (Table 1). The 16.5% of non-gang crimes reported to GRYD IR were likely considered gang-related at the point of initial notification, but were subsequently reclassified. Because GRYD IR confronts a mix of gang and non-gang crimes, we first tested whether GRYD IR had an impact against violent crimes in general. We fit the spatial-temporal multivariate model [2], with an exponential kernel for g (see Materials and Methods), to the full comple-

Table 1. Gang and non-gang crimes reported only to the LAPD and to the LAPD and GRYD IR in 2014-2015.

	LAPD-only				LAPD + GRYD IR				Grand Total	
	Gang N	Non-Gang N	TOTAL N	% Gang	Gang N	Non-Gang N	TOTAL N	% Gang	TOTAL N	% Gang
Agg. Assault	1,249	3,918	5,167	24.2%	470	99	569	82.6%	5,736	30.0%
Homicide	41	29	70	58.6%	152	24	176	86.4%	246	78.5%
TOTAL	1,290	3,947	5,237	24.6%	622	123	745	83.5%	5,982	32.0%

ment of 5,982 crimes (Fig. 2B). The estimate of parameter k_{11} indicates each aggravated assault or homicide exposed to the treatment triggered on average 0.1401 retaliations subsequently known to both LAPD and GRYD IR. Parameter k_{01} indicates each control crimes triggered on average 0.0526 retaliations. The estimated 62.5% higher rate of retaliation stemming from events exposed to the treatment condition is statistically significant ($p = 0.0092$) (30)(SI Appendix). Pathway k_{10} shows that treatment crimes also triggered on average 0.2841 retaliations known only to the LAPD, which is of equivalent magnitude to pathway k_{00} with 0.2824 retaliations ($p = 0.486$). In practical terms, the fitted model suggests that every 100 LAPD + GRYD IR (treatment) aggravated assaults and homicides triggered on average 42.4 retaliatory violent crimes ($k_{11} + k_{10}$), compared to an average 33.5 retaliations triggered by LAPD-only (control) crimes ($k_{01} + k_{00}$). GRYD IR may reduce retaliation among the broader set of gang and non-gang violent crimes, but not to levels characteristic of the LAPD control case, with its more generous mix of non-gang crimes and inherently lower risk of retaliation.

We therefore restricted analyses to gang aggravated assaults and homicides to ensure that test conditions were evaluated for events with similar potential for spawning retaliation. Against this set of crimes GRYD IR had a substantial impact (Fig. 2C). Pathway k_{11} triggered an average of 0.0015 retaliations for any one treatment gang crime. By contrast, pathway k_{01} triggered 0.0621 retaliations for any one control gang crime. This represents a 97.6% reduction in retaliation associated with GRYD IR notification ($p < 10^{-6}$). Pathway k_{10} triggered an average of 0.1483 retaliations for any one treatment gang crime. Pathway k_{00} triggered 0.2116 retaliations for any one control gang crime. The 29.9% reduction in retaliations is also significant ($p = 0.0163$). Every 100 LAPD + GRYD IR (treatment) gang crimes triggered an average of 15.0 retaliations ($k_{11} + k_{10}$) compared to 27.4 retaliations triggered by LAPD-only (control) crimes ($k_{01} + k_{00}$). Overall, the notification of GRYD IR was associated with a 45.3% decrease in retaliatory gang crimes.

To better understand the spatial dynamics of retaliatory gang violence we mapped the background intensity and self-excitation parts of the conditional intensity along with the distributions of background and retaliatory crimes determined via stochastic declustering (31) (SI Appendix) (Fig. 3A-F). The background risk of gang violence is characterized by numerous compact, but widely distributed hot spots (Fig. 3A), consistent with the observation that the opportunities for violence and strengths of gang rivalries are geographically variable (8, 32). The risk of retaliation is concentrated in more continuous bands (Fig. 3D), bridging the discrete areas of background risk. Notably there is a prominent North-South corridor of retaliatory risk that maps to an area locally known as 'death alley' (33). The patterns of risk influence the distribution of gang violent crimes (Fig. 3C and F). The density of background crimes forms five distinct hot spots (Fig. 3B) suggesting that background crimes are of local, neighborhood origin. The density of retaliatory crimes occupies two distinct hot spots (Fig. 3E), suggesting that retaliation spreads contagiously beyond immediate neighborhood contexts.

Stochastic declustering also allows us to evaluate differences in the frequency of retaliation by crime type across test conditions (Table 2). Background crimes make up 76.6% of all gang aggravated assaults and homicides for both test conditions combined. Retaliatory crimes are proportionally more common among events assigned to the LAPD-only control condition. This imbalance is pronounced for gang aggravated assaults (46.3% retaliation for LAPD-only vs. 10.3% for LAPD + GRYD IR), but particularly extreme for homicides (24.2% LAPD-only vs. 0.7% for LAPD + GRYD IR), especially considering the baseline bias towards notifying GRYD IR of most gang homicides.

We use the estimated treatment effects along with the results of stochastic declustering to compute numbers of prevented crimes (Table 2)(SI Appendix). The sum $(k_{11} + k_{01})$ is the average number of retaliations known to both the LAPD and GRYD IR produced by any one event (Fig. 1B). Similarly, $(k_{10} + k_{00})$ is the average number of retaliations known only to the LAPD produced by any one event. Note that $(k_{11} + k_{01})$ and $(k_{10} + k_{00})$ are measured directly from data and therefore are observed outcomes. We now define two counterfactuals. Let $(k_{01} + k_{01})$ be the average number of retaliations that would have been triggered in the absence of GRYD IR for those events assigned to LAPD + GRYD IR. Let $(k_{00} + k_{00})$ be the average number of retaliations that would have been triggered in the absence of GRYD IR for those events assigned to LAPD-only. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. From stochastic declustering, the observed number of gang retaliations arising from pathways $(k_{11} + k_{01})$ is 45 and from $(k_{10} + k_{00})$ is 403 (Table 2). The counterfactual conditions suggest that retaliatory gang crimes would have been 48.8% and 15.0% higher in the absence of GRYD IR for observed pathways $(k_{11} + k_{01})$ and $(k_{10} + k_{00})$, respectively. GRYD IR prevented an estimated total 82.2 retaliatory gang crimes, of which 77.8 are expected to have been aggravated assaults and 4.4 are expected to have been homicides.

Discussion

The multivariate self-exciting point process model allows us to quantify how GRYD IR disrupts retaliatory gang crimes. Instead of determining the causal structure of each event, productivities $K_{u_i u}$ are averaged over the entire time series, which yields very robust and stable estimates. This is particularly important in real-world empirical settings where control and test conditions are often mixed by structural and practical conditions beyond the control of the observer. The multivari-

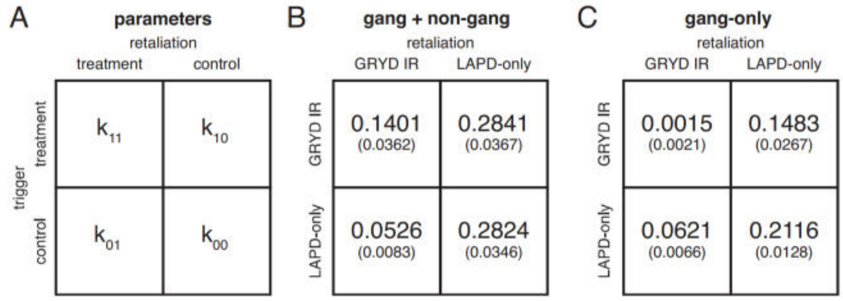


Fig. 2. (A) Matrix representation of productivity K_{ij} with the corresponding triggering pathways noted. Matrix entries are the average number of retaliations assigned to an experimental condition i triggered by an event assigned to condition j. (B). The productivity for combined gang and non-gang aggravated assaults and homicides in South Los Angeles for 2014-15. (C). The productivity for gang-only aggravated assaults and homicides in South Los Angeles for 2014-15. The control condition includes violent crimes known to the LAPD (LAPD-only). The treatment condition includes crimes known to the LAPD that were also reported to GRYD IR (LAPD + GRYD IR). Standard errors of parameter estimates are shown in parentheses.

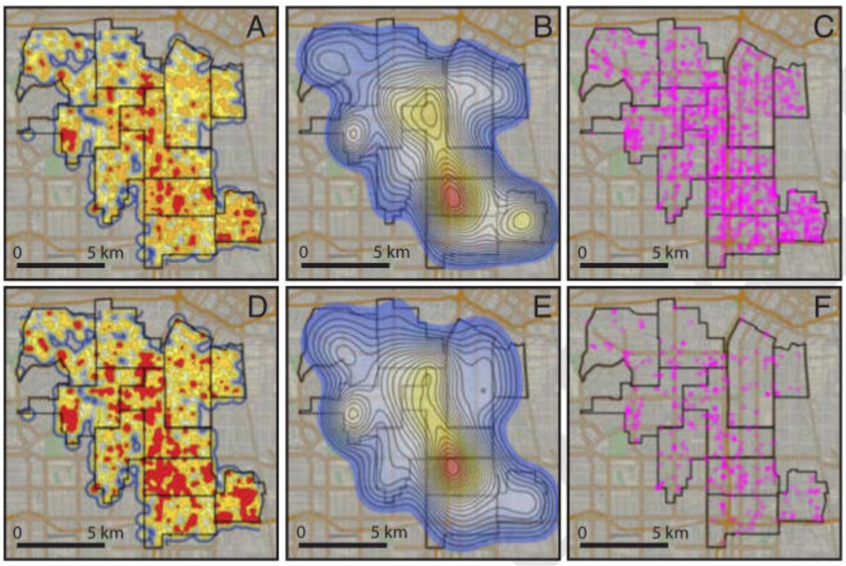


Fig. 3. Stochastic declustering of gang crimes in South Los Angeles. (A). The log of spatial-temporal background intensity function μ for gang violent crimes mapped over space. (B) Contour plot of the density of background gang aggravated assaults and homicides determined by declustering. (C) Point locations of background gang aggravated assaults and homicides determined by declustering. (D) The log of spatial-temporal self-excitation of retaliation $\lambda - \mu$ mapped over space. (E) Contour plot of the density of retaliatory gang aggravated assaults and homicides determined by declustering. (F) Point locations of retaliatory gang aggravated assaults and homicides determined by declustering. Boundaries for the ten GRYD IR Zones in South Los Angeles are outlined in black.

Table 2. Number of retaliatory and background aggravated assaults and homicides in South Los Angeles in 2014-2015 separated by test condition.

	LAPD-only				LAPD + GRYD IR			
	Retaliation	Background	TOTAL	% Retaliation	Retaliation	Background	TOTAL	% Retaliation
Agg. Assault	395	854	1,249	46.3%	44	426	470	10.3%
Homicide	8	33	41	24.2%	1	151	152	0.7%
TOTAL	403	887	1,290	45.4%	45	577	622	7.8%

ate modeling framework embraces this limitation and allows for interactions between conditions to proceed as part of the analysis. In the present case, we have the added benefit that the process by which GRYD IR was notified of gang violent crimes generated *as if random* assignment of events between the two test conditions. We were thus able to more completely disentangle the effects of intervention from other confounds that might lead some events to be preferentially treated only by the police and others by both the police and GRYD IR. We find that GRYD IR reduced gang retaliation by 45.3%, which corresponds to 82.2 fewer gang aggravated assaults and homicides over a two year period. Recent estimates place the overall cost of a single aggravated assault at $240,000 and a single homicide at $8.98 million (34). Over the two-year period in 2014-15, the potential savings from gang aggravated assaults and homicides prevented by GRYD IR is estimated at $49.0 million in South Los Angeles alone.

These results have important implications for the design, implementation and evaluation of gang violence intervention programs. CIWs provide tangible benefits in reducing gang violence that complement, rather than compete with the effects of policing. The impact likely traces to the intimate connection CIWs share with the community they serve. CIWs are usually individuals who grew up in the community and may themselves have had some prior involvement with gangs. This high degree of social embedding and their prior experience gives CIWs a so-called *license to operate* in the community and from which they derive authority to mediate conflicts at the street level. Civilian intervention workers operating at the street level are not a new phenomenon (35). But it has been difficult to establish their effectiveness independent of other violence-prevention efforts and the actions of police. The results presented here indicate that finding, cultivating and retaining such individuals would seem to be an essential part of designing gang violence interruption programs.

Our study leverages noise in the gang crime reporting system to detect and quantify the effects of violence interruption. In the ideal case, GRYD IR would receive notification for each and every gang crime occurring within their area of operation. In practice, however, notification is made only for a fraction of all events. This random process of assignment is serendipitous for making causal inferences, but also produces missed opportunities to further reduce gang retaliation. Specifically, eligible events that did not receive GRYD IR attention due to noisy notification processes, likely produced more retaliations than would otherwise have been the case. Implementation of violence interruption programs would benefit from improvement in notification procedures that ensures more events are brought to the attention of CIWs.

Finally, our approach is distinctive from studies that seek to measure average treatment effects of anti-violence programs before and after the onset of treatment, or between geographically non-overlapping control and treatment regions (17). In the present case, stochastic declustering suggests that a majority of gang crimes are actually background events (Table 2). Violence interruption efforts to reduce retaliation may have little impact on the independent processes that generate these background events. Given the low frequency of retaliatory crimes relative to background events, average treatment effects tied to violence interruption may be diluted beyond detection by these background processes. We model the processes of gang crime retaliation and interventions at the scale of the events themselves. This micro-scale focus allows us to maintain the important distinction between background and retaliatory events and consequently establish that violence interruption does indeed work to reduce gang retaliation.

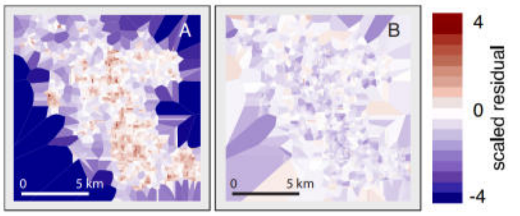


Fig. 4. Voronoi residuals for spatial-temporal Poisson (A) and Hawkes process models (B). Color scaling identifies the spatial locations where a model is over or under estimated. Cells with blue shades (negative residuals) indicate an overestimation of the intensity, while cells with red shades (positive residuals) indicate an underestimation.

Materials and Methods

Prior research on self-exciting point process models of crime suggest particular parametric equations for μ and g (36, 37):

$$\mu_u(x, y) = \sum_{i=1}^{N} \frac{\beta_{u_i u}}{2\pi\sigma^2 T} \times \exp\left(-\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma^2}\right).$$

$$g(x, y, t) = \omega \exp(-\omega t) \times \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

where β is a weights matrix for the degree to which events assigned to condition u_i contribute to the background rate for events assigned to condition u, T is the total time period represented by the sample of gang crimes, and σ is the spatial scale of influence for background events and retaliatory events. The kernel g is transformed to polar coordinates. The time scale of self-excitation is governed by ω (see Fig. 1A).

We use expectation maximization (EM) to estimate model [2] (27, 38). The expectation step is used to compute initial probabilities that an event i causes event j via either the background rate μ or the self-exciting kernel g. These expectations are then fed to the maximization step where a new set of parameter values (for iteration $k+1$) are determined by maximizing the expected probability with respect to the observed data. This maximization is done for all parameters taking into consideration the condition u to which a gang crime is assigned. The algorithm alternates between expectation and maximization until there is no further change in the parameter values. The full EM algorithm is presented in the SI Appendix.

Once a model is estimated, Voronoi residuals provide a powerful technique for evaluating model performance (28) (Fig. 4). Voronoi residuals measure the differences between the modeled conditional intensity and the observed number of points within spatially adaptive Voronoi cells. Comparison to a color scale defined by a null Poisson model further helps interpret performance (29). The null model corresponds to a case where all gang crimes are statistically independent background events. Voronoi residuals from the fitted intensity for the multivariate Hawkes model (Fig. 4B) has muted colors as compared to the null model (Fig. 4A), indicating improved performance. The fitted Hawkes process model performed well throughout the spatial window. The Poisson model underestimates the conditional intensity in areas such as the so-called 'death alley' where the Hawkes model identifies the greatest amount of retaliation (see Fig. 3D).

To address the question of whether our conclusion is dependent on the choice of parametric model, we constructed a model-independent multivariate self-exciting point process following the methods of (21, 26). We estimated this point process on the same dataset non-parametrically. The results presented in the SI Appendix reinforce the conclusion that GRYD IR reduces gang retaliations.

We tested the parametric model on synthetic datasets corresponding to different intervention scenarios. We extended the methods in (39) to allow simulation of space-time multivariate Hawkes processes with different triggering matrices K. We chose triggering matrices corresponding to hypothetical positive, neutral and negative effects of GRYD IR intervention relative to the control. The results presented in the SI Appendix show that our methods are effectively able to recover ground-truth model parameters from simulated events and statistically distinguish between different intervention effects.

ACKNOWLEDGMENTS. This research was made possible by the City of Los Angeles Mayor's Office of Gang Reduction and Youth Development (GRYD) and Los Angeles City contract #C-128086 ("Research Consulting for GRYD"). The authors thank GRYD Director Anne Tremblay, GRYD Office Staff and GRYD Intervention Incident Response Providers. We are also grateful to the Los Angeles Police Department for providing crime data. Thank you to California State University Los Angeles and University of California Los Angeles for their ongoing support. Any opinions, findings, conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of the GRYD Office.

1. Decker SH (1996) Collective and normative features of gang violence. *Justice Quarterly* 13(2):243–264.
2. Klein MW, Maxson CL (2006) *Street gang patterns and policies.* (Oxford University Press, New York).
3. Hughes LA, Short JF (2005) Disputes involving youth street gang members: Micro-social contexts. *Criminology* 43(1):43–76.
4. Jacobs BA, Wright R (2006) *Street justice: Retaliation in the criminal underworld.* (Cambridge University Press).
5. Bjerregaard B, Lizotte AJ (1995) Gun ownership and gang membership. *The Journal of Criminal Law and Criminology (1973-)* 86(1):37–58.
6. Howell JC (2011) *Gangs in America's communities.* (Sage).
7. Rosenfeld R, Bray TM, Egley A (1999) Facilitating violence: A comparison of gang-motivated, gang-affiliated, and nongang youth homicides. *Journal of Quantitative Criminology* 15(4):495–516.
8. Papachristos AV (2009) Murder by structure: Dominance relations and the social structure of gang homicide 1. *American Journal of Sociology* 115(1):74–128.
9. Stewart EA, Simons RL (2010) Race, code of the street, and violent delinquency: A multilevel investigation of neighborhood street culture and individual norms of violence. *Criminology* 48(2):569–605.
10. Dymnicki AB, Henry D, Quintana E, Wisnieski E, Kane C (2013) Outreach workers' perceptions of positive and negative critical incidents: Characteristics associated with successful and unsuccessful vioence interruption. *Journal of Community Psychology* 41(2):200–217.
11. Skogan WG, Hartnett SM, Bump N, Dubois J (2009) *Evaluation of CeaseFire-Chicago.* (U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington, D.C.).
12. Webster DW, Whitehill JM, Vernick JS, Curriero FC (2013) Effects of baltimore's safe streets program on gun violence: A replication of chicago's ceasefire program. *Journal of Urban Health* 90(1):27–40.
13. Whitehill JM, Webster DW, Frattaroli S, Parker EM (2014) Interrupting violence: how the ceasefire program prevents imminent gun violence through conflict mediation. *Journal of Urban Health* 91(1):84–95.
14. Wilson JM, Chermak S (2011) Community-driven violence reduction programs. *Criminology & Public Policy* 10(4):993–1027.
15. Butts JA, Roman CG, Bostwick L, Porter JR (2015) Cure violence: a public health model to reduce gun violence. *Annual review of public health* 36:39–53.
16. Klein MW (2011) Comprehensive gang and violence reduction programs. *Criminology & Public Policy* 10(4):1037–1044.
17. Papachristos AV (2011) Too big to fail. *Criminology & Public Policy* 10(4):1053–1061.
18. Cressie N, Wikle CK (2015) *Statistics for spatio-temporal data.* (John Wiley & Sons).
19. Delay D, Vere-Jones D (2003) An introduction to the theory of point processes. vol. i.
20. Diggle PJ (2013) *Statistical analysis of spatial and spatio-temporal point patterns.* (CRC Press).
21. Marsan D, Lengline O (2008) Extending earthquakes' reach through cascading. *Science* 319(5866):1076–1079.
22. Hawkes AG (1971) Spectra of some self-exciting and mutually exciting point processes. *Biometrika* pp. 83–90.
23. Short M, Mohler G, Brantingham PJ, Tita G (2014) Gang rivalry dynamics via coupled point process networks. *Discrete & Continuous Dynamical Systems-Series B* 19(5).
24. Green B, Horel T, Papachristos AV (2017) Modeling contagion through social networks to explain and predict gunshot violence in chicago, 2006 to 2014. *JAMA internal medicine.*
25. Wiley SA, Levy MZ, Branas CC (2016) *The Impact of Violence Interruption on the Diffusion of Violence: A Mathematical Modeling Approach*, eds. Letzter G, et al. (Springer International Publishing, Cham), pp. 225–249.
26. Fox EW, Schoenberg FP, Gordon JS (2016) Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric hawkes point process models of earthquake occurrences. *The Annals of Applied Statistics* 10(3):1725–1756.
27. Veen A, Schoenberg FP (2008) Estimation of space–time branching process models in seismology using an em–type algorithm. *Journal of the American Statistical Association* 103(482):614–624.
28. Bray A, Wong K, Barr CD, Schoenberg FP, , et al. (2014) Voronoi residual analysis of spatial point process models with applications to california earthquake forecasts. *The Annals of Applied Statistics* 8(4):2247–2267.
29. Gordon JS, Clements RA, Schoenberg FP, Schorlemmer D (2015) Voronoi residuals and other residual analyses applied to csep earthquake forecasts. *Spatial Statistics* 14:133–150.
30. Ogata Y (1978) The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics* 30(1):243–261.
31. Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association* 97(458):369–380.
32. Tita G, Ridgeway G (2007) The impact of gang formation on local patterns of crime. *Journal of Research in Crime and Delinquency* 44(2):208–237.
33. Santa Cruz N, Schwenke K (2014) South vermont avenue: L.a. county's 'death alley'.
34. McCollister KE, French MT, Fang H (2010) The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and alcohol dependence* 108(1):98–109.
35. Klein M (1971) *Street Gangs and Street Workers.* (Prentice Hall, Englewood Cliffs).
36. Mohler G, Short M, Brantingham P, Schoenberg F, Tita G (2011) Self-exciting point process modeling of crime. *Journal of the American Statistical Association* 106(493):100–108.
37. Mohler G, et al. (2015) Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association* 110(512):1399–1411.
38. Mohler G (2014) Marked point process hotspot maps for homicide and gun crime prediction in chicago. *International Journal of Forecasting* 30(3):491–497.
39. Lewis PAW, Shedler GS (1979) Simulation of nonhomogeneous poisson processes by thinning. *Naval Research Logistics Quarterly* 26(3):403–413.

Supporting Information

Brantingham et al.

Materials

Data Sources. The analyses presented rely on data collected by the Los Angeles Police Department (LAPD) as well as the City of Los Angeles Mayor's Office of Gang Reduction Youth Development (GRYD). Data provided by the LAPD include only officially reported crimes that have been through the Department's standard process of verification and quality control. Neither calls for service data, nor suspect and arrest data were used. The LAPD data includes records for all reported crime types ranging from public disorder to homicide. Most of these crime types are not directly relevant to understanding the dynamics of gang violence and the impact of GRYD Intervention Incident Response (GRYD IR). Our analyses focus on gang aggravated assaults and homicides, which are more frequently committed for the benefit of the gang and are particularly prone to initiating retaliation. We adopt an inclusive approach and treat any aggravated assault or homicide as gang-related if the LAPD or GRYD program labels it as such.

GRYD and GRYD IR. The City of Los Angeles Mayor's Office of Gang Reduction Youth Development (GRYD) deploys a comprehensive strategy aimed at reducing gang involvement and gang violence. It focuses on the provision of gang prevention and intervention services, violence interruption activities and proactive peace-making. Prevention services are aimed at providing alternatives for youth before they join gangs. Intervention services are aimed at provide pathways out of gang life for youth that have already become involved. Violence interruption seeks to disrupt retaliation in the aftermath of gang violent crimes. Proactive peacemaking reflects continuous efforts of Community Intervention Workers (CIWs) to tamp down general community tensions not tied to any one event. Violence interruption within the GRYD program is termed GRYD Intervention Incident Response (GRYD IR) and is the focus of our analysis. The GRYD Office was first established in 2007, community-based service provision began in 2009, and the GRYD comprehensive strategy was created in 2011.

The GRYD program design overlaps in part with Chicago's Operation Ceasefire (now Cure Violence) (1). Violence interruption in Chicago's Cease Fire and its daughter programs (2, 3), concentrates both on preventing a range of conflict types, many of which may not be formally reported, from escalating into retaliatory shootings as well as retaliation for prior shootings and homicides. By contrast, GRYD IR is focused on responding to reported shootings or homicides and stopping retaliation in the aftermath of these events. This difference allows us to focus on the statistical dependence between reported gang crimes.

South Los Angeles gang violence & GRYD IR Zones.

We focus our analysis on violent crimes occurring in an 87.2 km^2 (33.7 sq mi) area of South Los Angeles during 2014-2015 (Fig. S1). As of mid-2015, ten GRYD Zones operated in South Los Angeles forming a contiguous case study area. Prior to mid-2015, only seven of the ten regions shown were formally recognized. Nevertheless, GRYD IR received notifications and responded to crimes over this entire geographic area throughout 2014 and 2015. For example, between July 1, 2014 and December 31, 2014, GRYD IR was notified of 109 events over the South Los Angeles region (Fig. S2). Between July 1, 2015 and December 31, 2015, GRYD IR was notified of 129 events over this region, only an 18% increase in notifications that occurred against a backdrop of increasing violent crime across the region. The 77th 3 GRYD Zone is the only possible exception. This GRYD Zone received few notifications prior to its formal addition. In all other locations, GRYD IR crimes were recorded regardless of whether there was a formal GRYD Zone in place or not.

We therefore treat the South Los Angeles GRYD Zones as a single continuous study region for 2014 and 2015. This region is well bounded, but still expansive enough to understand the spatial dynamics of gang retaliatory violence. There is no official tally of the number of gangs present in the area. The GRYD ETO database notes 54 unique gang names in association with the suspects and victims of gang crimes in the area. This estimate is likely a lower bounds as many events lack information about suspect and victim gang affiliation. Taken at face value, the area hosts 0.6 gangs per km^2. During this period, there were 5,982 aggravated assaults and homicides crimes reported to the LAPD strictly inside the South Los Angeles GRYD Zones (main text Table 1). Of these, 1,912 were flagged as gang related. GRYD IR was notified on 745 of the violent crimes, with 622 of the notifications flagged as gang related.

GRYD IR Notification and Field Deployment. Gang crimes are typically reported to the LAPD first by members of the public. GRYD IR receives notification from the LAPD and from CIWs, when they are independently contacted by the community. However, notification of GRYD IR only occurs for a fraction of all reported crimes. We include all events where GRYD IR is notified, rather than restricting analysis to events where there is also some record of field activity by CIWs. This analytical choice is made due to uncertainty surrounding the measurement of dosage associated with CIW activities (1: 5-3, 4, 5).

The resulting hypotheses based on notification of GRYD IR are conservative. On the one hand, we suppose that CIWs can only have a direct impact on retaliation if they know about a potential triggering event. Thus, notification of GRYD IR is a logical precursor for treatment effects. If notification did not lead to interventions in the field by CIWs then we would expect there to be no difference between the control and treatment conditions. Specifically, even though events were labeled as LAPD + GRYD IR, the absence of field intervention would ensure that such events were no different than LAPD-only ones. On the other, if CIWs were able to source information about events on their own and self-deploy, operating outside of the normative channel of communication, then we would expect this contamination to also bias the results towards finding no difference between treatment and control. In other words, some fraction of events labeled as LAPD-only would include the effects of GRYD IR without formal recognition. Thus, LAPD-only interventions would appear similar to LAPD + GRYD IR because of this hidden activity. Because we are able to document a significant difference between LAPD-only and LAPD + GRYD IR intervention effects (main text Fig. 2C) we conclude such confounds have a minimal impact.

Methods

Runs tests and K-S test. GRYD IR was notified on only a fraction of the total number of gang crimes occurring in the South Los Angeles area. We hypothesized that the notifications were random and independent, establishing test conditions approximating a randomized controlled trial. We performed one-sample runs tests to evaluate whether notifications received by GRYD IR occur at random. We ran two-sample KS tests to test whether GRYD IR and LAPD events have the same spatial-temporal distribution, based on the distance between the empirical distribution functions of the two samples.

Consider a gang crime reported to the LAPD. With probability ρ, GRYD IR is notified of the event. We use LAPD + GRYD IR as a short-hand to identify these events. With probability $1 - \rho$ the crime remains known only to the LAPD. These events are noted as LAPD-only. Over a large sequence of reported gang crimes, we expect a fraction of crimes proportional to ρ to fall into condition LAPD + GRYD IR and a fraction proportional $1 - \rho$ to fall into condition LAPD-only. However, we also expect the assignment of any one event to be random and independent of other assignments. To test this hypothesis, reported gang crimes were arranged in ascending order of the date and time received and computed a runs test. A run is defined as a sequence of notifications of the same condition, uninterrupted by notifications of the other condition. For example, the sequence of coin tosses HHTHHHTTTT consists of four runs {HH}{T}{HHH}{TTTT}. Let R be the observed number of runs in a sample of size N. The null hypothesis is that the observed number of runs R is the product of random and mutually independent notifications. Note that ρ need not equal to $1 - \rho$ as would be the case in fair coin toss. All that is required is that each GRYD IR notification occur at probability ρ independently.

Results for different subsets of the observed data fail to reject the null hypothesis (Table S1). For gang aggravated assaults and homicides combined, the number of observed runs (R = 852) is not statistically different from the expected number of runs under random assignment (E[R] = 840.31) (Z = 0.609, p = 0.54). For gang aggravated assaults alone, the observed number of runs (R = 672) is not statistically different from the expected number under random assignment (E[R] = 683.99) (Z = 0.73, p = 0.47). Similarly, for gang homicides the observed number of runs (R = 65) is not statistically different from the expected number under random assignment (E[R] = 65.58)(Z = 0.125, p = 0.90). While the percent of crimes that lead to GRYD IR notifications differs by crime type, these notifications are random on a per event basis. Furthermore, the results of two-sample KS tests fail to reject the null hypothesis that the data come from the same underlying temporal (KS = 0.069, p = 0.20) and spatial distributions (KS = 0.054, p = 0.17) Therefore, we conclude that the assignment of events to different conditions approximates and randomized experimental protocol.

Expectation Maximization (EM). The estimation procedure for model [2] in the main text is a type of Maximum Likelihood Estimation (MLE) known as expectation maximization (EM) (6, 7). The expectation step of the EM algorithm is used to compute initial probabilities p^b_{ij} and p_{ij} that an event i causes event j via either the background rate μ or the self-exciting kernel g, respectively. These expectations are then fed to the maximization step where a new set of parameter values (for iteration $k + 1$) are determined by maximizing the expected probability with respect to the observed data. This maximization is done for all parameters taking into consideration whether gang crimes are known only to the LAPD or to both the LAPD and

GRYD IR. The algorithm alternates between expectation and maximization until there is no further change in the parameter values.

For completeness, the EM algorithm is structured as below. Note that $n_{\hat{u}}$ is the number of events that belongs to experimental condition type $\hat{u}$:

Complete Data Likelihood Function:

$$Q(\Omega)=\sum_{i=1}^{N}\sum_{j=1}^{N}p_{ij}^{b}log(\frac{\beta_{u_iu_j}}{2\pi\eta^2T}e^{\left(-\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2\eta^2}\right)})-\sum_{u=1}^{U}\sum_{i=1}^{N}\beta_{u_iu}$$

$$+\sum_{i<j}p_{ij}log(\omega K_{u_iu_j}e^{-\omega(t_j-t_i)}\frac{1}{2\pi\sigma^2}e^{-\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2\sigma^2}})-\sum_{u=1}^{U}\sum_{i=1}^{N}K_{u_iu}(1-e^{-w(T-t_i)})$$

Expectation Step:

$$p_{ij}=K_{u_iu_j}\omega\frac{\exp\left(-\omega(t_j-t_i)\right)}{2\pi\sigma^2}\times\frac{\exp\left(-\frac{(x_j-x_i)^2+(y_j-y_i)^2}{2\sigma^2}\right)}{\lambda_{u_j}(x_j,y_j,t_j)}$$

$$p_{ij}^{b}=\frac{\beta_{u_iu_j}}{2\pi\eta^2T}\frac{\exp\left(-\frac{(x_j-x_i)^2+(y_j-y_i)^2}{2\eta^2}\right)}{\lambda_{u_j}(x_j,y_j,t_j)}$$

Maximization Step:

$$\omega^{(k+1)}=\frac{\sum_{j<i}p_{ij}^{(k)}}{\sum_{i<j}p_{ij}^{(k)}(t_j-t_i)+\sum_{u=1}^{U}\sum_{i=1}^{N}K_{u_iu}(T-t_i)e^{-\omega(T-t_i)}}$$

$$\sigma^{2(k+1)}=\frac{\sum_{i<j}p_{ij}^{(k)}\left((x_i-x_j)^2+(y_i-y_j)\right)^2}{2\sum_{i<j}p_{ij}^{(k)}}$$

$$\eta^{2(k+1)}=\frac{\sum_{i,j=1}^{N}p_{ij}^{b(k+1)}\left((x_i-x_j)^2+(y_i-y_j)^2\right)}{2\sum_{i,j=1}^{N}p_{ij}^{b(k+1)}}$$

$$K_{\hat{u}u}^{(k+1)}=\frac{\sum_{i,j\in A_{\hat{u}u}}p_{ij}}{\sum_{l=1}^{n_{\hat{u}}}(1-e^{-w(T-t_{il})})}$$

$$\beta_{\hat{u}u}^{(k+1)}=\frac{\sum_{i,j\in B_{\hat{u}u}}p_{ij}^{b(k)}}{n_{\hat{u}}}$$

$$A_{\hat{u}u}=\left\{i,j\text{ index of events}\mid t_i<t_j,u_i=\hat{u},u_j=u\right\}$$
$$B_{\hat{u}u}=\left\{i,j\text{ index of events}\mid u_i=\hat{u},u_j=u\right\}$$

Non-parametric Model Fitting. Our parametric model choices were based on previous research on crime patterns indicating that exponential kernels provide a good description of the data (8). We extended the fully non-parametric model estimation methods in (9) and (10) to the multivariate case. The non-parametric model is similar to the parametric form:

$$\lambda_u(x,y) = \mu_u(x,y) + \sum_{t_i<t} K_{u_i u} v(t-t_i, x-x_i, y-y_i).$$

But here

$$\mu_u(x,y) = \gamma_u \tau(x,y) = \frac{\gamma_u}{T}\sum_{i=1}^{N}\frac{p_{ii}}{2\pi d_i^2} exp(-\frac{(x-x_i)^2+(y-y_i)^2}{2d_i^2})$$

and we assume $v(x,y,t) = g(t)f(x,y)$, which will be estimated non-parametrically. The term d_i is computed by finding the radius of the smallest disk centered at (x_i, y_i) that contains at least n_p other events, and is greater than some small value ϵ representing the location error. In (11) they suggest taking n_p between 15-100 and $\epsilon = 0.02$ degrees.

The log-likelihood function is:

$$l = \sum_{u=1}^{U}\left(\sum_{i=1}^{N} log(\lambda_u(t_i,x_i,y_i)) - \int_0^T \int\int_S \lambda_u(t,x,y)dsdt\right).$$

From EM we determine the nonparametric algorithm. We define p_{ij} as the probability that event i triggers j for $t_i < t_j$ and p_{ii} as the probability that i is from background and $p_{ij} = 0$ for $t_i > t_j$. We define n_t^{bins} as the number of bins in time and n_r^{bins} as the number of bins in space. C_k is the set of events pairs (i, j) such that t_j - t_i belongs to the k^{th} bin. D_k is the set of events pairs (i, j) such that r_{ij}, the distance between i and j, belongs to the k^{th} bin. N_α is the number of events with type u. Finally, δ_t is the size of k^{th} bin in time and δ_r is the size of the k^{th} bin in space. Further discussion of parameters can be found in (10).

The algorithm for our nonparametric method is:

Step 1: Initialize the $P^{(0)} = (p_{ij})$ matrix randomly, index $v = 0$.

Step 2: Update

$$\gamma_u^{(v)} = \frac{\sum_{u_i=u} p_{ii}^{(v)}}{Z^{(v)}}$$

where u_i is the type of event i and

$$\frac{1}{Z^{(v)}}\int_0^T \int\int_S \tau(x,y)dsdt = 1.$$

Step 3: Update

$$K_{\alpha\beta}^{(v)} = \frac{\sum_{u_i=\alpha}\sum_{u_j=\beta} p_{ij}}{N_\alpha}, \quad g_k^{(v)} = \frac{\sum_{i,j\in C_k} p_{ij}^{(v)}}{\delta t_k \sum_{i,j} p_{ij}^{(v)}}, \quad \text{and } h_k^{(v)} = \frac{\sum_{i,j\in D_k} p_{ij}^{(v)}}{\delta r_k \sum_{i,j} p_{ij}^{(v)}},$$

where $\alpha = 1$ and $\beta = 2$ in our case, $k = 1, \ldots, n_t^{bins}$ and $k = 1, \ldots, n_r^{bins}$ for $g_k^{(v)}$ and $h_k^{(v)}$, respectively.

Step 4: Update

$$p_{ij}^{(v+1)} = K_{u_i u_j}^{(v)} g^{(v)}(t_j - t_i) f^{(v)}(r_{ij}) \quad \text{for } t_i < t_j,$$

and

$$p_{jj}^{(v+1)} = \mu_{u_j}(x_j, y_j).$$

Then normalize such that for any j $\sum_{i=1}^{N} p_{ij} = 1$. Here $2\pi r f^{(v)} = h^{(v)}(r)$.

Step 5: If $max_{ij} \left\| p_{ij}^{(v+1)} - p_{ij}^{(v)} \right\| < \epsilon$, then the algorithm has converged. In practice, we take $\epsilon = 10^{-3}$. Otherwise, set $v \longleftarrow v + 1$ and repeat Steps 2–5 until convergence.

Voronoi Residuals Analysis. A powerful technique for evaluating model performance is Voronoi residuals (12). Voronoi residuals allow the examination of differences between the modeled conditional intensity and the observed number of points within spatially adaptive Voronoi cells. Voroni residuals for a Poisson process model provides a baseline case where all crimes are assumed to be statistically independent background events (13). Maps of Voronoi residuals use color scaling allows to see the spatial locations where a model is over or under estimated.

Forward simulation of multivariate Hawkes and parameter recovery via EM. Here we describe a procedure of simulating multivariate Hawkes process based on (14) Algorithm C. The multivariate Hawkes process is defined as:

$$\lambda_u(x,y) = \mu_u(x,y) + \sum_{t_i<t} K_{u_i u} v(t - t_i, x - x_i, y - y_i).$$

Here $v(t - t_i, x - x_i, y - y_i) = f(x,y)g(t)$.

Step 1: Generate background events. Draw the number of background events in type u from Poisson distribution with $\lambda = \gamma_u * T$. Then draw each background point in each type u with spatial uniform distribution and temporal uniform distribution in [0, T].

Step 2: Generate self-excited events. For each point i of type u, the number of triggered points N_i is drawn from Poisson distribution with $\lambda = \sum_{u'=i}^{U} K_{uu'}$, where $U = 2$ in our case. Then draw each triggered point with spatial distribution $f(x - x_i, y - y_i)$ and temporal distribution $g(t - t_i)$. Finally, this new point belongs to type u' with probability:

$$\frac{K_{uu'}}{\sum_{u'=i}^{U} K_{uu'}}.$$

- **Step 3:** Stop criterion. For each event i, if $t_i > T$ or $N_i < 0.5$, then stop. Otherwise repeat step 2.

Using this forward simulation process we generate synthetic datasets to test whether we are able to accurately recover parameter values and detect intervention effects.

Stochastic declustering. Gang crimes occurring in a given area represent a mixture of those that are background events and those that are retaliatory in response to other crimes. We wish to sort events into these two groups to understand how important background and retaliatory processes are for gang violence overall.

Stochastic declustering is a suite of methods developed in the study of earthquake catalogs where the goal is to distinguish between background seismicity and aftershocks (11). The same methods can be applied to the study of crime (15).

Starting with a self-exciting point process model like the one developed here, stochastic declustering proceeds through a thinning procedure that removes events probabilistically classified as retaliations. The events remaining after thinning represent the background events generated by a spatially non-homogeneous Poisson process $\lambda(t, x, y) = \mu(x, y)$. Specifically, in the univariate case, the probability that an event j is a retaliation is given by

$$\rho_j = \frac{\sum_{t_i < t_j} Kg(t_j - t_i, x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j)}.$$

The probability that an event j is a background event is therefore

$$1 - \rho_j = \frac{\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j)}.$$

For a catalog of N total crimes and a point process model fit to those events, the simplest procedure is to generate N uniform random variables $U_1, U_2, \ldots, U_N$ in the range [0,1]. An event is classified as a background crime when $U_j < 1 - \rho_j$, otherwise it is removed and classified as a retaliation (11).

Note that the assignment of an event to being background or retaliation is a probabilistic classification. On average the relative mixture of background and retaliation events is correct for a given time window and spatial region, but we cannot say with absolute certainty whether any specific event is or is not a retaliation.

Estimating Statistical Significance. Our null hypothesis is that the GRYD IR interventions have no impact. That is GRYD IR does not reduce gang violence. If this null hypothesis is true then the ground truth values of the matrix K should be $k_{11} = k_{01}$ and $k_{10} = k_{00}$. Given estimates k_{11} and k_{01} and standard errors s_{11} and s_{01}, for example, the quantity

$$\frac{k_{11} - k_{01}}{\sqrt{s_{11}^2 + s_{01}^2}}$$

should be t-distributed and for large samples approximately standard normal (16). The p-value associated with the magnitude of the observed difference between estimates of k_{11} and k_{01} can be computed directly against the cumulative distribution for the standard normal. We reject the null hypothesis based on standard probability criteria.

Estimating prevented crimes. The parameter values from the spatio-temporal model can be used to estimate the number of retaliations prevented by GRYD IR notifications. The sum (k_{11} + k_{01}) is the average number of retaliations known to both the LAPD and GRYD IR produced by the two types of triggers, LAPD + GRYD IR and LAPD-only. Similarly, (k_{10} + k_{00}) is the average number of retaliations known only to the LAPD produced by the two types of triggers. Note that (k_{11} + k_{01}) and (k_{10} + k_{00}) are actually measured directly from data and therefore are the observed outcome.

We then define two counterfactual situations. Let (k_{01} + k_{01}) be the average number of retaliations that *would have occurred* in the absence of GRYD IR notification for those retaliations known to LAPD + GRYD IR. Here we simply replace k_{11} with a second instance of k_{01}. Thus, we suppose that the LAPD + GRYD IR effect is replaced with the LAPD-only effect in the absence of GRYD IR notification. Similarly, let (k_{00} + k_{00}) be the average number of retaliations that *would have occurred* in the absence of GRYD IR notification for those retaliations known only to the LAPD. We then compute the relative effect of GRYD IR notification on the average number of retaliations as:

Relative effect of GRYD IR notification on LAPD + GRYD IR retaliations	Relative effect of GRYD IR notification on LAPD-only retaliations
$\dfrac{(k_{11}+k_{01})-(k_{01}+k_{01})}{(k_{01}+k_{01})}$	$\dfrac{(k_{10}+k_{00})-(k_{00}+k_{00})}{(k_{00}+k_{00})}$

We use the results of stochastic declustering to estimate the observed number of retaliations known only to the LAPD and those know both to the LAPD and GRYD IR (see Table 2 main text). The estimated number of prevented retaliations is the number of retaliations from stochastic declustering removing the relative effects of GRYD. City-wide data suggests that homicides make up on average 5.4% of all gang retaliations. We use this figure to estimate the number of prevented homicides from the total prevented retaliations.

Results

Nonparametric Model Estimates. The nonparametric estimates for the triggering pathways for each of our cases are substantially similar to the parametric model. However, it is not possible to estimate standard errors for the nonparametric model.

LAPD + GRYD IR vs. LAPD-only violent crime (gang and non-gang)	LAPD + GRYD IR vs. LAPD-only gang crime
$K_{u_i u} = \begin{pmatrix} 0.0674 & 0.3474 \\ 0.0550 & 0.3941 \end{pmatrix}$	$K_{u_i u} = \begin{pmatrix} 0.0640 & 0.2320 \\ 0.1187 & 0.3039 \end{pmatrix}$

Voronoi Residuals. Fig. 4 in the main text shows the results of analyses for a Poisson process model compared to the Hawkes process model. Voronoi residuals from the fitted intensity for the proposed Hawkes model has muted colors as compared to the null model, indicating improved performance. Fig. S3 compares residual values in histogram form. Voronoi residuals for the Poisson model are spread over a wide range with a bias toward model underestimation of intensity (positive residuals). Voronoi residuals for the Hawkes model are more compact by comparison with a slight bias toward model overestimation of intensity (negative residuals).

Forward Simulation Parameter Recovery & Intervention Effects. Using the above forward simulation procedures, we produce synthetic data sets for three different kinds of *K* matrices:

Intervention has no effect: $K_{u_i u} = \begin{pmatrix} 0.10 & 0.30 \\ 0.10 & 0.30 \end{pmatrix}$

Intervention reduces retaliation: $K_{u_i u} = \begin{pmatrix} 0.05 & 0.15 \\ 0.10 & 0.30 \end{pmatrix}$

Intervention increases retaliation: $K_{u_i u} = \begin{pmatrix} 0.20 & 0.45 \\ 0.10 & 0.30 \end{pmatrix}$

We then simulate the multivariate Hawkes process with additional parameter values $\sigma = 0.0082$, $\mu = [0.5, 0.7]$, $\omega = 1.5996$, $T = 1000$, and $(x, y) \in [0,1] \times [0,1]$. For each theoretical case of *K*, we simulate the point processes 100 times. Parameter estimates are not significantly different from the ground-truth values for both the parametric non-parametric cases. They are underestimates, however, due to boundary effects. Standard errors of parameter estimates are in parentheses.

	Parametric Estimate	**Nonparametric Estimate**
Intervention has no effect:	$K_{u_iu} = \begin{pmatrix} \underset{(0.0192)}{0.0964} & \underset{(0.0405)}{0.2854} \\ \underset{(0.0146)}{0.0944} & \underset{(0.0361)}{0.2882} \end{pmatrix}$	$K_{u_iu} = \begin{pmatrix} \underset{(0.0175)}{0.0923} & \underset{(0.0265)}{0.2854} \\ \underset{(0.0127)}{0.0949} & \underset{(0.0179)}{0.2877} \end{pmatrix}$
Intervention reduces retaliation:	$K_{u_iu} = \begin{pmatrix} \underset{(0.0140)}{0.0441} & \underset{(0.0262)}{0.1426} \\ \underset{(0.0156)}{0.0948} & \underset{(0.0360)}{0.2820} \end{pmatrix}$	$K_{u_iu} = \begin{pmatrix} \underset{(0.0157)}{0.0475} & \underset{(0.0201)}{0.1377} \\ \underset{(0.0119)}{0.0928} & \underset{(0.0221)}{0.2868} \end{pmatrix}$
Intervention increases retaliation:	$K_{u_iu} = \begin{pmatrix} \underset{(0.0280)}{0.1922} & \underset{(0.0553)}{0.4395} \\ \underset{(0.0156)}{0.0993} & \underset{(0.0347)}{0.2936} \end{pmatrix}$	$K_{u_iu} = \begin{pmatrix} \underset{(0.0201)}{0.1902} & \underset{(0.0317)}{0.4368} \\ \underset{(0.0119)}{0.0978} & \underset{(0.0183)}{0.2904} \end{pmatrix}$

Number of Prevented Crimes. We use these estimated effects along with the results of stochastic declustering in South Los Angeles to compute numbers of prevented crimes. Stochastic declustering identified a total of 45 LAPD + GRYD IR gang aggravated assaults and homicides in 2014-2015 as retaliatory (see Table 2 in the main text). The remaining 577 gang aggravated assaults and homicides were statistically defined as background events. Similarly, declustering identified a total of 403 LAPD-only gang aggravated assaults and homicides in 2014-2015 as retaliatory (see Table 2 in the main text). The remaining 877 gang aggravated assaults and homicides were statistically identified as background events. The counterfactual conditions suggest that retaliatory gang aggravated assaults and homicides would have been 48.8% and 15.0% higher in the absence of GRYD IR for events recorded, respectively, as LAPD + GRYD IR and LAPD-only. Thus GRYD IR prevented an estimated total 82.2 retaliatory gang aggravated assaults and homicides. An estimate based on city-wide data, suggests that homicides make up on average 5.4% of all gang retaliations. Thus, in South Los Angeles, the prevented crimes are expected to include 4.4 retaliatory homicides and 77.8 retaliatory aggravated assaults.

We refer to McCollister et al. (17) for estimates of the costs of crime. In their work the total cost of a single homicide to government, victims and suspects is approximately $8.98 million. The cost of a single aggravated assault is $240,000. We simply multiply these figures by the estimated number of prevent aggravated assaults and homicides, respectively. The estimated number of homicides prevented by GRYD IR may add up to savings between $39.4 million over two years. The savings from prevented gang aggravated assaults in South Los Angeles may amount to an additional $9.5 million over two years. The combined savings per year in South Los Angeles alone may amount to $49.0 million.

Table S1. Results of one sample runs tests for three different categories of gang crimes in South Los Angeles.

	Gang Aggravated Assault + Homicide	Gang Aggravated Assault	Gang Homicide
LAPD + GRYD IR (n1)	622	470	152
LAPD-only (n2)	1290	1249	41
Total (N)	1912	1719	193
Percent LAPD + GRYD IR	32.5%	27.3%	78.7%
Runs	852	672	65
Expected Mean	840.31	683.99	65.58
Expected SD	19.18	16.47	4.62
Z	0.609	0.73	0.125
two-tailed p-value	0.54	0.47	0.90

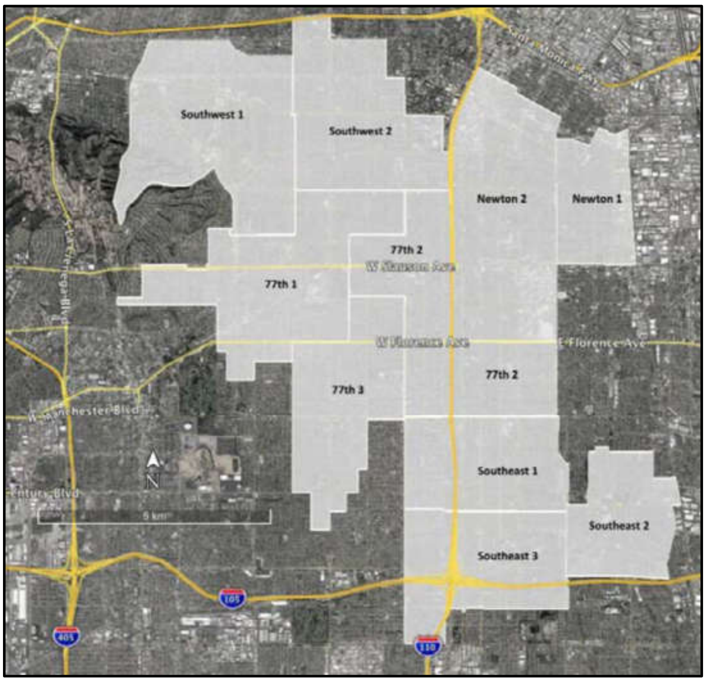


Fig. S1. Google Earth map of South Los Angeles showing the ten GRYD IR Zones in operation as of mid-2015.

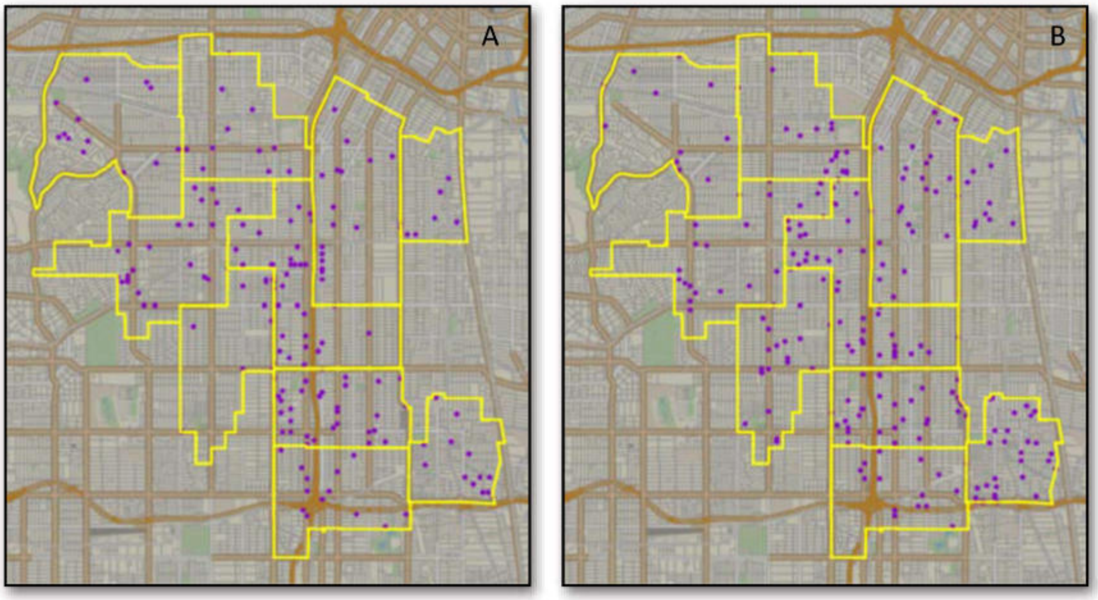


Fig. S2. Locations of GRYD IR Violent Crimes (purple) in a comparable six-month period before (A) and after (B) the July 2015 expansion of GRYD Zones in South Los Angeles. The time periods cover July-December 2014 (A) and July-December 2015 (B).

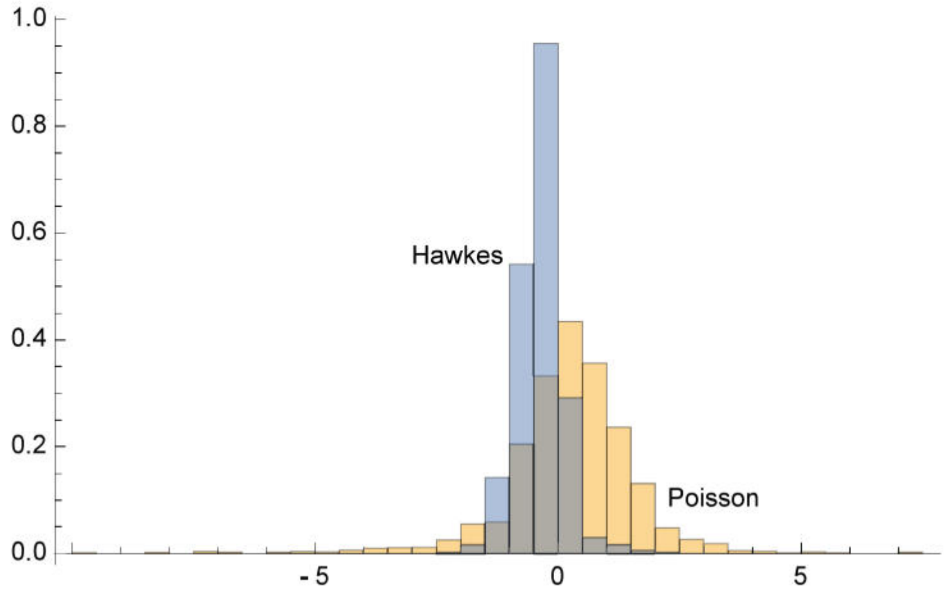


Fig. S3. Frequency histogram of Voronoi residuals for the fit Poisson (orange)l and Hawkes process (blue) models.

References Cited

1. Skogan WG, Hartnett SM, Bump N, & Dubois J (2009) *Evaluation of CeaseFire-Chicago* (U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, Washington, D.C.).

2. Whitehill JM, Webster DW, Frattaroli S, & Parker EM (2014) Interrupting Violence: How the CeaseFire Program Prevents Imminent Gun Violence through Conflict Mediation. *Journal of Urban Health* 91(1):84-95.
3. Webster DW, Whitehill JM, Vernick JS, & Curriero FC (2013) Effects of Baltimore's Safe Streets Program on Gun Violence: A Replication of Chicago's CeaseFire Program. *Journal of Urban Health* 90(1):27-40.
4. Klein MW (1971) *Street Gangs and Street Workers* (Prentice Hall, Englewood Cliffs).
5. Papachristos AV (2011) Too big to fail. *Criminology & Public Policy* 10(4):1053-1061.
6. Veen A & Schoenberg FP (2008) Estimation of Space-Time Branching Process Models in Seismology Using an EM-Type Algorithm. *Journal of the American Statistical Association* 103(482):614-624.
7. Mohler G (2014) Marked point process hotspot maps for homicide and gun crime prediction in Chicago. *International Journal of Forecasting* 30(3):491-497.
8. Mohler G, Short M, Brantingham PJ, Schoenberg F, & Tita G (2011) Self-exciting point process modeling of crime. *Journal of the American Statistical Association* 106(493):100-108.
9. Marsan D & Lengline O (2008) Extending earthquakes' reach through cascading. *Science* 319(5866):1076-1079.
10. Fox EW, Schoenberg FP, & Gordon JS (2016) Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric Hawkes point process models of earthquake occurrences. *The Annals of Applied Statistics* 10(3):1725-1756.
11. Zhuang J, Ogata Y, & Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association* 97(458):369-380.
12. Bray A, Wong K, Barr CD, & Schoenberg FP (2014) Voronoi cell based residual analysis of spatial point process models with applications to Southern California earthquake forecasts. *Annals of Applied Statistics* 8(4):2247-2267.
13. Gordon JS, Clements RA, Schoenberg FP, & Schorlemmer D (2015) Voronoi residuals and other residual analyses applied to CSEP earthquake forecasts. *Spatial Statistics* 14, Part B:133-150.
14. Zhuang J, Ogata Y, & Vere-Jones D (2004) Analyzing earthquake clustering features by using stochastic reconstruction. *Journal of Geophysical Research: Solid Earth* 109(B5):n/a-n/a.
15. Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, & Tita GE (2011) Self-exciting point process modeling of crime. *Journal of the American Statistical Association* 106(493):100-108.
16. Ogata Y (1978) The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics* 30(1):243-261.
17. McCollister KE, French MT, & Fang H (2010) The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and Alcohol Dependence* 108(1–2):98-109.

Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial

P. Jeffrey Brantingham

UCLA Department of Anthropology, Los Angeles, CA 90095, USA.

E-mail: branting@ucla.edu

Matthew Valasik

Louisiana State University Department of Sociology, Baton Rouge, LA 70803, USA.

George O. Mohler

Indiana University-Purdue University Indianapolis Department of Computer and Information Science, Indianapolis IN 46202, USA.

Summary.
A randomized controlled field trial of place-based predictive policing conducted in Los Angeles showed crime reductions when police used algorithmic forecasts compared to controls based on crime analysis best practice. The algorithmic forecasting methods tested did not make use of arrest data, nor did they include personal identifying information, general environmental characteristics or demographic measures as part of forecasting procedures. Rather, forecasts relied only on on the time and location of officially reported crimes. While the potential for bias may be low given these narrow data inputs, there is still reasonable concern that such methods encourage directed police patrols to target minority communities with discriminatory consequences for minority individuals. Here we test for such biases using arrest data from the Los Angeles predictive policing experiments. We find not only that the numbers of arrests declined or remained unchanged during predictive policing deployments, but also that there were no significant differences in arrests by racial-ethnic group between control and treatment conditions. Arrests also did not contribute significantly to the observed impact of predictive policing on crime.

Keywords: crime forecasting; civil liberties; police bias; regression discontinuity design

1. Introduction

Place-based predictive policing is based on two core ideas: (1) mathematical forecasting methods can be used to anticipate future crime risk in narrowly proscribed geographic areas; and (2) the delivery of police resources to those prediction locations disrupts the opportunity for crime (Mohler et al., 2011; Bowers et al., 2004). Randomized controlled experiments of predictive policing conducted in Los Angeles provided evidence that algorithmic methods not only predict two-times as much crime as existing best practice, but also double the amount of crime prevented (Mohler et al., 2015). While this treatment effect can be measured in the field, the specific mechanism by which predictive policing delivers greater crime reduction is not immediately obvious.

The prevailing view, derived from experiments in hot spot policing (Sherman et al., 1989; Braga and Bond, 2008), is that the presence of police in a given place removes opportunities for crime even without any direct contact with potential offenders (Sherman and Weisburd, 1995; Weisburd, 2008; Loughran et al., 2011). This general deterrent effect persists for some time after police have departed (Koper, 1995; Sherman and Weisburd, 1995), and appears to diffuse into nearby areas where the police were not concentrating their efforts (Clarke and Weisburd, 1994; Weisburd et al., 2006; Telep et al., 2014). General deterrence is not the only mechanism by which crime might be prevented by police patrol, however. Direct interference via stops, searches, detentions short of arrest, and arrest, may prevent crime by physically incapacitating potential offenders (Sherman and Eck, 2002; Weisburd and Eck, 2004). This use of selective incapacitation may have immediate effects on crime (Wyant et al., 2012), especially if prolific offenders are the ones being arrested. Incapacitation may have longer term effects if those prolific offenders are subsequently removed from the community.

Considerable evidence, reviewed below, suggests that explicit and implicit bias can have a major impact on *who* gets stopped, searched and detained. Reasonable concern therefore exists that predictive policing can exacerbate such biases and reinforce any tendency for police to target minority individuals and communities. Such concern exists even if the forecasting methods used to drive predictive policing refrain from incorporating data that would be an explicit source of bias. If predictive policing indirectly exacerbates bias, any crime control benefits would need to be weighed in terms of their discriminatory costs. In the worst case, documented benefits might be derived solely from bias in the predictions. In other words, predictions absent such bias would yield no crime control benefits at all. Here we seek to evaluate whether predictive policing leads to pattens of arrest biased against minority individuals.

2. Bias in Police Patrol

Observed racial disparities in police law enforcement strategies are as polarizing today as at any point over the last 20 years (Beckett, 2012; Engel et al., 2012). Research has demonstrated that a racial bias exists in the business of policing including the racial profiling of vehicles (Baumgartner et al., 2016; Epp et al., 2014; Farrell and McDevitt, 2006; Harris, 1999; Horrace and Rohlin, 2016; Meehan and Ponder, 2002; Novak, 2004; Ridgeway, 2006; Smith and Petrocelli, 2001; Warren et al., 2006), pedestrian stops (Fagan et al., 2015, 2010; Gelman et al., 2007; Goel et al., 2016; Harris, 1994; Rios, 2011; Stuart, 2016), traffic tickets (Dunn, 2009), drug enforcement and arrests (Beckett et al., 2006; Lynch et al., 2013), use of force (Nix et al., 2017; Buehler, 2017; Legewie, 2016; Schuck, 2004), and even in the decision to shoot white or black criminal suspects while in a training simulator (Geller and Toch, 1995; Plant and Peruche, 2005). While the mechanisms driving these observed patterns of racial disparity (i.e., racial profiling, stereotyping/cognitive bias, deployment, racial animus/prejudice) remain difficult to disentangle, as Warren and colleagues Warren et al. (2006) attest, there is little doubt that racial disparities in police contacts with civilians do exist.

A prime example of racially disparate outcomes is the controversial use of stop, question, and frisk (SQF) by the New York Police Department (NYPD) (Fagan et al., 2010, 2015; Gelman et al., 2007; Laniyonu, 2017). The premise behind SQF is closely related to broken windows policing (Wilson and Kelling, 1982) in that more serious crimes may be prevented by stopping, searching and arresting suspicious individuals for minor criminal offenses (White and Fradella, 2016; Greene, 1999; Zimring, 2012). While some argue that racial disparities in SQF stops are an unintended consequence (MacDonald, 2002), Fagan et al. (2015) have shown a pattern of disproportionate impacts on people of color, controlling for population share and race-specific criminal offending. SQF directed at people of color does not yield increases in arrests, nor does it yield higher rates of detection of illegal weapons (Coviello and Persico, 2015). Any crime reductions associated with SQF last for only a few months (Rosenfeld and Fornango, 2017).

Research has also identified racial disparities in routine traffic enforcement practices (Baumgartner et al., 2016; Dunn, 2009; Epp et al., 2014; Farrell and McDevitt, 2006; Harris, 1999). People of color, particularly black motorists, are disproportionately more likely to be stopped by law enforcement than white motorists (Gaines, 2002; Lundman and Kaufman, 2003; Smith and Petrocelli, 2001). These studies highlight, however, a great deal of variation in racial disparity across jurisdictions. For instance, Warren et al. (2006) find that local police agencies in North Carolina are significantly more likely to target blacks motorists than the State Highway Patrol, who may be unable to identify the race of a motorist traveling at high speeds.

America's "war on drugs" has also shaped policing in urban centers. In contrast to traffic stops, drug law enforcement is much more proactive and more clearly indicates an agency's institutional policies and practices (Lynch et al., 2013). Research reveals that even in more progressive cities (e.g., San Francisco, Seattle), racially disparate drug law enforcement practices ensnare more people of color (Beckett et al., 2006; Beckett and Herbert, 2008; Rios, 2011). Lynch et al. (2013) reveal how race and place are not independent in the geography of drug law enforcement. Police tend to intercede in spaces that are contested (e.g., gentrifying neighborhoods, skid row tracts) as law enforcement attempts to maintain a municipality's economic and political interests (Stuart, 2016).

Given the empirical record there is genuine concern that place-based predictive policing may exacerbate racial biases (Ferguson, in press). The premise is that forecasts direct police officers into locations where they have an increased opportunity to exercise bias. Critically, such opportunities would not have been available, or only available in diminished frequency, in the absence of those forecasts. The implication is that biased outcomes should increase with the implementation of place-based predictive policing. We test several derived hypotheses using arrests recorded during the Los Angeles predictive policing experiment.

3. Predictive Policing Experiments in Los Angeles

A randomized controlled trial of predictive policing was conducted in three divisions of the Los Angeles Police Department (LAPD) between November 2011 and January 2013. The three participating divisions were Foothill, North Hollywood and Southwest. Only a brief outline of the experiment is present here. Details of the algorithmic procedures, experimental design and main effects are presented in Mohler et al. (2015).

Each day of the experiment police patrol officers were handed patrol maps with twenty target areas marked as 500 x 500 foot boxes. Officers were informed that the target areas were locations where the risk of crime was highest for their shift. They were encouraged to patrol target areas during any available discretionary time. What officers did not know was that the mission maps distributed to them each day were designed either by an algorithmic forecasting method (see Mohler et al., 2011, 2015), or by an analyst from within the division using all of the technological and intelligence assets at their disposal. Which mission map officers received on any given day was randomized creating a treatment condition (algorithmic forecast) and control condition (analyst forecast). In this repeated-measures experimental design, treatment days were considered exchangeable with control days (Mohler et al., 2015).

The outcome of interest was the difference in reported crime between control and treatment days. The crime types targeted were burglary, car theft and burglary theft from vehicle (BTFV). Historically, these crime types account for as much as 60% of the crime in the City of Los Angeles. In addition to this outcome measure, we collected information on the amount of time police officers spent in prediction areas under each of the experimental conditions (Mohler et al., 2015). Officers used their in-car computer terminals to register when they were entering and exiting prediction locations. This "dosage" was aggregated by day for a total amount of time (in minutes) spent in prediction locations.

Across the three test divisions, patrol officers using the algorithmic predictions produced an average 7.4% drop in crime as a function of patrol dosage. By contrast, use of the best-practice predictions failed yield a significant reduction in crime. The evidence presented in Mohler et al. (2015) is consistent with the conclusion that police patrol, when influenced by accurate predictions about the timing and location of crime, is effective at deterring crime. We reproduce the main regression results from Mohler et al. (2015) in Table 4.

We now turn to a consideration of potential biases induced by predictive policing. Specifically, we seek answers to the following empirical questions: (1) Were there differences in arrest rates for minority individuals in the test divisions before and after exposure? (2) Were there differences in arrest rates for minority individuals under control and treatment conditions during the testing? and (3) Did arrests contribute to observed changes in crime and were there differences by racial-ethnic group?

3.1. Empirical Data

3.1.1. Defining Pre- and Post-Exposure Periods

Predictive policing experiments in each of the three LAPD divisions started at different times and ran for different total durations. Pre- and post-exposure periods for each respective division overlap to a considerable degree (see Figure 1). The experiment in Foothill Division ran for 172 days from November 7, 2011 to April 27, 2012. The pre-exposure period was therefore chosen to be the 172 days prior to the start of the experiment, giving a pre-exposure start date of May 18, 2011. In North Hollywood, the experiment ran for 167 days from March 31, 2012 to Sept 14, 2012, giving a pre-exposure start date of Oct 15, 2011. In Southwest Division, the experiment ran for 239 days between May 16, 2012 to January 10, 2013, giving a s a pre-exposure start date of Sept 19, 2011.

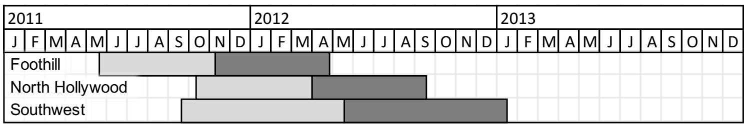


Fig. 1. The pre- and post-deployment exposure periods in Foothill, North Hollywood and Southwest Divisions of the LAPD.

3.1.2. Defining Control and Treatment Days

Control and treatment missions were designed independently, but in parallel each day of the experiment. Recall that treatment missions were based on algorithmic forecasting, while the control missions were based on existing best practice of analysts. Once mission designs were finalized, a control or treatment mission was chosen randomly for deployment. This randomization was done independently each day for each division taking part in the experiment. On occasion, the analyst was not present on a randomly designated control day and therefore control missions were not available for those days. We exclude treatment days from these days to ensure fair comparison. In Foothill Division, there were a total of 124 test days with successful random assignment, after discarding days on which the analyst was not present to design control missions. The 124 test days were evenly divided with 62 control and 62 treatment days. There were 152 total test days in North Hollywood Division. These included 82 control and 70 treatment days. In Southwest Division, there were 234 total days, including 117 control and 117 treatment days.

3.1.3. Defining Arrests

An arrest is generally understood to mean the taking into custody of an individual by the police given probable cause that a violation of the law has occurred. An arrest, as recorded by the LAPD, should not be conflated with other down-stream processes of the criminal justice system. An arrest does not imply booking, continued detention, nor whether those individuals are ultimately prosecuted for a crime. Arrests also should not be conflated with contacts between the public and police that did not result in arrest, even if such contacts were contentious. In general, police can exercise many alternatives to arrest in seeking to enforce laws and ensure order including behavioral directives, warnings and brief detention without arrest. On average, the LAPD makes about 1.5 million public contacts per year, but only about 24,000 of these contacts (1.6%) are arrests (Beck, 2016). Here arrests are taken at face value, without considering anything beyond the official record that an individual was taken into custody.

We do not distinguish between arrests for different types of crimes. In 2012, the LAPD made arrests under 520 different criminal codes representing 25 broad classes of crimes such as aggravated assault, robbery, burglary and larceny. Our primary focus is on whether the practice of policing introduces new biases into arrest patterns, not whether bias might be differentially present in arrests for different types of crimes.

3.1.4. Defining Racial-Ethnic Groups

The LAPD collects demographic information as part of the arrest process including age, sex and race-ethnicity of the individuals arrested. This information may be elicited from the individual or inferred by the arresting officer. The LAPD recognizes the categories Asian, black, Latino, white and other, which combined constitute 97.7% of all arrests on average. Occasionally, other categories such as Filipino, Korean, and Pacific Islander appear within the data. Given the sometimes fraught history between the LAPD and Latino and black communities (see Herbert, 1997; Martinez, 2016; Muiz, 2015) we focus on patterns in the arrest of black and Latino individuals and therefore report results for these two groups and for arrests overall.

4. Methods

We use three principal methods to address the questions outlined above. To examine whether arrests rates differed between pre- and post-exposure periods we use regression discontinuity design (RDD) (MacDonald et al., 2016; Imbens and Lemieux, 2008) and pooled proportions tests. RDD is chosen to test for short-range changes in arrest rates coinciding with the onset of predictive policing deployments. The assumption is that underlying environmental processes are similar over narrow time windows on either side of the boundary except for the introduction of predictive policing protocols. Underlying processes may be quite different over a larger time window in addition to the difference in experimental conditions. Pooled proportions tests are used to test for stable shifts in aggregate arrests between pre- and post-exposure periods. Pooled proportions tests are are also used to detect differences in arrests rates between control and treatment conditions. Finally, we use multiple regression to assess whether arrests play a role in observed impacts on crime under different experimental conditions. The multiple regression model forms are introduced along with the analysis.

Our null hypotheses are: (1) arrests rates for minority individuals do not differ between pre- and post-exposure periods; (2) arrests rates for minority individuals do not differ between control and treatment conditions during the experiment; and (3) arrest of minority individuals did not contribute to observed impacts on crime.

5. Results

5.1. Pre-Post Exposure to Predictive Policing

We first test for differences in the volume of arrests between pre- and post-exposure periods. Regression discontinuity models provide little indication that there was a shift in arrest volume accompanying the start of predictive policing experiments in any of the LAPD test divisions. In Foothill Division, the continuity in total arrests, black arrests, and Latino arrests is visibly apparent (Figure 2). Analyses of the local treatment effect confirm this visual impression (Table 1). In North Hollywood, there is continuity in black arrests, but a significant drop in Latino arrests at the boundary. By contrast, in Southwest, there is a significant increase in black arrests that coincides with the start of predictive policing, but continuity in Latino arrests. The magnitude and timing of each of these jumps is linked primarily to the overall volatility of arrests, rather than

Table 1. Regression discontinuity analysis of arrests.

Division & Race	Bandwidth (days)	# Observ. (N)	LATE Estimate (N crimes)	Std. Error	z value	$\Pr(>\lvert z\rvert)$
FH All	10.893	21	0.6094	4.303	0.1416	0.8874
FH Black	7.636	15	-0.425	0.6565	-0.6473	0.5174
FH Latino	11.317	24	-1.4365	2.443	-0.5881	0.5565
NH All	13.196	27	-11.23	9.729	-1.1545	0.2483
NH Black	8.007	17	-2.03742	2.298	-0.88663	0.37528
NH Latino	10.832	21	-12.986	5.686	-2.2838	0.022384*
SW All	12.047	25	17.72	4.288	4.133	3.587e-05***
SW Black	12.556	25	12.54	4.548	2.758	5.812e-03**
SW Latino	9.707	19	2.2033	3.849	0.5724	0.567

any persistent change in arrest patterns associated with predictive policing. Complete examination of all hypothetical cut points show there are numerous jumps of equal or greater magnitude. In North Hollywood Division, for example, six of the eight events with a drop in Latino arrests at least as large as that shown in Figure 2, occurred *before* the onset of the predictive policing experiments. In Southwest Division, 15 of the 23 events with an increase in black arrests at least as large as that observed occurred *before* the onset of predictive policing.

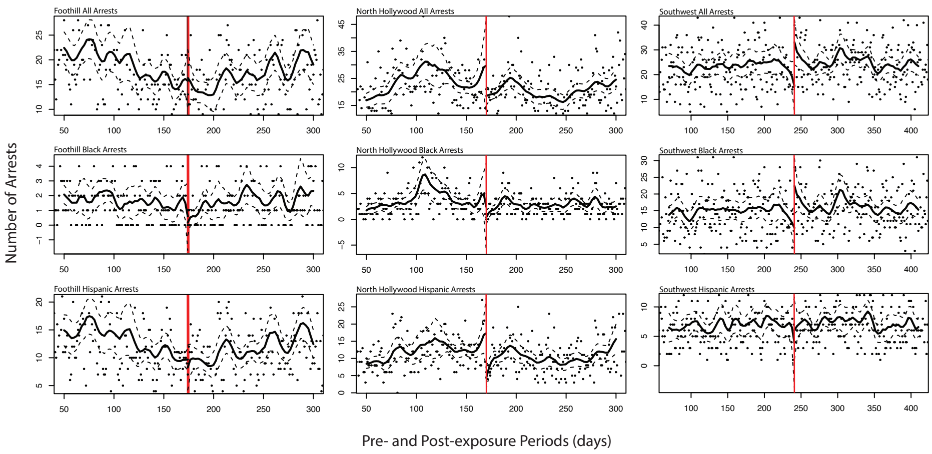


Fig. 2. Local polynomial smoothing of crime per day during predictive policing pre- and post-exposure periods in three divisions of the LAPD. Plots represent all (top row), black (middle row) and Latino (bottom row) arrests

That there was not a sharp change in arrests associated with the start of predictive policing does not exclude the possibility that there was a less pronounced, but longer-term shift towards greater numbers of arrests. We therefore test for changes in the proportions of arrests over the entire pre- and post-exposure periods (Table 2). Total

Table 2. Numbers of arrests and pool proportions tests for pre- and post-exposure periods.

	Total Arrests (N)			Black Arrests (N)				Latino Arrests (N)			
	pre	post	% change	pre	post	% change	p-val	pre	post	% change	p-val
FH	3493	2930	-16.1%	315	284	-9.8%	0.35	2413	1990	-17.5%	0.32
NH	6919	3551	-48.7%	1041	486	-53.3%	0.06	3446	1868	-45.8%	0.0067**
SW	5630	5730	1.8%	3626	3744	3.3%	0.30	1600	1685	5.3%	0.25

Table 3. Arrests on control and treatment days in three LAPD Divisions. † raw counts are lower due to the few treatment days compared to control.

Division	Race/ Ethnicity	Control Arrests (N)	Treatment Arrests (N)	Percent Control Total	Percent Treatment Total	Difference	p-val
FH	Black	117	118	11.0%	10.4%	-0.6%	0.68
FH	Latino	700	756	66.1%	66.8%	0.7%	0.36
FH	All	1059	1131	100.0%	100.0%		
NH	Black	241	203†	13.7%	13.8%	0.1%	0.46
NH	Latino	927	783†	52.7%	53.3%	0.6%	0.36
NH	All	1759	1468†	100.0%	100.0%		
SW	Black	1887	1818	65.4%	65.8%	0.4%	0.74
SW	Latino	843	811	29.2%	29.4%	0.1%	0.91
SW	All	2884	2761	100.0%	100.0%		

arrests declined in Foothill and North Hollywood, comparing the entire post exposure period to the entire pre-exposure period. Black arrests declined as a proportion of all arrests in Foothill, but the change was not statistically significant. Black and Latino arrests both declined significantly as a proportion of all arrests in North Hollywood. Total arrests increased in Southwest Division for the entire post-exposure period (Table 2). There were slight increases in both black and Latino arrests as a proportion of all arrests, but these increases were not statistically significant.

5.2. Control-Treatment Comparisons

The LAPD experiment was designed to test for differences in predictive accuracy and impact on crime between control and treatment conditions. Here we examine arrest patterns on control and treatment days (Table 3). Total arrests were slightly higher on treatment days compared to control days in Foothill Division. Total arrests were lower in North Hollywood and Southwest Divisions on treatment days compared to control. In North Hollywood, the absolute magnitude of the difference is due in large measure to the lower number of treatment days ($n = 70$) compared to control days ($n = 82$) in the experiment. Adjusted for this difference, the total number of arrests in North Hollywood would have been approximately 1717.3, an estimated decrease over the control number of arrests. In proportion to all arrests, black arrests were lower on treatment days in Foothill, but higher in North Hollywood and Southwest Divisions. Latino arrests were higher on treatment days in all three Divisions. However, none of the differences between control and treatment arrests were statistically significant (Table 3).

Table 4. Parameter estimates for impact of policing dosage on crime.

group	Parameter Estimate	Standard Error	t-value	p-value
Treatment μ_{FH}	6.793967	0.432887	15.684	2.0*10^{-}16
Treatment μ_{NH}	9.354687	0.492241	5.202	4.16*10^{-}7
Treatment μ_{SW}	9.117100	0.4468645	5.199	4.23*10^{-}7
Treatment β_1	-0.000994279	0.00042298	-2.303	0.0221
Control μ_{FH}	6.5950493	0.4641932	14.208	2.0*10^{-}16
Control μ_{NH}	9.1953424	0.5193115	5.007	1.03*10^{-}6
Control μ_{SW}	8.7943222	0.4861717	4.524	9.29*10^{-}6
Control β_1 Control	-0.0004664	0.0005127	-0.91	0.382399

5.3. *Impact of Arrests on Crime*

Finally, we address the question of whether arrests had a differential impact on crime under control and treatment conditions. Our baseline for comparison is the experiment presented in Mohler et al. (2015). In that work, regression models of the form $Y_{ij} = \mu_j + \beta_k X_{ij} + \epsilon_{ij}$ were used to assess the relationship between daily crime volume and patrol time on mission. The outcome variable, Y_{ij} is the daily crime volume on day i in division j. Recall that the target crime types for the experimental deployment were burglary, car theft and burglary-theft from motor vehicle. The independent variable X_{ij} is the cumulative police patrol time in minutes on day i across all active prediction boxes in division j. The parameter ϵ_{ij} is the uncorrelated error. The coefficient μ_j is an estimate of the mean crime volume per day in division j, in the absence of directed patrol. The coefficient β_k is an estimate of the impact of increasing patrol dosage under experimental condition k. The baseline results from Mohler et al. (2015) are reproduced in Table 4. The treatment condition (algorithmic forecast) produced a statistically significant decrease in the target crime types as a function of increasing policing dosage. By contrast, the control condition (best practice) did not yield a significant decrease in crime as a function of increasing dosage. The difference between treatment and control corresponded to a doubling of crime reduction for the same amount of police effort.

To address the impact of arrests we extended the above model in a straightforward manner to read $Y_{ij} = \mu_j + \beta_{1k} X_{ij} + \beta_{2k} Z_{ij} + \epsilon_{ij}$. The coefficient β_{1k} is an estimate of the impact of increasing patrol dosage on mean crime volume per day under experimental condition k, and is the counterpart of β_k from above. The coefficient β_{2k} is an estimate of the impact of increasing arrests on mean crime volume per day under experimental condition k. In all but one case, increasing arrests have no significant effect on the mean crime volume per day when controlling for policing dosage (Table 5). The exception concerns Latino arrests under treatment conditions, where the effect is marginally significant. Crime volume per day increased with the number of Latino arrests. This effect runs in the opposite direction from the impact of policing dosage, which remains a statistically significant source of crime reduction. In other words, the increase in Latino arrests not responsible for responsible for the observed decrease in crime under treatment conditions. A separate regression of Latino treatment arrests on treatment dosage shows that the two processes behave independently of one another (Table 6). Increasing policing dosage in treatment prediction areas did not increase in Latino arrests.

Table 5. Multiple regression for dosage and arrests under treatment and control conditions.

group	estimate	SE	t-value	p-value
Treatment + All Arrests				
μ_{FH}	6.50188	0.565575	11.4961	$9.95{*}10^{-25}$
μ_{NH}	9.15754	0.501334	5.29719	$2.63{*}10^{-7}$
μ_{SW}	8.85567	0.460604	5.11023	$6.5{*}10^{-7}$
β_1 Treatment Dosage	-0.00100063	0.000423606	-2.36217	0.0189544
β_2 Treatment Arrests	0.0129482	0.0202237	0.64025	0.522611
Control + All Arrests				
μ_{FH}	6.2382	0.624003	9.99706	$4.44{*}10^{-20}$
μ_{NH}	9.0241	0.527495	5.28137	$2.74{*}10^{-7}$
μ_{SW}	8.50478	0.51667	4.38691	$1.689{*}10^{-5}$
β_1 Control Dosage	-0.000453612	0.000509099	-0.891009	0.373761
β_2 Control Arrests	0.0115781	0.0249507	0.464037	0.643015
Treatment + Black Arrests				
μ_{FH}	6.80608	0.436548	15.5907	$1.71{*}10^{-38}$
μ_{NH}	9.5373	0.498403	5.47993	$1.06{*}10^{-7}$
μ_{SW}	9.76503	0.636963	4.6454	$5.55{*}10^{-6}$
β_1 Treatment Dosage	-0.000982505	0.000422704	-2.32434	0.0209293
β_2 Treatment Arrests	-0.0410679	0.0340836	-1.20492	0.229403
Control + Black Arrests				
μ_{FH}	6.384	0.469567	13.5955	$4.69{*}10^{-32}$
μ_{NH}	9.18987	0.517777	5.41908	$1.38{*}10^{-7}$
μ_{SW}	8.32581	0.76121	2.55096	0.0113253
β_1 Control Dosage	-0.000460084	0.000508949	-0.903988	0.366852
β_2 Control Arrests	0.0292476	0.0421066	0.694607	0.487931
Treatment + Latino Arrests				
μ_{FH}	5.8825	0.617093	9.53259	$1.63{*}10^{-18}$
μ_{NH}	8.64369	0.496342	5.56308	$6.95{*}10^{-8}$
μ_{SW}	8.67682	0.488814	5.71653	$3.16{*}10^{-8}$
β_1 Treatment Dosage	-0.000994498	0.000420657	-2.36416	0.0188553
β_2 Treatment Arrests	0.0698846	0.0362634	1.92714	0.0551233
Control + Latino Arrests				
μ_{FH}	6.37857	0.642941	9.92094	$7.71{*}10^{-20}$
μ_{NH}	9.21484	0.516391	5.49248	$9.55{*}10^{-8}$
μ_{SW}	8.74811	0.512742	4.62132	$6.04{*}10^{-6}$
β_1 Control Dosage	-0.000445898	0.000509063	-0.875919	0.381895
β_2 Control Arrests	0.00474212	0.0398681	0.118945	0.905412

Table 6. Regression of Latino treatment arrests on treatment dosage.

group	Parameter Estimate	Standard Error	t-value	p-value
μ_{FH}	12.1918	0.758458	16.0745	$3.47{*}10^{-40}$
μ_{NH}	11.18384	0.872066	-1.15583	0.248877
μ_{SW}	6.76718	0.788364	-6.88086	$4.94{*}10^{-11}$
Treatment Dosage	3.13543*10ˆ-6	0.0007411	0.00423078	0.996628

6. Discussion

The stated goal of the analyses present above was to assess the degree to which arrest rates were impacted by the introduction of predictive policing by the LAPD. Special attention was paid to arrest rates by the race-ethnicity of the individuals detained. Our null hypotheses were: (1) arrests did not change with the deployment of predictive policing; (2) arrests did not differ between control and treatment conditions; and (3) arrests did not contribute to observed crime reductions. The evidence presented does not allow us to reject these null hypotheses. Clearly, arrests are a common part of day-to-day police operations. However, the introduction of predictive policing did not induce biases sufficient to alter arrest patterns. Moreover, there is no evidence that observed crime declines were achieved through biased arrests.

The present study has several important limitations. First, it is possible that there was a temporal lag between experimental conditions and arrests. The lag could operate in two different ways. On the one hand, police patrol activities on day i might be responsible for arrests on day $i + l_1$, where $l_1 > 0$ is the lag measured in days. On the other, arrests on day i might impact on crime on day $i + l_2$, where again $l_2 > 0$ is the lag measured in days. The randomization of experimental conditions by day dilutes our ability to detect such lagged effects. To wit, because each day i was randomly assigned to control or treatment, day $i + l$ was also randomly assigned to control or treatment. Any lagged bias therefore should be equally distributed across experimental conditions. For example, assume that each day there is a baseline number of arrests a regardless of experimental condition, plus an additional bias b induced by the treatment. If the bias applies strictly to day $i + l$, then the expected number of arrests for *both* control and treatment days is $a + b/2$. The same logic applies if we believe that the effect of arrests on crime are lagged. These possibilities are very difficult to disentangle given the volatility of daily crime.

Second, arrests are an imperfect proxy for other types of police contacts including stops, searches and detentions short of arrest. It is possible that predictive policing induced increases in these other categories of police contacts, without a concomitant impact on arrests. For this to hold true it would have to be the case that the rate of arrest actually declined as these other precursor contacts increased, leaving overall arrest rates unchanged. This hypothetical downward adjustment in arrests would have to hold not only for the experimental deployment period overall, but also for randomly assigned treatment days. We do not have sufficient data to exclude such dynamics, but they seem improbable on the face of it.

Finally, the analyses do not provide any guidance on whether arrests are themselves systemically biased. Such could be the case, for example, if black and Latino individuals experienced arrest at a rate disproportionate to their share of the population and their share of offending (Rosenfeld and Fornango, 2014). The current study is only able to ascertain that arrest rates for black and Latino individuals were not impacted, positively or negatively, by using predictive policing. Future research could seek to test whether the situational conditions surrounding arrests and final dispositions differ in the presence of predictive policing.

References

Baumgartner, F. R., Epp, D. A., Shoub, K. and Love, B. (2016) Targeting young men of color for search and arrest during traffic stops: evidence from north carolina, 20022013. *Politics, Groups, and Identities*, 1–25.

Beck, C. (2016) Use of force year-end review executive summary 2015. *Report*, Los Angeles Police Department.

Beckett, K. (2012) Race, drugs, and law enforcement. *Criminology & Public Policy*, **11**, 641–653.

Beckett, K. and Herbert, S. (2008) Dealing with disorder: Social control in the post-industrial city. *Theoretical Criminology*, **12**, 5–30.

Beckett, K., Nyrop, K. and Pfingst, L. (2006) Race, drugs, and policing: Understanding disparities in drug delivery arrests. *Criminology*, **44**, 105–137.

Bowers, K. J., Johnson, S. D. and Pease, K. (2004) Prospective hot-spotting the future of crime mapping? *British Journal of Criminology*, **44**, 641–658.

Braga, A. A. and Bond, B. J. (2008) Policing crime and disorder hot spots: A randomized controlled trial*. *Criminology*, **46**, 577–607.

Buehler, J. W. (2017) Racial/ethnic disparities in the use of lethal force by us police, 20102014. *American journal of public health*, e1–e3.

Clarke, R. and Weisburd, D. (1994) Diffusion of crime control benefits: Observations on the reverse of displacement. *Crime Prevention Studies*, **2**, 165–184.

Coviello, D. and Persico, N. (2015) An economic analysis of black-white disparities in the new york police departments stop-and-frisk program. *The Journal of Legal Studies*, **44**, 315–360.

Dunn, R. A. (2009) Measuring racial disparities in traffic ticketing within large urban jurisdictions. *Public Performance & Management Review*, **32**, 537–561.

Engel, R. S., Smith, M. R. and Cullen, F. T. (2012) Race, place, and drug enforcement. *Criminology & Public Policy*, **11**, 603–635.

Epp, C. R., Maynard-Moody, S. and Haider-Markel, D. P. (2014) *Pulled over: How police stops define race and citizenship.* Chicago, IL: University of Chicago Press.

Fagan, J., Braga, A. A., Brunson, R. K. and Pattavina, A. (2015) Stops and stares: Street stops, surveillance and race in the new policing.

Fagan, J., Geller, A., Davies, G. and West, V. (2010) *Street Stops and Broken Windows Revisited: The Demography and Logic of Proactive Policing in a Safe and Changing City.* Race, Ethnicity and Policing. New York, NY: New York University.

Farrell, A. and McDevitt, D. J. (2006) Rhode island traffic stop statistics data collection study 2004-2005. *Report*, Institute on Race and Justice.

Ferguson, A. G. (in press) Policing predictive policing. *Washington University Law Review*, **94**.

Gaines, L. K. (2002) An analysis of traffic stop data in the city of riverside. *Report*.

Geller, W. A. and Toch, H. (1995) *And justice for all: Understanding and controlling police abuse of force*. Police Executive Research Forum Washington, DC.

Gelman, A., Fagan, J. and Kiss, A. (2007) An analysis of the new york city police department's stop-and-frisk policy in the context of claims of racial bias. *Journal of the American Statistical Association*, **102**, 813–823.

Goel, S., Rao, J. M. and Shroff, R. (2016) Precinct or prejudice? understanding racial disparities in new york citys stop-and-frisk policy. *The Annals of Applied Statistics*, **10**, 365–394.

Greene, J. A. (1999) Zero tolerance: A case study of police policies and practices in new york city. *Crime & Delinquency*, **45**, 171–187.

Harris, D. A. (1994) Frisking every suspect: The withering of terry. *UC Davis L. Rev.*, **28**, 1.

— (1999) The stories, the statistics, and the law: Why driving while black matters. *Minn. L. Rev.*, **84**, 265.

Herbert, S. K. (1997) *Policing space: Territoriality and the Los Angeles police department*. U of Minnesota Press.

Horrace, W. C. and Rohlin, S. M. (2016) How dark is dark? bright lights, big city, racial profiling. *Review of Economics and Statistics*, **98**, 226–232.

Imbens, G. W. and Lemieux, T. (2008) Regression discontinuity designs: A guide to practice. *Journal of econometrics*, **142**, 615–635.

Koper, C. S. (1995) Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, **12**, 649–672. URL: `http://dx.doi.org/10.1080/07418829500096231`.

Laniyonu, A. (2017) Coffee shops and street stops: Policing practices in gentrifying neighborhoods. *Urban Affairs Review*, 1078087416689728.

Legewie, J. (2016) Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. *American Journal of Sociology*, **122**, 379–424.

Loughran, T. A., Paternoster, R., Piquero, A. R. and Pogarsky, G. (2011) On ambiguity in perceptions of risk: Implications for criminal decision making and deterrence. *Criminology*, **49**, 1029–1061. URL: `http://dx.doi.org/10.1111/j.1745-9125.2011.00251.x`.

Lundman, R. J. and Kaufman, R. L. (2003) Driving while black: Effects of race, ethnicity, and gender on citizen selfreports of traffic stops and police actions. *Criminology*, **41**, 195–220.

Lynch, M., Omori, M., Roussell, A. and Valasik, M. (2013) Policing the progressivecity: The racialized geography of drug law enforcement. *Theoretical criminology*, **17**, 335–357.

MacDonald, H. (2002) *Are cops racist? How the war against the police harms Black Americans*. Ivan R. Dee.

MacDonald, J. M., Klick, J. and Grunwald, B. (2016) The effect of private police on crime: evidence from a geographic regression discontinuity design. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**, 831–846. URL: `http://dx.doi.org/10.1111/rssa.12142`.

Martinez, C. (2016) *The Neighborhood Has Its Own Rules: Latinos and African Americans in South Los Angeles*. NYU Press.

Meehan, A. J. and Ponder, M. C. (2002) Race and place: The ecology of racial profiling african american motorists. *Justice Quarterly*, **19**, 399–430.

Mohler, G., Short, M., Brantingham, P., Schoenberg, F. and Tita, G. (2011) Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, **106**, 100–108.

Mohler, G., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L. and Brantingham, P. J. (2015) Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, **110**, 1399–1411.

Muiz, A. (2015) *Police, Power, and the Production of Racial Boundaries*. Rutgers University Press.

Nix, J., Campbell, B. A., Byers, E. H. and Alpert, G. P. (2017) A bird's eye view of civilians killed by police in 2015. *Criminology & Public Policy*, **16**, 309–340. URL: `http://dx.doi.org/10.1111/1745-9133.12269`.

Novak, K. J. (2004) Disparity and racial profiling in traffic enforcement. *Police Quarterly*, **7**, 65–96.

Plant, E. A. and Peruche, B. M. (2005) The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, **16**, 180–183.

Ridgeway, G. (2006) Assessing the effect of race bias in post-traffic stop outcomes using propensity scores. *Journal of quantitative criminology*, **22**, 1–29.

Rios, V. M. (2011) *Punished: Policing the lives of Black and Latino boys*. NYU Press.

Rosenfeld, R. and Fornango, R. (2014) The impact of police stops on precinct robbery and burglary rates in new york city, 2003-2010. *Justice Quarterly*, **31**, 96–122. URL: `http://dx.doi.org/10.1080/07418825.2012.712152`.

— (2017) The relationship between crime and stop, question, and frisk rates in new york city neighborhoods. *Justice Quarterly*, 1–21.

Schuck, A. M. (2004) The masking of racial and ethnic disparity in police use of physical force: The effects of gender and custody status. *Journal of Criminal Justice*, **32**, 557–564.

Sherman, L. and Weisburd, D. (1995) General deterrent effects of police patrol in crime ?hot spots?: A randomized, controlled trial. *Justice Quarterly*, **12**, 625–648.

Sherman, L. W. and Eck, J. E. (2002) *Policing for crime prevention*, 295–397.

Sherman, L. W., Gartin, P. R. and Buerger, M. E. (1989) Hot spots of predatory crime: Routine activities and the criminology of place*. *Criminology*, **27**, 27–56.

Smith, M. R. and Petrocelli, M. (2001) Racial profiling? a multivariate analysis of police traffic stop data. *Police Quarterly*, **4**, 4–27.

Stuart, F. (2016) *Down, out, and under arrest: Policing and everyday life in Skid Row.* University of Chicago Press.

Telep, C., Weisburd, D., Gill, C., Vitter, Z. and Teichman, D. (2014) Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, **10**, 515–548. URL: `http://dx.doi.org/10.1007/s11292-014-9208-5`.

Warren, P., Tomaskovic-Devy, D., Smith, W., Zingraff, M. and Mason, M. (2006) Driving while black: Bias processes and racial disparity in police stops. *Criminology*, **44**, 709–738.

Weisburd, D. (2008) Place-based policing. *Ideas in American Policing*, **9**, 1–15.

Weisburd, D. and Eck, J. E. (2004) What can police do to reduce crime, disorder, and fear? *The ANNALS of the American Academy of Political and Social Science*, **593**, 42–65.

Weisburd, D., Wyckoff, L. A., Ready, J., Eck, J. E., Hinkle, J. C. and Gajewski, F. (2006) Does crime just move around the corner? a controlled study of spatial displacement and diffusion of crime control benefits*. *Criminology*, **44**, 549–592.

White, M. D. and Fradella, H. F. (2016) *Stop and frisk: The use and abuse of a controversial policing tactic.* NYU Press.

Wilson, J. Q. and Kelling, G. (1982) Broken windows: The police and neighborhood safety. *The Atlantic Monthly*, **127**, 29–38.

Wyant, B. R., Taylor, R. B., Ratcliffe, J. H. and Wood, J. (2012) Deterrence, firearm arrests, and subsequent shootings: A micro-level spatio-temporal analysis. *Justice Quarterly*, **29**, 524–545. URL: `<Go to ISI>://WOS:000307240700003`.

Zimring, F. E. (2012) *The city that became safe: New York's lessons for urban crime and its control.* Oxford University Press.

Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial

P. Jeffrey Brantingham

UCLA Department of Anthropology, Los Angeles, CA 90095, USA.

E-mail: branting@ucla.edu

Matthew Valasik

Louisiana State University Department of Sociology, Baton Rouge, LA 70803, USA.

George O. Mohler

Indiana University-Purdue University Indianapolis Department of Computer and Information Science, Indianapolis IN 46202, USA.

Summary.
A randomized controlled field trial of place-based predictive policing conducted in Los Angeles showed crime reductions when police used algorithmic forecasts compared to controls based on crime analysis best practice. The algorithmic forecasting methods tested did not make use of arrest data, nor did they include personal identifying information, general environmental characteristics or demographic measures as part of forecasting procedures. Rather, forecasts relied only on on the time and location of officially reported crimes. While the potential for bias may be low given these narrow data inputs, there is still reasonable concern that such methods encourage directed police patrols to target minority communities with discriminatory consequences for minority individuals. Here we test for such biases using arrest data from the Los Angeles predictive policing experiments. We find not only that the numbers of arrests declined or remained unchanged during predictive policing deployments, but also that there were no significant differences in arrests by racial-ethnic group between control and treatment conditions. Arrests also did not contribute significantly to the observed impact of predictive policing on crime.

Keywords: crime forecasting; civil liberties; police bias; regression discontinuity design

1. Introduction

Place-based predictive policing is based on two core ideas: (1) mathematical forecasting methods can be used to anticipate future crime risk in narrowly proscribed geographic areas; and (2) the delivery of police resources to those prediction locations disrupts the opportunity for crime (Mohler et al., 2011; Bowers et al., 2004). Randomized controlled experiments of predictive policing conducted in Los Angeles provided evidence that algorithmic methods not only predict two-times as much crime as existing best practice, but also double the amount of crime prevented (Mohler et al., 2015). While this treatment effect can be measured in the field, the specific mechanism by which predictive policing delivers greater crime reduction is not immediately obvious.

The prevailing view, derived from experiments in hot spot policing (Sherman et al., 1989; Braga and Bond, 2008), is that the presence of police in a given place removes opportunities for crime even without any direct contact with potential offenders (Sherman and Weisburd, 1995; Weisburd, 2008; Loughran et al., 2011). This general deterrent effect persists for some time after police have departed (Koper, 1995; Sherman and Weisburd, 1995), and appears to diffuse into nearby areas where the police were not concentrating their efforts (Clarke and Weisburd, 1994; Weisburd et al., 2006; Telep et al., 2014). General deterrence is not the only mechanism by which crime might be prevented by police patrol, however. Direct interference via stops, searches, detentions short of arrest, and arrest, may prevent crime by physically incapacitating potential offenders (Sherman and Eck, 2002; Weisburd and Eck, 2004). This use of selective incapacitation may have immediate effects on crime (Wyant et al., 2012), especially if prolific offenders are the ones being arrested. Incapacitation may have longer term effects if those prolific offenders are subsequently removed from the community.

Considerable evidence, reviewed below, suggests that explicit and implicit bias can have a major impact on *who* gets stopped, searched and detained. Reasonable concern therefore exists that predictive policing can exacerbate such biases and reinforce any tendency for police to target minority individuals and communities (Ferguson, in press). Such concern exists even if the forecasting methods used to drive predictive policing refrain from incorporating data that would be an explicit source of bias. If predictive policing indirectly exacerbates bias, any crime control benefits would need to be weighed in terms of their discriminatory costs. In the worst case, documented benefits might be derived solely from bias induced by predictions. In other words, predictions absent such bias would yield no crime control benefits at all. Here we seek to evaluate whether predictive policing leads to pattens of arrest biased against minority individuals.

2. Bias in Police Patrol

Observed racial disparities in police law enforcement strategies are as polarizing today as at any point over the last 20 years (Beckett, 2012; Engel et al., 2012). Research has demonstrated that a racial bias exists in the business of policing including the racial profiling of vehicles (Baumgartner et al., 2016; Epp et al., 2014; Farrell and McDevitt, 2006; Harris, 1999; Horrace and Rohlin, 2016; Meehan and Ponder, 2002; Novak, 2004; Ridgeway, 2006; Smith and Petrocelli, 2001; Warren et al., 2006), pedestrian stops (Fagan et al., 2015, 2010; Gelman et al., 2007; Goel et al., 2016; Harris, 1994; Rios, 2011; Stuart, 2016), traffic tickets (Dunn, 2009), drug enforcement and arrests (Beckett et al., 2006; Lynch et al., 2013), use of force (Nix et al., 2017; Buehler, 2017; Legewie, 2016; Schuck, 2004), and even in the decision to shoot white or black criminal suspects while in a training simulator (Geller and Toch, 1995; Plant and Peruche, 2005). While the mechanisms driving these observed patterns of racial disparity (i.e., racial profiling, stereotyping/cognitive bias, deployment, racial animus/prejudice) remain difficult to disentangle, as Warren et al. (2006) attest, there is little doubt that racial disparities in policing outcomes do exist.

A prime example of racially disparate outcomes is the controversial use of stop, question, and frisk (SQF) by the New York Police Department (NYPD) (Fagan et al., 2010, 2015; Gelman et al., 2007; Laniyonu, 2017). The premise behind SQF is closely related to "broken windows policing" (Wilson and Kelling, 1982), which holds that more serious crimes may be prevented by stopping, searching and arresting suspicious individuals for minor criminal offenses (White and Fradella, 2016; Greene, 1999; Zimring, 2012). While some argue that racial disparities in SQF stops are an unintended consequence (MacDonald, 2002), Fagan et al. (2015) have shown a pattern of disproportionate impacts on people of color, controlling for population share and race-specific criminal offending. SQF directed at people of color does not yield increases in arrests, nor does it yield higher rates of detection of illegal weapons (Coviello and Persico, 2015). Any crime reductions associated with SQF last for only a few months (Rosenfeld and Fornango, 2017).

Research has also identified racial disparities in routine traffic enforcement practices (Baumgartner et al., 2016; Dunn, 2009; Epp et al., 2014; Farrell and McDevitt, 2006; Harris, 1999). People of color, particularly black motorists, are disproportionately more likely to be stopped by law enforcement than white motorists (Gaines, 2002; Lundman and Kaufman, 2003; Smith and Petrocelli, 2001). These studies highlight, however, a great deal of variation in racial disparity across jurisdictions. For instance, Warren et al. (2006) find that local police agencies in North Carolina are significantly more likely to target blacks motorists than the State Highway Patrol, who may be unable to identify the race of a motorist traveling at high speeds.

America's "war on drugs" has also shaped policing in urban centers. In contrast to traffic stops, drug law enforcement is much more proactive and more clearly indicates an agency's institutional policies and practices (Lynch et al., 2013). Even in nominally more progressive cities (e.g., San Francisco, Seattle), racially disparate drug law enforcement practices ensnare more people of color (Beckett et al., 2006; Beckett and Herbert, 2008; Rios, 2011). Lynch et al. (2013) reveal how race and place are not independent in the geography of drug law enforcement. Police tend to intercede in spaces that are contested (e.g., gentrifying neighborhoods, skid row tracts) as law enforcement attempts to maintain a municipality's economic and political interests (Stuart, 2016).

Given the empirical record there is genuine worry that place-based predictive policing may exacerbate racial biases (Ferguson, in press). The reasoning is that forecasts direct police officers into locations where they have an increased opportunity to exercise bias. Such opportunities would not have been available, or only available in diminished frequency, in the absence of those forecasts. Moreover, knowing that one is in a prediction area is expected to heighten awareness in ways that amplify bias (Ferguson, 2012). Biased outcomes should therefore increase with the implementation of place-based predictive policing. We test several derived hypotheses using arrests recorded during the Los Angeles predictive policing experiment.

3. Predictive Policing Experiments in Los Angeles

A randomized controlled trial of predictive policing was conducted in three divisions of the Los Angeles Police Department (LAPD) between November 2011 and January 2013. The three participating divisions were Foothill (FH), North Hollywood (NH) and Southwest (SW). Only a brief outline of the experiment is present here. Details of the algorithmic procedures, experimental design and main effects are presented in Mohler et al. (2015).

Each day of the experiment police patrol officers were handed patrol maps with twenty target areas marked as 500 x 500 foot boxes. Officers were informed that the target areas were locations where the risk of crime was highest for their shift. They were encouraged to patrol target areas during any available discretionary time. What officers did not know was that the mission maps distributed to them each day were designed either by an algorithmic forecasting method (see Mohler et al., 2011, 2015), or by an analyst from within the division using all of the technological and intelligence assets at their disposal. Which mission map officers received on any given day was randomized creating a treatment condition (algorithmic forecast) and control condition (analyst forecast). In this repeated-measures experimental design, treatment days were considered exchangeable with control days (Mohler et al., 2015).

The outcome of interest was the difference in reported crime between control and treatment days. The crime types targeted were burglary, car theft and burglary theft from vehicle (BTFV). Historically, these crime types account for as much as 60% of the crime in the City of Los Angeles. In addition to this outcome measure, we collected information on the amount of time police officers spent in prediction areas under each of the experimental conditions (Mohler et al., 2015). Officers used their in-car computer terminals to register when they were entering and exiting prediction locations. This "dosage" was aggregated by day for a total amount of time (in minutes) spent in prediction locations.

Across the three test divisions, patrol officers using the algorithmic predictions produced an average 7.4% drop in crime as a function of patrol dosage. By contrast, use of the best-practice predictions failed yield a significant reduction in crime. The evidence presented in Mohler et al. (2015) is consistent with the conclusion that police patrol, when influenced by accurate predictions about the timing and location of crime, is effective at deterring crime. We reproduce the main regression results from Mohler et al. (2015) in Table 4.

We now turn to a consideration of potential biases induced by predictive policing. Specifically, we seek answers to the following empirical questions: (1) Were there differences in arrest rates for minority individuals in the test divisions before and after exposure? (2) Were there differences in arrest rates for minority individuals under control and treatment conditions during the testing? and (3) Did arrests contribute to observed changes in crime and were there differences by racial-ethnic group?

3.1. Empirical Data

3.1.1. Defining Pre- and Post-Exposure Periods

Predictive policing experiments in each of the three LAPD divisions started at different times and ran for different total durations. Pre- and post-exposure periods for each respective division overlap to a considerable degree (see Figure 1). The experiment in Foothill Division ran for 172 days from November 7, 2011 to April 27, 2012. The pre-exposure period was therefore chosen to be the 172 days prior to the start of the experiment, giving a pre-exposure start date of May 18, 2011. In North Hollywood, the experiment ran for 167 days from March 31, 2012 to Sept 14, 2012, giving a pre-exposure start date of Oct 15, 2011. In Southwest Division, the experiment ran for 239

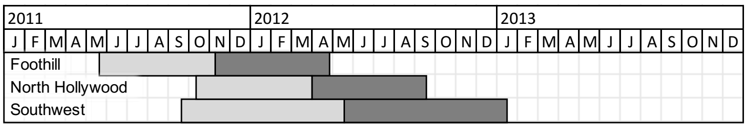


Fig. 1. The pre- and post-deployment exposure periods in Foothill, North Hollywood and Southwest Divisions of the LAPD.

days between May 16, 2012 to January 10, 2013, giving a s a pre-exposure start date of Sept 19, 2011.

3.1.2. Defining Control and Treatment Days

Control and treatment missions were designed independently, but in parallel each day of the experiment. Recall that treatment missions were based on algorithmic forecasting, while the control missions were based on existing best practice of analysts. Once mission designs were finalized, a control or treatment mission was chosen randomly for deployment. This randomization was done independently each day for each division taking part in the experiment. On occasion, the analyst was not present on a randomly designated control day and therefore control missions were not available for those days. We exclude treatment days from these days to ensure fair comparison. In Foothill Division, there were a total of 124 test days with successful random assignment, after discarding days on which the analyst was not present to design control missions. The 124 test days were evenly divided with 62 control and 62 treatment days. There were 152 total test days in North Hollywood Division. These included 82 control and 70 treatment days. In Southwest Division, there were 234 total days, including 117 control and 117 treatment days.

3.1.3. Defining Arrests

An arrest is generally understood to mean the taking into custody of an individual by the police given probable cause that a violation of the law has occurred. An arrest, as recorded by the LAPD, should not be conflated with other down-stream processes of the criminal justice system. An arrest does not imply booking, continued detention, nor whether those individuals are ultimately prosecuted for a crime. Arrests also should not be conflated with contacts between the public and police that did not result in arrest, even if such contacts were contentious. In general, police can exercise many alternatives to arrest in seeking to enforce laws and ensure order including behavioral directives, warnings and brief detention without arrest. On average, the LAPD makes about 1.5 million public contacts per year, but only about 24,000 of these contacts (1.6%) are arrests (Beck, 2016). Here arrests are taken at face value, without considering anything beyond the official record that an individual was taken into custody.

We do not distinguish between arrests for different types of crimes. In 2012, the LAPD made arrests under 520 different criminal codes representing 25 broad classes of crimes such as aggravated assault, robbery, burglary and larceny. Our primary focus is on whether the practice of policing introduces new biases into arrest patterns, not whether bias might be differentially present in arrests for different types of crimes.

3.1.4. Defining Racial-Ethnic Groups

The LAPD collects demographic information as part of the arrest process including age, sex and race-ethnicity of the individuals arrested. This information may be elicited from the individual or inferred by the arresting officer. The LAPD recognizes the categories Asian, black, Latino, white and other, which combined constitute 97.7% of all arrests on average. Occasionally, other categories such as Filipino, Korean, and Pacific Islander appear within the data. Given the sometimes fraught history between the communities of color (see Herbert, 1997; Martinez, 2016; Muiz, 2015) we focus on patterns in the arrest of black and Latino individuals and therefore report results for these two groups and for arrests overall.

4. Methods

We use three principal methods to address the questions outlined above. To examine whether arrests rates differed between pre- and post-exposure periods we use regression discontinuity design (RDD) (MacDonald et al., 2016; Imbens and Lemieux, 2008) and pooled proportions tests. RDD is chosen to test for short-range changes in arrest rates coinciding with the onset of predictive policing deployments. The assumption is that underlying environmental processes are similar over the narrow time windows on either side of the boundary except for the introduction of predictive policing protocols. Underlying processes may be quite different over a larger time window in addition to the difference in experimental conditions. Pooled proportions tests are used to test for stable shifts in aggregate arrests between pre- and post-exposure periods. Pooled proportions tests are are also used to detect differences in arrests rates between control and treatment conditions. Finally, we use multiple regression to assess whether arrests play a role in observed impacts on crime under different experimental conditions. The multiple regression model forms are introduced along with the analysis.

Our null hypotheses are: (1) arrests rates for minority individuals do not differ between pre- and post-exposure periods; (2) arrests rates for minority individuals do not differ between control and treatment conditions during the experiment; and (3) arrest of minority individuals did not contribute to observed impacts on crime.

5. Results

5.1. Pre-Post Exposure to Predictive Policing

We first test for differences in the volume of arrests between pre- and post-exposure periods. Regression discontinuity models provide little indication that there was a shift in arrest volume accompanying the start of predictive policing experiments in any of the LAPD test divisions. In Foothill Division, the continuity in total arrests, black arrests, and Latino arrests is visibly apparent (Figure 2). Analyses of the local treatment effect confirm this visual impression (Table 1). In North Hollywood, there is continuity in black arrests, but a significant drop in Latino arrests at the boundary. By contrast, in Southwest, there is a significant increase in black arrests that coincides with the start of predictive policing, but continuity in Latino arrests. The magnitude and timing of each of these observed jumps is linked to the overall volatility of arrests, rather than any persistent change in arrest patterns associated with predictive policing. Complete examination of all hypothetical cut points across the time series reveals numerous jumps of equal or greater magnitude. In North Hollywood Division, for example, six of the eight events with a decrease in Latino arrests at least as large as that shown in Figure 2, occurred *before* the onset of the predictive policing experiments. In Southwest Division, 15 of the 23 events with an increase in black arrests at least as large as that observed occurred *before* the onset of predictive policing. If anything, there is a reduction in the volatility of arrests following the deployment of predictive policing.

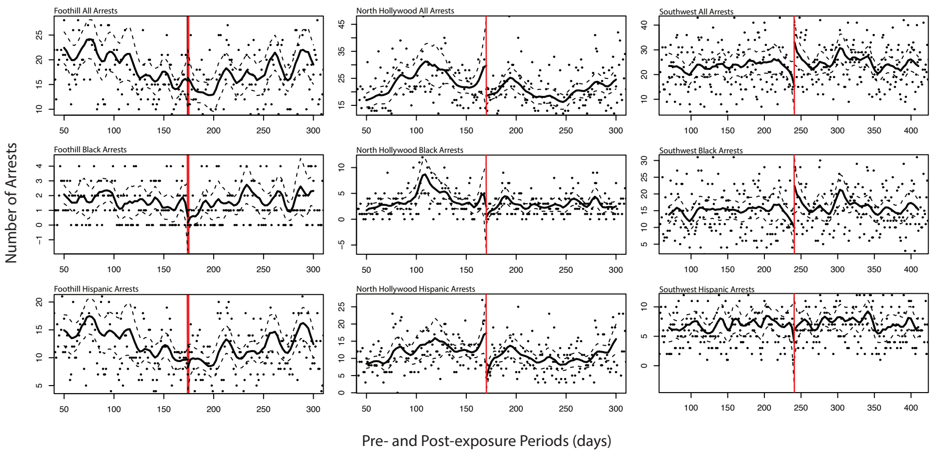


Fig. 2. Local polynomial smoothing of crime per day during predictive policing pre- and post-exposure periods in three divisions of the LAPD. Plots represent all (top row), black (middle row) and Latino (bottom row) arrests. The deployment cut points are shown as a red line.

That there was not a sharp change in arrests at the start of predictive policing deployments does not exclude the possibility that there was a less pronounced, but longer-term shift towards greater numbers of arrests. We therefore test for changes in the proportions of arrests over the entire pre- and post-exposure periods (Table 2). Total arrests declined in Foothill and North Hollywood, comparing the entire post exposure period to the entire pre-exposure period. Black arrests declined as a proportion of all arrests in Foothill, but the change was not statistically significant. Black and Latino arrests both declined significantly as a proportion of all arrests in North Hollywood. Total arrests increased in Southwest Division for the entire post-exposure period (Table 2). There were slight increases in both black and Latino arrests as a proportion of all arrests, but these increases were not statistically significant.

Table 1. Regression discontinuity analysis of arrests.

Division & Race	Bandwidth (days)	# Observ. (N)	LATE Estimate (N crimes)	Std. Error	z value	Pr(> $\lvert z \rvert$)
FH All	10.893	21	0.6094	4.303	0.1416	0.8874
FH Black	7.636	15	-0.425	0.6565	-0.6473	0.5174
FH Latino	11.317	24	-1.4365	2.443	-0.5881	0.5565
NH All	13.196	27	-11.23	9.729	-1.1545	0.2483
NH Black	8.007	17	-2.03742	2.298	-0.88663	0.37528
NH Latino	10.832	21	-12.986	5.686	-2.2838	0.022384*
SW All	12.047	25	17.72	4.288	4.133	3.587e-05***
SW Black	12.556	25	12.54	4.548	2.758	5.812e-03**
SW Latino	9.707	19	2.2033	3.849	0.5724	0.567

Table 2. Numbers of arrests and pool proportions tests for pre- and post-exposure periods.

	Total Arrests (N)			Black Arrests (N)				Latino Arrests (N)			
	pre	post	% change	pre	post	% change	p-val	pre	post	% change	p-val
FH	3493	2930	-16.1%	315	284	-9.8%	0.35	2413	1990	-17.5%	0.32
NH	6919	3551	-48.7%	1041	486	-53.3%	0.06	3446	1868	-45.8%	0.0067**
SW	5630	5730	1.8%	3626	3744	3.3%	0.30	1600	1685	5.3%	0.25

5.2. Control-Treatment Comparisons

The LAPD experiment was designed to test for differences in predictive accuracy and impact on crime between control and treatment conditions. Here we examine arrest patterns on control and treatment days (Table 3). Total arrests were slightly higher on treatment days compared to control days in Foothill Division. Total arrests were lower in North Hollywood and Southwest Divisions on treatment days compared to control. In North Hollywood, the absolute magnitude of the difference is due in large measure to the lower number of treatment days ($n = 70$) compared to control days ($n = 82$) in the experiment. Adjusted for this difference, the total number of arrests in North Hollywood would have been approximately 1717.3, an estimated decrease over the control number of arrests. In proportion to all arrests, black arrests were lower on treatment days in Foothill, but higher in North Hollywood and Southwest Divisions. Latino arrests were higher on treatment days in all three Divisions. However, none of the differences between control and treatment arrests were statistically significant (Table 3).

5.3. Impact of Arrests on Crime

Finally, we address the question of whether arrests had a differential impact on crime under control and treatment conditions. Our baseline for comparison is the analysis presented in Mohler et al. (2015). In that work, regression models of the form $Y_{ij} = \mu_j + \beta_k X_{ij} + \epsilon_{ij}$ were used to assess the relationship between daily crime volume and patrol time on mission. The outcome variable, Y_{ij} is the daily crime volume on day i in division j. Recall that the target crime types for the experimental deployment were burglary, car theft and burglary-theft from motor vehicle. The independent variable

Table 3. Arrests on control and treatment days in three LAPD Divisions. † raw counts are lower due to the few treatment days compared to control.

Division	Race/ Ethnicity	Control Arrests (N)	Treatment Arrests (N)	Percent Control Total	Percent Treatment Total	Difference	p-val
FH	Black	117	118	11.0%	10.4%	-0.6%	0.68
FH	Latino	700	756	66.1%	66.8%	0.7%	0.36
FH	All	1059	1131	100.0%	100.0%		
NH	Black	241	203†	13.7%	13.8%	0.1%	0.46
NH	Latino	927	783†	52.7%	53.3%	0.6%	0.36
NH	All	1759	1468†	100.0%	100.0%		
SW	Black	1887	1818	65.4%	65.8%	0.4%	0.74
SW	Latino	843	811	29.2%	29.4%	0.1%	0.91
SW	All	2884	2761	100.0%	100.0%		

Table 4. Parameter estimates for impact of policing dosage on crime.

group	Parameter Estimate	Standard Error	t-value	p-value
Treatment μ_{FH}	6.793967	0.432887	15.684	2.0*10$^-$16
Treatment μ_{NH}	9.354687	0.492241	5.202	4.16*10$^-$7
Treatment μ_{SW}	9.117100	0.4468645	5.199	4.23*10$^-$7
Treatment β_1	-0.000994279	0.00042298	-2.303	0.0221
Control μ_{FH}	6.5950493	0.4641932	14.208	2.0*10$^-$16
Control μ_{NH}	9.1953424	0.5193115	5.007	1.03*10$^-$6
Control μ_{SW}	8.7943222	0.4861717	4.524	9.29*10$^-$6
Control β_1 Control	-0.0004664	0.0005127	-0.91	0.382399

X_{ij} is the cumulative police patrol time in minutes on day i across all active prediction boxes in division j. The parameter ϵ_{ij} is the uncorrelated error. The coefficient μ_j is an estimate of the mean crime volume per day in division j, in the absence of directed patrol. The coefficient β_k is an estimate of the impact of increasing patrol dosage under experimental condition k. The baseline results from Mohler et al. (2015) are reproduced in Table 4. The treatment condition (algorithmic forecast) produced a statistically significant decrease in the target crime types as a function of increasing policing dosage. By contrast, the control condition (best practice) did not yield a significant decrease in crime as a function of increasing dosage. The difference between treatment and control corresponded to a doubling of crime reduction for the same amount of police effort.

To address the impact of arrests we extended the above model in a straightforward manner to read $Y_{ij} = \mu_j + \beta_{1k}X_{ij} + \beta_{2k}Z_{ij} + \epsilon_{ij}$. The coefficient β_{1k} is an estimate of the impact of increasing patrol dosage on mean crime volume per day under experimental condition k, and is the counterpart of β_k from above. The coefficient β_{2k} is an estimate of the impact of increasing arrests on mean crime volume per day under experimental condition k. In all but one case, increasing arrests have no significant effect on the mean crime volume per day when controlling for policing dosage (Table 5). The exception concerns Latino arrests under treatment conditions, where the effect is marginally significant. Crime volume per day increased with the number of Latino arrests. This effect runs in the opposite direction from the impact of policing dosage, which remains a statistically significant source of crime reduction. In other words, the increase in Latino arrests was not responsible for responsible for the observed decrease in crime under treatment conditions. A separate regression of Latino treatment arrests on treatment dosage shows that the two processes behave independently of one another (Table 6). Increasing policing dosage in treatment prediction areas did not increase in Latino arrests.

6. Discussion

The stated goal of the analyses present above was to assess the degree to which arrest rates were impacted by the introduction of predictive policing in three divisions patrolled by the LAPD. Special attention was paid to arrest rates by the race-ethnicity of the individuals detained. Our null hypotheses were: (1) arrests did not change with the deployment of predictive policing; (2) arrests did not differ between control and treatment conditions; and (3) arrests did not contribute to observed crime reductions. The evidence presented does not allow us to reject these null hypotheses. Clearly, arrests are a common part of day-to-day police operations. However, the introduction of predictive policing did not induce biases sufficient to alter arrest patterns. Moreover, there is no evidence that observed crime declines were achieved through biased arrests.

The present study has several important limitations. First, it is possible that there was a temporal lag between experimental conditions and arrests. The lag could operate in two different ways. On the one hand, police patrol activities on day i might be responsible for arrests on day $i + l_1$, where $l_1 > 0$ is the lag measured in days. On the other, arrests on day i might impact on crime on day $i + l_2$, where again $l_2 > 0$ is the lag measured in days. The randomization of experimental conditions by day dilutes our ability to detect such lagged effects. To wit, the experimental condition assigned on day i and day $i + l$ was random and independent. Any lagged bias therefore should be equally distributed across both experimental conditions. For example, assume that each day there is a baseline number of arrests a, plus an additional bias b induced by the treatment. If the bias applies strictly to day $i+l$, then the expected number of arrests is $a+b/2$ for *both* control and treatment days. The same logic applies if we believe that the effect of arrests on crime are lagged. These possibilities are very difficult to disentangle given the volatility of daily crime.

Second, arrests are an imperfect proxy for other types of police contacts including stops, searches and detentions short of arrest. It is possible that predictive policing induced increases in these other categories of police contacts, without a concomitant impact on arrests. For this to hold true it would have to be the case that the rate of arrest actually declined as these other precursor contacts increased, leaving overall arrest rates unchanged. This hypothetical downward adjustment in arrests would have to hold not only for the experimental deployment period overall, but also for randomly assigned treatment days. We do not have sufficient data to exclude such dynamics, but they seem improbable on the face of it.

Finally, the analyses do not provide any guidance on whether arrests are themselves systemically biased. Such could be the case, for example, if black and Latino individuals

Table 5. Multiple regression for dosage and arrests under treatment and control conditions.

group	estimate	SE	t-value	p-value
Treatment + All Arrests				
μ_{FH}	6.50188	0.565575	11.4961	$9.95*10^-25$
μ_{NH}	9.15754	0.501334	5.29719	$2.63*10^-7$
μ_{SW}	8.85567	0.460604	5.11023	$6.5*10^-7$
β_1 Treatment Dosage	-0.00100063	0.000423606	-2.36217	0.0189544
β_2 Treatment Arrests	0.0129482	0.0202237	0.64025	0.522611
Control + All Arrests				
μ_{FH}	6.2382	0.624003	9.99706	$4.44*10^-20$
μ_{NH}	9.0241	0.527495	5.28137	$2.74*10^-7$
μ_{SW}	8.50478	0.51667	4.38691	$1.689*10^-5$
β_1 Control Dosage	-0.000453612	0.000509099	-0.891009	0.373761
β_2 Control Arrests	0.0115781	0.0249507	0.464037	0.643015
Treatment + Black Arrests				
μ_{FH}	6.80608	0.436548	15.5907	$1.71*10^-38$
μ_{NH}	9.5373	0.498403	5.47993	$1.06*10^-7$
μ_{SW}	9.76503	0.636963	4.6454	$5.55*10^-6$
β_1 Treatment Dosage	-0.000982505	0.000422704	-2.32434	0.0209293
β_2 Treatment Arrests	-0.0410679	0.0340836	-1.20492	0.229403
Control + Black Arrests				
μ_{FH}	6.384	0.469567	13.5955	$4.69*10^-32$
μ_{NH}	9.18987	0.517777	5.41908	$1.38*10^-7$
μ_{SW}	8.32581	0.76121	2.55096	0.0113253
β_1 Control Dosage	-0.000460084	0.000508949	-0.903988	0.366852
β_2 Control Arrests	0.0292476	0.0421066	0.694607	0.487931
Treatment + Latino Arrests				
μ_{FH}	5.8825	0.617093	9.53259	$1.63*10^-18$
μ_{NH}	8.64369	0.496342	5.56308	$6.95*10^-8$
μ_{SW}	8.67682	0.488814	5.71653	$3.16*10^-8$
β_1 Treatment Dosage	-0.000994498	0.000420657	-2.36416	0.0188553
β_2 Treatment Arrests	0.0698846	0.0362634	1.92714	0.0551233
Control + Latino Arrests				
μ_{FH}	6.37857	0.642941	9.92094	$7.71*10^-20$
μ_{NH}	9.21484	0.516391	5.49248	$9.55*10^-8$
μ_{SW}	8.74811	0.512742	4.62132	$6.04*10^-6$
β_1 Control Dosage	-0.000445898	0.000509063	-0.875919	0.381895
β_2 Control Arrests	0.00474212	0.0398681	0.118945	0.905412

Table 6. Regression of Latino treatment arrests on treatment dosage.

group	Parameter Estimate	Standard Error	t-value	p-value
μ_{FH}	12.1918	0.758458	16.0745	$3.47*10^-40$
μ_{NH}	11.18384	0.872066	-1.15583	0.248877
μ_{SW}	6.76718	0.788364	-6.88086	$4.94*10^-11$
Treatment Dosage	3.13543*10ˆ-6	0.0007411	0.00423078	0.996628

experienced arrest at a rate disproportionate to their share of the population and their share of offending (Rosenfeld and Fornango, 2014). The current study is only able to ascertain that arrest rates for black and Latino individuals were not impacted, positively or negatively, by using predictive policing. Future research could seek to test whether the situational conditions surrounding arrests and final dispositions differ in the presence of predictive policing.

References

Baumgartner, F. R., Epp, D. A., Shoub, K. and Love, B. (2016) Targeting young men of color for search and arrest during traffic stops: evidence from north carolina, 20022013. *Politics, Groups, and Identities*, 1–25.

Beck, C. (2016) Use of force year-end review executive summary 2015. *Report*, Los Angeles Police Department.

Beckett, K. (2012) Race, drugs, and law enforcement. *Criminology & Public Policy*, **11**, 641–653.

Beckett, K. and Herbert, S. (2008) Dealing with disorder: Social control in the post-industrial city. *Theoretical Criminology*, **12**, 5–30.

Beckett, K., Nyrop, K. and Pfingst, L. (2006) Race, drugs, and policing: Understanding disparities in drug delivery arrests. *Criminology*, **44**, 105–137.

Bowers, K. J., Johnson, S. D. and Pease, K. (2004) Prospective hot-spotting the future of crime mapping? *British Journal of Criminology*, **44**, 641–658.

Braga, A. A. and Bond, B. J. (2008) Policing crime and disorder hot spots: A randomized controlled trial*. *Criminology*, **46**, 577–607.

Buehler, J. W. (2017) Racial/ethnic disparities in the use of lethal force by us police, 20102014. *American journal of public health*, e1–e3.

Clarke, R. and Weisburd, D. (1994) Diffusion of crime control benefits: Observations on the reverse of displacement. *Crime Prevention Studies*, **2**, 165–184.

Coviello, D. and Persico, N. (2015) An economic analysis of black-white disparities in the new york police departments stop-and-frisk program. *The Journal of Legal Studies*, **44**, 315–360.

Dunn, R. A. (2009) Measuring racial disparities in traffic ticketing within large urban jurisdictions. *Public Performance & Management Review*, **32**, 537–561.

Engel, R. S., Smith, M. R. and Cullen, F. T. (2012) Race, place, and drug enforcement. *Criminology & Public Policy*, **11**, 603–635.

Epp, C. R., Maynard-Moody, S. and Haider-Markel, D. P. (2014) *Pulled over: How police stops define race and citizenship*. Chicago, IL: University of Chicago Press.

Fagan, J., Braga, A. A., Brunson, R. K. and Pattavina, A. (2015) Stops and stares: Street stops, surveillance and race in the new policing.

Fagan, J., Geller, A., Davies, G. and West, V. (2010) *Street Stops and Broken Windows Revisited: The Demography and Logic of Proactive Policing in a Safe and Changing City.* Race, Ethnicity and Policing. New York, NY: New York University.

Farrell, A. and McDevitt, D. J. (2006) Rhode island traffic stop statistics data collection study 2004-2005. *Report*, Institute on Race and Justice.

Ferguson, A. G. (2012) Predictive policing and reasonable suspicion. *Emory LJ*, **62**, 259–324.

— (in press) Policing predictive policing. *Washington University Law Review*, **94**.

Gaines, L. K. (2002) An analysis of traffic stop data in the city of riverside. *Report*.

Geller, W. A. and Toch, H. (1995) *And justice for all: Understanding and controlling police abuse of force.* Police Executive Research Forum Washington, DC.

Gelman, A., Fagan, J. and Kiss, A. (2007) An analysis of the new york city police department's stop-and-frisk policy in the context of claims of racial bias. *Journal of the American Statistical Association*, **102**, 813–823.

Goel, S., Rao, J. M. and Shroff, R. (2016) Precinct or prejudice? understanding racial disparities in new york citys stop-and-frisk policy. *The Annals of Applied Statistics*, **10**, 365–394.

Greene, J. A. (1999) Zero tolerance: A case study of police policies and practices in new york city. *Crime & Delinquency*, **45**, 171–187.

Harris, D. A. (1994) Frisking every suspect: The withering of terry. *UC Davis L. Rev.*, **28**, 1.

— (1999) The stories, the statistics, and the law: Why driving while black matters. *Minn. L. Rev.*, **84**, 265.

Herbert, S. K. (1997) *Policing space: Territoriality and the Los Angeles police department.* U of Minnesota Press.

Horrace, W. C. and Rohlin, S. M. (2016) How dark is dark? bright lights, big city, racial profiling. *Review of Economics and Statistics*, **98**, 226–232.

Imbens, G. W. and Lemieux, T. (2008) Regression discontinuity designs: A guide to practice. *Journal of econometrics*, **142**, 615–635.

Koper, C. S. (1995) Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, **12**, 649–672. URL: `http://dx.doi.org/10.1080/07418829500096231`.

Laniyonu, A. (2017) Coffee shops and street stops: Policing practices in gentrifying neighborhoods. *Urban Affairs Review*, 1078087416689728.

Legewie, J. (2016) Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. *American Journal of Sociology*, **122**, 379–424.

Loughran, T. A., Paternoster, R., Piquero, A. R. and Pogarsky, G. (2011) On ambiguity in perceptions of risk: Implications for criminal decision making and deterrence. *Criminology*, **49**, 1029–1061. URL: `http://dx.doi.org/10.1111/j.1745-9125.2011.00251.x`.

Lundman, R. J. and Kaufman, R. L. (2003) Driving while black: Effects of race, ethnicity, and gender on citizen selfreports of traffic stops and police actions. *Criminology*, **41**, 195–220.

Lynch, M., Omori, M., Roussell, A. and Valasik, M. (2013) Policing the progressivecity: The racialized geography of drug law enforcement. *Theoretical criminology*, **17**, 335–357.

MacDonald, H. (2002) *Are cops racist? How the war against the police harms Black Americans*. Ivan R. Dee.

MacDonald, J. M., Klick, J. and Grunwald, B. (2016) The effect of private police on crime: evidence from a geographic regression discontinuity design. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**, 831–846. URL: `http://dx.doi.org/10.1111/rssa.12142`.

Martinez, C. (2016) *The Neighborhood Has Its Own Rules: Latinos and African Americans in South Los Angeles*. NYU Press.

Meehan, A. J. and Ponder, M. C. (2002) Race and place: The ecology of racial profiling african american motorists. *Justice Quarterly*, **19**, 399–430.

Mohler, G., Short, M., Brantingham, P., Schoenberg, F. and Tita, G. (2011) Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, **106**, 100–108.

Mohler, G., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L. and Brantingham, P. J. (2015) Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, **110**, 1399–1411.

Muiz, A. (2015) *Police, Power, and the Production of Racial Boundaries*. Rutgers University Press.

Nix, J., Campbell, B. A., Byers, E. H. and Alpert, G. P. (2017) A bird's eye view of civilians killed by police in 2015. *Criminology & Public Policy*, **16**, 309–340. URL: `http://dx.doi.org/10.1111/1745-9133.12269`.

Novak, K. J. (2004) Disparity and racial profiling in traffic enforcement. *Police Quarterly*, **7**, 65–96.

Plant, E. A. and Peruche, B. M. (2005) The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, **16**, 180–183.

Ridgeway, G. (2006) Assessing the effect of race bias in post-traffic stop outcomes using propensity scores. *Journal of quantitative criminology*, **22**, 1–29.

Rios, V. M. (2011) *Punished: Policing the lives of Black and Latino boys*. NYU Press.

Rosenfeld, R. and Fornango, R. (2014) The impact of police stops on precinct robbery and burglary rates in new york city, 2003-2010. *Justice Quarterly*, **31**, 96–122. URL: `http://dx.doi.org/10.1080/07418825.2012.712152`.

— (2017) The relationship between crime and stop, question, and frisk rates in new york city neighborhoods. *Justice Quarterly*, 1–21.

Schuck, A. M. (2004) The masking of racial and ethnic disparity in police use of physical force: The effects of gender and custody status. *Journal of Criminal Justice*, **32**, 557–564.

Sherman, L. and Weisburd, D. (1995) General deterrent effects of police patrol in crime ?hot spots?: A randomized, controlled trial. *Justice Quarterly*, **12**, 625–648.

Sherman, L. W. and Eck, J. E. (2002) *Policing for crime prevention*, 295–397.

Sherman, L. W., Gartin, P. R. and Buerger, M. E. (1989) Hot spots of predatory crime: Routine activities and the criminology of place*. *Criminology*, **27**, 27–56.

Smith, M. R. and Petrocelli, M. (2001) Racial profiling? a multivariate analysis of police traffic stop data. *Police Quarterly*, **4**, 4–27.

Stuart, F. (2016) *Down, out, and under arrest: Policing and everyday life in Skid Row*. University of Chicago Press.

Telep, C., Weisburd, D., Gill, C., Vitter, Z. and Teichman, D. (2014) Displacement of crime and diffusion of crime control benefits in large-scale geographic areas: a systematic review. *Journal of Experimental Criminology*, **10**, 515–548. URL: `http://dx.doi.org/10.1007/s11292-014-9208-5`.

Warren, P., Tomaskovic-Devy, D., Smith, W., Zingraff, M. and Mason, M. (2006) Driving while black: Bias processes and racial disparity in police stops. *Criminology*, **44**, 709–738.

Weisburd, D. (2008) Place-based policing. *Ideas in American Policing*, **9**, 1–15.

Weisburd, D. and Eck, J. E. (2004) What can police do to reduce crime, disorder, and fear? *The ANNALS of the American Academy of Political and Social Science*, **593**, 42–65.

Weisburd, D., Wyckoff, L. A., Ready, J., Eck, J. E., Hinkle, J. C. and Gajewski, F. (2006) Does crime just move around the corner? a controlled study of spatial displacement and diffusion of crime control benefits*. *Criminology*, **44**, 549–592.

White, M. D. and Fradella, H. F. (2016) *Stop and frisk: The use and abuse of a controversial policing tactic*. NYU Press.

Wilson, J. Q. and Kelling, G. (1982) Broken windows: The police and neighborhood safety. *The Atlantic Monthly*, **127**, 29–38.

Wyant, B. R., Taylor, R. B., Ratcliffe, J. H. and Wood, J. (2012) Deterrence, firearm arrests, and subsequent shootings: A micro-level spatio-temporal analysis. *Justice Quarterly*, **29**, 524–545. URL: `<Go to ISI>://WOS:000307240700003`.

Zimring, F. E. (2012) *The city that became safe: New York's lessons for urban crime and its control*. Oxford University Press.

RIPS 2017 Industry Sponsors Guide

Thank you for your interest in sponsoring a project in IPAM's Research in Industrial Projects for Students (RIPS) Program this summer! We have developed this short guide for sponsors, so that they know what to expect and understand the commitment they will be making. Please contact us (see page 6) with any questions or concerns about your participation.

Choosing an Industry Mentor

The industry mentor is a mathematician, scientist or engineer at the sponsoring organization who has the most familiarity with the proposed project. This will likely be the same person who writes the project description (below) but may be another researcher in the same area/office with similar expertise. A team of mentors is fine and even desirable, as long as one researcher is assigned to be the main point of contact for the students. The industry mentor must be available throughout the summer (June 19 – Aug. 18, 2017) to meet with the students in person or remotely at least weekly, as well as to respond to their questions by email. He or she must also attend opening day at IPAM (UCLA) on Monday, June 19. When choosing an industry mentor, please confirm that the person can make this commitment. If the industry mentor will be traveling during the summer, he or she must still be able to communicate with the students by email or phone, or must find a qualified substitute industry mentor for that period of time.

Industry mentors who do not live in LA area should plan to spend a few extra days at UCLA after Opening Day. Face-to-face interaction at the start of the program is important to foster a productive relationship with the students and a good experience overall. We require that the industry mentor resides in the United States, if not California; we have found that a time difference of more than a few hours is impractical, because the students cannot get timely responses to their questions.

Project Description

The project description explains the problem and the reason the outcome is important to the sponsor. The students should be able to appreciate the scope of the project and how mathematics is involved. Based on the project description and their own background research, the students will formulate a *statement of work* in the first week or two of the program. The project description should be provisional in nature; that is, not so specific that it proscribes the methodology, but not so general that the students are unable to identify specific objectives. It should list tasks you want the students to undertake, but should not insist on a specific methodology or outcome. A typical project description is 2 or 3 pages long and we strongly recommend that it includes elements we have identified in project descriptions from previous years that have been evaluated positively. The following suggestions emphasize the content and format we believe are most conducive to a successful project and a successful participant experience:

- Title of project and sponsor, and name of the industry mentor
- Introduction: A brief description of the problem and the motivation to work on the proposed project, including its relevance for the sponsor and its clients or society. Also, include a brief overview of the company –*i.e.*, history, type of business/operations, customers, etc.
- Technical Background: Provide a detailed description of the problem both in words and in a more abstract/technical formulation, *i.e.*, include relevant mathematical models, enunciate physical laws, or outline of algorithms if/when appropriate. Use graphics if you believe they'll help students understand the problem. Suggest the best possible approaches to solve the proposed problem.
  If more than one, a brief discussion of the strengths, weaknesses, or limitations of each approach may help students set realistic goals and plan their work most efficiently.
- Your Expectations: Set clear, reasonable, and attainable goals for the project. If possible, provide intermediate goals/milestones. That way, if students find difficulties reaching the ultimate goal of the project, they may still report results and conclusions, and provide future lines of work to reach the original goal. If the research objective proves to be impossible, a work statement may be re-negotiated by the team and sponsor.
- Recommended Reading: Please list two or three articles or chapters that the students may read for back- ground information. You may send one or two of them to your team in advance.
- Software Packages and Special Requirements: For any specific software tools you expect the students to use, provide references to manuals, user's guides, and, when available, online tutorials, including, if they exist, videos and webinars. Please also state if any data (or other inputs) will be provided or required.
- References: Provide a complete list of references.

We will need a draft project description by **February 14** (so that we may select and assign students to your team), and the final version by **June 1**.

Selection of Students

Along with the project description, please tell us if you have any **prerequisites or desired qualification** for your team along with the project description; for example, do you want students who have taken a particular course in math or science, or have experience with a specific programming language? We will try to accommodate your request. Additionally, please notify us as early as possible if your organization cannot work with students who aren't US citizens or permanent residents.

IPAM will receive more than 600 applicants for 36 spots. We will make offers to the most qualified students. Once most of the spots are filled, we will begin assigning teams. We will do our best to assign students to your team according to the guidelines you provide us, while also considering the interests of the students. Because we may have cancellations, we typically wait until May 1 to notify the students of their team assignment. We give them the project description about two weeks prior to the start of the program. At that point, we will encourage you to introduce yourself by email to your students, send them an article to read in preparation for opening day, and ask them about specific skills or background.

Technical, Security, and Legal Considerations

IPAM provides its students with dual-boot Windows 7 (or 10)/Ubuntu Mint Linux desktop computers and a variety of software, including a limited number of Matlab licenses and toolboxes, C and C++ compilers, LaTeX, Maple and Mathematica. If your students will need specialized hardware or software that we do not typically provide, please notify IPAM **by May 1, 2017** so that we can discuss the request and, if necessary, make appropriate purchases it in time for the start of RIPS. We welcome offers from the sponsor to provide or purchase these items for the team.

Additionally, if you will be providing propriety data or software to the students and will require them to sign a **nondisclosure agreement**, this must be reviewed by IPAM/UCLA prior to the start of RIPS. Please prepare and send us the draft agreement by May 1. As this process involves another UCLA office and may require several iterations, we need to allow sufficient time; we believe a May 1 deadline is early enough to have the document approved by Opening Day.

If your students will need **data** from you, please get permission from your company in advance so the students can proceed with their analysis immediately. If you will need IPAM to provide additional or specific **security** measures to protect your data, please be sure to bring this to our attention by May 1 as well. We may not be able to accommodate late requests. Finally, please notify your students of any rules concerning the use of data.

Finally, RIPS teams may be able to access UCLA's Hoffman Computational Cluster for extra data analysis power; however, as this resource is shared with other researchers there are some limitations on the types and amount of analysis that can be done. If your company has access to high-end computational facilities, please consider providing your students access to your own corporate computing systems. In either case, please let us know if you think you might have needs in this area.

During the Program

All industry mentors will meet with their team for a few hours on Opening Day. Those who are not local will ideally continue the meeting the following day, then continue to discuss the project by conference call or Skype regularly after that. Industry mentor who are local may meet with their students in person at IPAM or at the company's offices throughout the summer. It is fundamental to the success of the program that Industry Mentors provide students and academic mentors with technical support, and that they show involvement in the project and interest in the progress of their teams.

Your team will have an academic mentor (AM), typically a postdoc or junior faculty member, who will be in residence at IPAM for 20 hours per week. The AM will choose a Project Manager (PM) from among the four students. The industry mentor will communicate primarily with the team's PM throughout the summer. The industry mentor will not give specific instructions to the students or AM; rather, the students will seek to understand the problem and find for themselves a formulation of the problem and path for a solution, with the help of the AM. The PM's role includes leading team meetings, monitoring the team's progress, and delegating tasks as needed. Please remember that the program is an educational experience; in addition to conducting research, the students will also write a Statement of Work and Final Report, and give polished mid-term and final presentations. The standards for these program components are quite high, and will require a significant amount of their time.

The industry mentor will arrange for the students to have a "site visit" towards the end of the summer. (We suggest you schedule it during the seventh or ninth week of the program.) On the site visit, the students will present their research to an audience at the company made up of scientists and others interested in their work. We also recommend that you arrange for a tour and meeting with some or your scientists so the students can learn about other research that your organization sponsors.

We expect that the industry mentor will attend Projects Day if at all possible. If it is not possible, he or she can watch via live-stream video.

A Note on Sponsor Expectations

Please remember that RIPS is primarily an educational experience for undergraduates. For some students, this is their first opportunity to do research. The work statement, which the students prepare and present to the sponsor in the second week, helps to set expectations. If the research objective proves to be impossible, a work statement may be re-negotiated by the team and sponsor. We recommend that you suggest milestones to your team, so that if they do not reach the final goal, you can still walk away with some useful results. Throughout the summer, active dialogue between the industry mentor and the team is critical.

Industry Sponsor Fee and Other Expenses

IPAM will send each sponsor an invoice for the sponsorship fee. UCLA requests payment of the invoice within 45 days of receipt.

IPAM and the sponsor will share the cost of the students' site visit if it requires air travel and/or a hotel room. This may be negotiated in advance and included in the sponsorship fee, or the sponsor can reimburse the students directly for some of their expenses.

The industry sponsor may incur additional costs associated with the travel, parking, meals, and accommodations of the industry mentor(s) to attend opening day, projects day, and other activities at IPAM.

IPAM has a limited budget for software, hardware, and other expenses that are project-specific. If your team will require purchases, please discuss with IPAM early. We may ask the sponsor to help with the cost if it is significant.

RIPS 2016 TIMELINE

February 14, 2017	IPAM's deadline for students to apply. We will begin making offers shortly after this date.
February 14	Provide IPAM with the first draft of your project description and additional information as described under Selection of Students.
April 15-30	IPAM will put you in touch with the academic mentor for your team. The academic mentor may have valuable feedback on the project description that you can incorporate into the final version.
May 1	Notify IPAM of hardware, software and/or security requirements for your project, and present IPAM with an NDA, if one is needed.
June 1	Final version of the project description is due.
June 19	RIPS Opening Day
June 26-30	Team will present industry mentor with Work Statement.
July 15	Choose a date for your team's site visit and begin making arrangements.
July 17-21	Midterm presentation take place this week.
August 7-15	Site visits will likely take place during this time.
August 17	RIPS Projects Day

CONTACT INFORMATION

Stacey Beggs, Assistant Director
sbeggs@ipam.ucla.edu
310-794-0069

Russ Caflisch, Director
rcaflisch@ipam.ucla.edu
310-983-3297

James Kimmick, IT Manager
jkimmick@ipam.ucla.edu
310-794-7844

Dimi Mavalski, RIPS Program Coordinator
dmavalski@ipam.ucla.edu
310-794-7708

Tom Nykiel, Chief Financial Officer
tnykiel@ipam.ucla.edu
310-267-5247

American Sociological Review
1–32

DOI: 10.1177/0003122417725865
journals.sagepub.com/home/asr

Big Data Surveillance: The Case of Policing

Sarah Brayne[a]

Abstract

This article examines the intersection of two structural developments: the growth of surveillance and the rise of "big data." Drawing on observations and interviews conducted within the Los Angeles Police Department, I offer an empirical account of how the adoption of big data analytics does—and does not—transform police surveillance practices. I argue that the adoption of big data analytics facilitates amplifications of prior surveillance practices and fundamental transformations in surveillance activities. First, discretionary assessments of risk are supplemented and quantified using risk scores. Second, data are used for predictive, rather than reactive or explanatory, purposes. Third, the proliferation of automatic alert systems makes it possible to systematically surveil an unprecedentedly large number of people. Fourth, the threshold for inclusion in law enforcement databases is lower, now including individuals who have not had direct police contact. Fifth, previously separate data systems are merged, facilitating the spread of surveillance into a wide range of institutions. Based on these findings, I develop a theoretical model of big data surveillance that can be applied to institutional domains beyond the criminal justice system. Finally, I highlight the social consequences of big data surveillance for law and social inequality.

Keywords

police, big data, inequality, crime, law

In the past decade, two major structural developments intersected: the proliferation of surveillance in everyday life and the rise of "big data." Emblematic of the expansion of surveillance (Lyon 2003; Marx 2016; Rule 2007) is the rapid growth of the criminal justice system since 1972 (Carson 2015; Garland 2001; Wakefield and Uggen 2010; Western 2006). At the same time, facilitated by the mass digitization of information, there has been a rise in the computational analysis of massive and diverse datasets, known as "big data." Big data analytics have been taken up in a wide range of fields, including finance, health, social science, sports, marketing, security, and criminal justice. The use of big data in police surveillance activities is the subject of contentious debate in policy, media, legal, regulatory, and academic circles. However, discourse on the topic is largely speculative, focusing on the *possibilities*, good and bad, of new forms of data-based surveillance. The technological capacities for surveillance far outpace empirical research on the new data landscape. Consequently, we actually know very little about how big data is used in

[a]The University of Texas at Austin

Corresponding Author:
Sarah Brayne, Assistant Professor, Department of Sociology, The University of Texas at Austin, 305 E. 23rd St., A1700, Austin, TX 78712-1086
Email: sbrayne@utexas.edu

surveillance activities and to what consequence.

This article provides a case study of a large urban police department—the Los Angeles Police Department (LAPD)—to investigate the relationship between big data analytics and surveillance. In particular, it asks whether and how the adoption of big data analytics transforms police surveillance practices. Moreover, it investigates implications of new surveillance practices not only for policing, but also for law, social inequality, and research on big data surveillance in other institutions. On the one hand, big data analytics may be a rationalizing force, with potential to reduce bias, increase efficiency, and improve prediction accuracy. On the other hand, use of predictive analytics has the potential to technologically reify bias and deepen existing patterns of inequality.

To shed light on the social dimensions of surveillance in the age of big data, I draw on original interview and observational data collected during fieldwork over the course of two and a half years with the LAPD. As an agency at the forefront of data analytics, the Department serves as a strategic site for understanding the interplay between technology, law, and social relations. I provide one of the first on-the-ground accounts of how the growing suite of big data systems and predictive analytics are used for surveillance within an organization, unpacking how in some cases, the adoption of big data analytics is associated with mere *amplifications* in prior practices, but that in others, it is associated with fundamental *transformations* in surveillance activities. I argue there are five key ways in which the adoption of big data analytics is associated with shifts in practice to varying degrees: (1) discretionary assessments of risk are supplemented and quantified using risk scores; (2) data are increasingly used for predictive, rather than reactive or explanatory, purposes; (3) the proliferation of automated alerts makes it possible to systematically surveil an unprecedentedly large number of people; (4) datasets now include information on individuals who have not had any direct police contact; and (5) previously separate data systems are merged into relational systems and include data originally collected in other, non-criminal-justice institutions. These shifts made possible by big data have implications for inequality, law, and organizational practice in a range of institutional domains.

THE INTENSIFICATION OF SURVEILLANCE

Surveillance is ubiquitous in modern societies (Giddens 1990; Lyon 1994, 2003, 2006, 2015; Marx 1974, 2002, 2016; Rule 1974). The practice of surveillance involves the collection, recording, and classification of information about people, processes, and institutions (Foucault 1977; Haggerty and Ericson 2000; Lyon 2003; Marx 2016). A wide range of scholars highlight the growing pervasiveness of surveillance, referring to the emergence of "mass surveillance" (Rule 1974) and "surveillance societies" (Lyon 1994). Although it has received more attention in recent decades, surveillance as a practice is not new. Processes of surveillance can be traced back to at least the sixteenth century, during the appearance of the nation-state (Marx 2016), through the transatlantic slave trade (Browne 2015), bureaucratization, rationalization, and modern management in the nineteenth and twentieth centuries (Braverman 1974; Rule 1974; Weber 1978), and as an axiomatic accompaniment to risk management practices in the twentieth century (Ericson and Haggerty 1997). The attacks on September 11th, 2001, further stimulated and legitimated the expansion of surveillance. Widely viewed as a case of information sharing failure in the intelligence community, 9/11 encouraged an alignment of actors with vested interests in enhancing surveillance operations (Ball and Webster 2007; Lyon 2015), spurred an infusion of tax dollars for development of new surveillance sensors and data mining programs to produce strategic intelligence (Gandy 2002), and accelerated the convergence of previously separate surveillance systems into a "surveillant assemblage" (Deleuze and Guattari 1987; Haggerty and Ericson 2000).

Surveillance scholars have documented a quantitative increase in surveillance, arguing that surveillance is one of the major institutional dimensions of modern societies (Ball and Webster 2007; Giddens 1990; Lyon 1994, 2003; Marx 1988, 2016). Although surveillance is growing in all areas of society, its penetration is unevenly distributed (Fiske 1998). Some individuals, groups, areas, and institutions are surveilled more than others, and different populations are surveilled for different purposes (Lyon 2003). On the one hand, there is a deepening of surveillance of "at-risk" groups, such as parolees and individuals on public assistance (Gilliom 2001; Gustafson 2011; Soss, Fording, and Schram 2011), who can increasingly be tracked across institutional boundaries. On the other hand, emerging "dragnet" surveillance practices—meaning those that collect data on everyone, rather than merely individuals under suspicion—result in increased monitoring of groups "previously exempt from routine surveillance" (Haggerty and Ericson 2000:606; see also Angwin 2014; Lyon 2015). Surveillance is therefore now both *wider* and *deeper*: it includes a broader swath of people and can follow any single individual across a greater range of institutional settings.

Surveillance is increasingly technologically mediated, and emergent technologies make it possible at an unprecedented scale (Ericson and Haggerty 1997; Lyon 1994; Marx 2016). With the development of computing, mass surveillance emerged alongside mass communication (Rule 1974). The mass digitization of information enables much of what Marx terms the "new surveillance": the "scrutiny of individuals, groups, and contexts using technical means to extract or create information" (Marx 2016:20). Although the goals of traditional and new surveillance are similar, the means are different. Whereas traditional surveillance is inductive, involving the "close observation, especially of a suspected person" (*Oxford American Dictionary of Current English* 1999), and relying on the unaided senses, new surveillance is more likely to be applied categorically, deductive, remote, low visibility or invisible, involuntary, automated, preemptive, and embedded into routine activity (Marx 2002, 2016). In fact, new forms of systematic surveillance have become quotidian to the point that it is now an unavoidable feature of everyday life (Ball and Webster 2007; Lyon 2003; Marx 2016). Consider the extent to which surveillance is a requisite of participating in today's world (Ball and Webster 2007): to use a bank, send an e-mail, obtain medical care, make a phone call, travel on a highway, or conduct an Internet search, individuals leave digital traces that are recorded and saved (see also Foucault 1977; Haggerty and Ericson 2000; Lyon 2003).

Technologically mediated surveillance has become routine organizational practice in a wide range of public and private domains. It is a tool of general governance (Lyon 2003), a basic prerogative of individuals in all sorts of private and public institutions, and a tool to accomplish "goals that meet particular interests" (Ericson and Haggerty 2006:22). Surveillance scholars have accordingly extended the core concerns of surveillance from policing and military contexts to other institutions, including finance, commerce, labor, health, education, insurance, immigration, and activism (Lyon 2003; Marx 2016; Rule 2007). According to David Lyon (2015:68–69), surveillance in each of these institutions, "cannot be understood without a sense of how the quest for 'big data' approaches are becoming increasingly central."

RISE OF BIG DATA

Big data is an emerging modality of surveillance. A wide range of organizations—from finance to healthcare to law enforcement—have adopted big data analytics as a means to increase efficiency, improve prediction, and reduce bias (Christin 2016). Despite its take-up, big data remains an ambiguous term whose precise definition can vary across fields and institutional contexts. Drawing on previous definitions (e.g., Laney 2001; Lazer and Radford 2017; Mayer-Schönberger and

Cukier 2013), the working definition of big data used in this research is that it is a data *environment* characterized by four features: it is vast, fast, disparate, and digital. First, big data analytics involve the analysis of large amounts of information, often measured in petabytes and involving tens of millions of observations. Second, big data typically involves high frequency observations and fast data processing. Third, big data is disparate—it comes from a wide range of institutional sensors and involves the merging of previously separate data sources. Fourth, big data is digital. The mass digitization of records facilitates the merging and sharing of records across institutions, makes storage and processing easier, and makes data more efficient to analyze and search remotely. These four characteristics are not limited to any one institutional context, and they enable the use of advanced analytics—such as predictive algorithms or network analysis—and complex data display—such as topical-, temporal-, or geo-analysis. For purposes of sociological research on the topic, this definition shifts the focus from features of the data itself to the social processes that give rise to big data collection and analysis (i.e., the data environment). For example, instead of focusing on the "variety" of big data (one of the "3 Vs" in Laney's [2001] definition), the focus here is on the disparate institutional data sources big data is gathered from.

There are numerous theories for why such a wide range of institutions adopted big data surveillance as an organizational practice, most of which fit into one of two theoretical perspectives: the technical/rational perspective and the institutional perspective. Both perspectives are premised on the notion that organizations are self-interested (Scott 1987), but actors within organizations may adopt big data analytics in response to different pressures. According to the technical perspective, big data is a means by which organizational actors improve efficiency through improving prediction, filling analytic gaps, and more effectively allocating scarce resources. By contrast, the institutional perspective (DiMaggio and Powell 1983; Meyer and Rowan 1977) questions the assumption that organizational structures stem from rational processes or technical imperatives (Scott 2004). Instead, it highlights the role of culture, suggesting organizations operate in technically ambiguous fields in which they adopt big data analytics not because of empirical evidence that it *actually* improves efficiency, but in response to wider beliefs of what organizations *should* be doing with big data (Willis, Mastrofski, and Weisburd 2007; see also Kling 1991 on computerization). In other words, using big data may confer legitimacy. If other institutions are marshalling big data and algorithmic predictions for decision-making—rather than relying on human discretion—there may be institutional pressure to conform.

Big data makes possible new forms of classification and prediction using machine learning algorithms. Applications of big data analytics range from spam and fraud detection to credit scoring, insurance pricing, employment decisions, and predictive policing. Although much of the appeal of algorithms lies in their replacement of human decision-making with automated decisions, this study highlights the ways in which humans remain integral to the analytic process.

In the big data environment, individuals contribute to a growing trove of data as they go about their daily lives (Ball and Webster 2007; Garland 2001). Every time people make a purchase using a credit card, drive through a tollbooth, or click on an advertisement online, they leave a digital trace (Rule 2007). The adoption of digital information and communications technologies transformed previously paper files and face-to-face data collection (Marx 1998), making it possible for records "initially introduced with limited intentions" to be "developed, refined and expanded to deal with new problems and situations" (Innes 2001:8). Function creep—the tendency of data initially collected for one purpose to be used for another often unintended or unanticipated purpose (Innes 2001)—is a fundamental component of the big data surveillant landscape.

BIG DATA SURVEILLANCE: THE CASE OF POLICING

This article provides a case study of policing, one of the many organizational contexts in which the use of big data surveillance has grown. More generally, criminal justice surveillance has increased dramatically in the United States in the past four decades. It has expanded at all levels, including incarceration (Travis, Western, and Redburn 2014), parole and probation (Bonczar and Herberman 2014), and policing (Carson 2015). For example, the Violent Crime Control and Law Enforcement Act of 1994 provided funds to hire 100,000 new police officers, and the Homeland Security Act of 2002 committed over 17 billion dollars for state and local governments to fund local law enforcement agencies (Roush 2012). More recently, federal funds have been targeted at improving and expanding law enforcement's use of technology. For example, the Smart Policing Initiative—a consortium of the Bureau of Justice Assistance, local police departments, and researchers—provides federal funds to more than 30 local law enforcement agencies (including the LAPD) to support new data-driven practices.

The use of data for decision-making in criminal justice is not new. In 1928, Ernest Burgess of the Chicago School designed an actuarial model that predicted the probability of parolees' reoffending (Harcourt 2006). In the courts, quantification was embedded into legal practices in the 1970s and 1980s through sentencing guidelines (Espeland and Vannebo 2007). In the past three decades, the criminal justice system experienced a shift toward "actuarial justice" (Feeley and Simon 1992), in which actors use criteria derived from risk management (Lyon 2003) to estimate probabilities of criminal risk (Ericson and Haggerty 1997). That said, although actuarial methods have existed in corrections and the courts for almost a century (Feeley and Simon 1992; Harcourt 2006; Lyon 2003), data-driven decision-making has become systematically incorporated into law enforcement practices only in recent decades.

In the 1970s, the dominant police patrol model was reactive (Reiss 1971), involving random patrols, rapid responses to 911 calls, and reactive investigations (Sherman 2013). However, practitioners and researchers became increasingly aware that these strategies had little effect on crime, catalyzing a shift from reactive to more proactive, evidence-based forms of policing, such as hot spots policing (Braga and Weisburd 2010; Sherman, Gartin, and Buerger 1989). In 1994, CompStat—a management model linking crime and enforcement statistics—was established in New York City (Weisburd et al. 2003). CompStat quickly spread to other cities, including Los Angeles in 2002, as a managerial model for identifying crime patterns, quantifying and incentivizing police activity, and directing police resources. The attacks on 9/11 spurred the development of "intelligence-led policing" (Ratcliffe 2008). Viewing local law enforcement agencies as actors on the front lines of the domestic war against terror (Waxman 2009), federal agencies provided considerable funding to local law enforcement agencies to collect, analyze, share, and deploy a wide range of new data. In 2008, William Bratton, then-Chief of the LAPD (and former Commissioner of the New York City Police Department) began working with federal agencies to assess the viability of a more predictive approach to policing. Today, predictive analytics are used for a wide range of law enforcement–related activities, including algorithms predicting when and where future crimes are most likely to occur (Perry et al. 2013), network models predicting individuals most likely to be involved in gun violence (Papachristos, Hureau, and Braga 2013), and risk models identifying law enforcement officers most likely to engage in at-risk behavior (U.S. Department of Justice 2001 [2015]).

What explains the proliferation of big-data-driven decision-making in organizations generally, and law enforcement specifically? Much like in other institutional domains, it has the potential to improve both efficiency and accountability. It may improve the prediction and preemption of behaviors by helping law enforcement deploy resources more efficiently, ultimately helping prevent and intercept crimes, thus reducing crime rates. Data-driven policing also holds potential as an accountability mechanism and response to criticisms organizations are facing over discriminatory practices. For example, in response to police violence, nationwide movements such as Black Lives Matter have brought racial tensions to the forefront of demands for police reform. Data-driven policing is being offered as a partial antidote to racially discriminatory practices in police departments across the country (e.g., see White House Police Data Initiative 2015).

However, although part of the appeal of big data lies in its promise of less discretionary and more objective decision-making (see Porter's [1995] work on mechanical objectivity; see also Espeland and Vannebo 2007; Hacking 1990), new analytic platforms and techniques are deployed in preexisting organizational contexts (Barley 1986, 1996; Kling 1991) and embody the purposes of their creators (boyd and Crawford 2012; Gitelman 2013; Kitchin 2014). Therefore, it remains an open empirical question to what extent the adoption of advanced analytics will reduce organizational inefficiencies and inequalities, or serve to entrench power dynamics within organizations. The present study sheds light on these questions and helps us understand the changing relationship between quantification, prediction, and inequality.

LIMITATIONS TO EXISTING LITERATURE

There are five key limitations to the existing literature related to big data surveillance. First, most work on actuarialism was written before big data analytics took hold. Second, although there is strong theoretical work in surveillance studies, how big data surveillance plays out on the ground remains largely an open empirical question. Third, the majority of sociological research on criminal justice surveillance focuses on the experiences and outcomes of individuals under surveillance, rather than the surveilling agents themselves. Therefore, offering an organizational perspective may generate new insight about this mediating level of analysis. Fifth, although there is a strong body of work demonstrating that marking someone in the criminal justice system is consequential for life outcomes and patterns of inequality (Becker 1963; Brayne 2014; Kohler-Hausmann 2013; Pager 2007; Rios 2011), we know relatively little about whether and how the marking process has changed in the age of big data. Consequently, there is a dearth of theoretically informed empirical research on the relationship between surveillance, big data, and the social consequences of the intersection of the two forces.

Many of these gaps in the policing context can be attributed to practical constraints—it is difficult for researchers to secure the degree of access to police departments necessary to obtain in-depth qualitative data on day-to-day police practices. Although classic police ethnographies exist (e.g., Bittner 1967; Manning and Van Maanen 1978; Wilson 1968), there have been only a handful of in-depth studies within police departments since data analytics became an integral part of police operations (for early exceptions, see Ericson and Haggerty 1997; Manning 2011; Moskos 2008; Skogan 2006; Willis et al. 2007). Although these studies offer important insight into the use of CompStat and crime mapping, they predate algorithmic policing. Consequently, we still know little about how big data policing is exercised in practice.

This article has three aims. First, it draws on unique, original data to analyze how a law enforcement organization conducts big data surveillance. Second, it forwards an original theoretical framework for understanding the changes—and continuities—in surveillance practices associated with the adoption of big data analytics that can be applied to other institutional domains. Finally, it highlights implications of big data surveillance for law and social inequality.

FIELDWORK

Over the course of two and a half years, I conducted a qualitative case study of the Los Angeles Police Department (LAPD). The LAPD is the third-largest local law enforcement agency in the United States, employing 9,947 sworn officers and 2,947 civilian staff (Los Angeles Police Department 2017). The Department covers an area of almost 500 square miles and a population of almost four million people. It consists of four bureaus—Central, South, Valley, and West—which are divided into a total of 21 geographic areas. There are also two specialized bureaus, Detective and Special Operations.

I conducted interviews and observations with 75 individuals, and conducted between one and five follow-up interviews with a subsample of 31 individuals to follow how certain technologies were disseminated and information was shared throughout the Department. Interviewees included sworn officers of various ranks and civilian employees working in patrol, investigation, and crime analysis. I was able to gain analytic leverage by exploiting different adoption temporalities within and between divisions during my fieldwork. Not all divisions adopted big data surveillance at the same time. Rather, there was considerable variation in whether and when different big data technologies were adopted in the area and specialized divisions.[1] For example, I was able to conduct interviews and observations in divisions that were not using predictive policing and other big data surveillant technologies at the beginning of my fieldwork, but were by the end. This variation enabled me to observe actual changes in practice, but also to talk to respondents in patrol, investigation, and analysis roles about how they interpreted their work changing in light of big data analytics. I also interviewed individuals in specialized divisions—including Robbery-Homicide, Information Technology, Records and Identification, Fugitive Warrants, Juvenile, Risk Management, and Air Support—and at the Real-Time Crime Analysis Center (see Figure 1).

Additionally, I conducted observations on ride-alongs in patrol cars and a helicopter to study how officers deploy data in the field. I also shadowed analysts as they worked with data, observing them responding to queries from detectives and supervisors and proactively analyzing data for patrol, investigations, and crime analysis.

To supplement my research within the LAPD, I interviewed individuals within the L.A. County Sheriff's Department (LASD), as it is an integral part of the broader ecosystem of public services in the region. In addition, I conducted interviews at the Joint Regional Intelligence Center (JRIC), the "fusion center" in Southern California. Fusion centers are multiagency, multidisciplinary surveillance organizations by state or local agencies that received considerable federal funding from the Department of Homeland Security and the Department of Justice (Monahan and Palmer 2009). JRIC is one of 78 federally funded fusion centers established across the country in the wake of 9/11. Individuals at JRIC conduct data collection, aggregation, and surveillance in conjunction with other fusion centers and agencies, including, but not limited to, the Department of Homeland Security (DHS), the Federal Bureau of Investigations (FBI), the Central Intelligence Agency (CIA), and Immigration and Customs Enforcement (ICE). I also conducted observations at surveillance industry conferences and interviewed individuals working at technology companies that design analytic platforms used by the LAPD, including Palantir and PredPol, and individuals working in federal agencies in Washington, DC, to understand how data on criminal and noncriminal activity are shared across agencies. I supplemented my fieldwork with archival research of law enforcement and military training manuals and surveillance industry literature. Triangulating across various sources of data provided the analytic leverage necessary to better understand how law enforcement uses big data in theory, how they use it in practice, and how they interpret and make meaning out of its changing role in daily operations.

Figure 1. Situation Room at the Real-Time Crime Analysis Center (RACR)
Source: Author's photo.

Site Selection

I selected the LAPD as a strategic site for studying big data surveillance because it is an agency at the forefront of data analytics. The LAPD invests heavily in its data collection, analysis, and deployment capacities, and offers international training sessions on how law enforcement can better harness big data. Therefore, practices within the Department may forecast broader trends that may shape other law enforcement agencies in the coming years.

In addition to being one of the largest law enforcement agencies in North America, there are additional, contextual reasons why the LAPD is on the leading edge of data analytics. The first factor relates to external pressures for transparency and accountability. The LAPD was involved in a number of high-profile scandals in the 1990s, including the Rampart Scandal[2] and the now infamous Rodney King beating, which led to investigations exposing an expansive web of corruption, training deficiencies, and civil rights violations within the Department. In response, the Department of Justice entered into a consent decree[3] with the LAPD from 2001 to 2009 that mandated, among other things, the creation and oversight of a new data-driven employee risk management system, TEAMS II. The legacy of the decree extends beyond employee risk management; it led to more information sharing and data-driven decision-making within the organization in general.

The second factor is the influence of state legislative decisions concerning offender management. In the wake of *Brown v. Plata*[4] and the associated order to dramatically reduce the prison population, the California Legislature passed AB 109, a bill that shifted the responsibility of supervising released non-violent, non-serious, non-sex offenders from state to local law enforcement and county probation officers. It also outsourced compliance checks to local law enforcement agencies, including the LAPD and LASD. As a result, local law enforcement agencies were responsible for approximately 500 additional individuals released into L.A. County each month. Therefore, they needed a means by which to efficiently stratify the post-release community supervision population according to risk, necessitating risk modeling and interagency data integration efforts across the region.

A third relevant factor to the LAPD's use of big data is the availability and adoption of new data integration technologies. In 2011, the LAPD began using a platform designed by Palantir Technologies. Palantir was founded in 2004 and has quickly grown into one of the premier platforms for compiling and analyzing massive and disparate data by law enforcement and intelligence agencies. Originally intended for use in national defense, Palantir was initially partially funded by In-Q-Tel, the CIA's venture capital firm. Palantir now has government and commercial customers, including the CIA, FBI, ICE, LAPD, NYPD, NSA, DHS, and J.P. Morgan. JRIC (the Southern California fusion center) started using Palantir in 2009, with the LAPD following shortly after. The use of Palantir has expanded rapidly through the Department, with regular training sessions and more divisions signing on each year. It has also spread throughout the greater L.A. region: in 2014, Palantir won the Request for Proposals to implement the statewide AB 109 administration program, which involves data integration and monitoring of the post-release community supervision population.

CHANGES ASSOCIATED WITH ADOPTION OF BIG DATA ANALYTICS

To what extent does the adoption of big data analytics change police surveillance? Based on my fieldwork, I argue that in some cases, the adoption of big data analytics is associated with mere *amplifications* in prior surveillance practices, but in others, it is associated with fundamental *transformations* in surveillance activities and daily operations. I empirically demonstrate five key ways in which the adoption of big data analytics is associated with shifts in surveillance practices to varying degrees. First, law enforcement supplements officers' discretionary assessments of risk with quantified risk scores. Second, there is an increase in the use of data analytics for predictive—rather than reactive or explanatory—purposes. Third, there is a proliferation in alert-based systems, which facilitates the passive, systematic surveillance of a larger number of individuals than is possible with traditional query-based systems. Fourth, the threshold for inclusion in law enforcement databases is lower, now including individuals who have not had direct police contact. Finally, previously separate data systems are merged into relational systems, making it possible for the police to use data originally collected in other, non–criminal justice contexts.

I offer an original conceptual framework for understanding the continuities and changes associated with the adoption of big data analytics within the police organization. Figure 2 depicts this framework and illustrates the migration of traditional police practices toward big data surveillance. Each line represents a continuum of surveillance practices, from traditional to big data surveillance. The five shifts in practice do not represent discrete either/or categories, but rather are better understood as continuous gradations of varying degrees between the extreme values of traditional and big data surveillance. The length of the black lines represents the degree of transformation in surveillance practices associated with the use of big data. For example, the first two shifts—from discretionary to quantified risk assessment, and explanatory to predictive analytics—are not particularly transformative; rather, they represent quantified recapitulations of traditional surveillance practices. By contrast, the last two shifts—the inclusion of data on individuals with no direct police contact, and from institutions typically not associated with crime control—represent fundamental transformations in surveillance activities. The shift from query-based systems to automated alerts is a moderate shift, representing, in part, an elaboration of existing practices, and in part, a new surveillance strategy. I will analyze each of these shifts in the following sections.

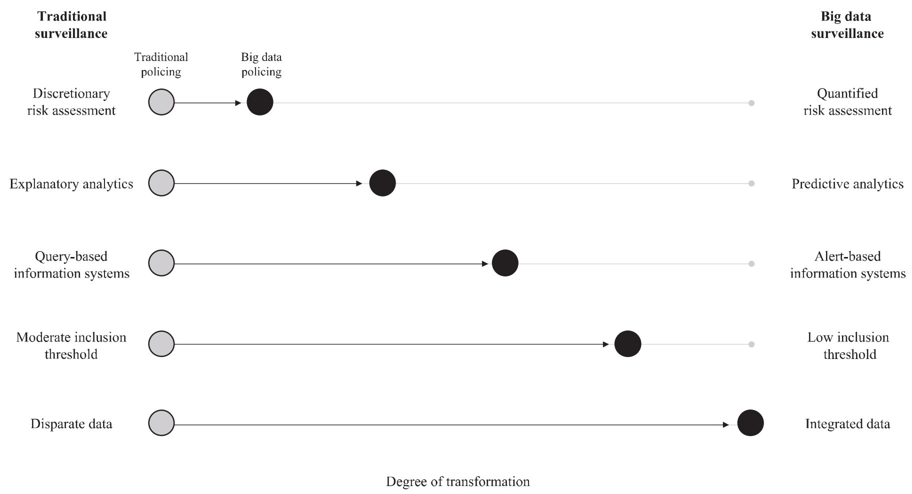


Figure 2. Migration of Traditional Police Practices toward Big Data Surveillance

The shift from traditional to big data surveillance is associated with a migration of law enforcement operations toward intelligence activities. The basic distinction between law enforcement and intelligence is as follows: law enforcement typically becomes involved once a criminal incident has occurred. Legally, the police cannot undertake a search and gather personal information until there is probable cause. Intelligence, by contrast, is fundamentally predictive. Intelligence activities involve gathering data; identifying suspicious patterns, locations, activity, and individuals; and preemptively intervening based on the intelligence acquired. Before discussing the findings, one caveat is worth noting: the migration of law enforcement toward intelligence was in its nascency before the use of big data, in part due to Supreme Court decisions dismantling certain criminal protections. Technically, the Fourth Amendment makes unreasonable searches and seizures illegal in the absence of probable cause. However, in practice, decisions such as *Terry v. Ohio* and *Whren v. United States* made it easier for law enforcement to circumvent the barrier of probable cause, ultimately contributing to the proliferation of pretext stops. In other words, the Supreme Court's dismantling of probable cause catalyzed the migration of law enforcement toward intelligence, and the adoption of big data analytics facilitated and accelerated this shift.

The Quantification of Individual Risk

The first shift in police practice is the quantification of civilians according to risk. Quantified knowledge is supplementing officers' experiential knowledge through the implementation of a new point system: Operation LASER (Los Angeles' Strategic Extraction and Restoration program). The program began in 2011 and was funded through the Smart Policing Initiative, a national initiative encouraging local police departments and researchers to use evidence-based, data-driven tactics. The strategy includes place-based and offender-based models. The offender-based strategy was implemented in a low-income, historically high-crime division in South Bureau. It is premised on the idea that a small percentage of high-impact players are disproportionately responsible for most violent crime. Therefore, identifying and focusing police resources on the "hottest" individuals should be an efficient means of reducing crime.[5]

The strategy begins by plotting crimes in the division. The Crime Intelligence Detail (CID), which is composed of three sworn officers and a civilian crime analyst, identifies a problem crime, which in this division is often armed robbery. Next, the CID shifts their unit of analysis from crimes to individuals. They gather intelligence daily from patrols, the Parole Compliance Unit, field interview (FI) cards (police contact cards), traffic citations, release from custody forms, crime and arrest reports, and criminal histories to generate a list of "chronic offenders," who are each assigned a point value and given a numerical rank according to that value. Individuals are assigned five points for a violent criminal history, five points for known gang affiliation, five points for prior arrests with a handgun, and five points if they are on parole or probation. One officer explained:

> We said ok, we need to decide who's the worst of the worst . . . we need something to pull them apart. So this was the important one, and this is really what gives the importance of FI-ing someone [filling out a field interview card] on a daily basis instead of just saying, okay, I saw that guy hanging out, I'm gonna give him two weeks and I'll go FI him again. *It's one point for every police contact*.[6]

As illustrated in Figure 3, FI cards include personal information such as name, address, physical characteristics, vehicle information, gang affiliations, and criminal history. On the back of the card, there is space for officers to include information on persons with the subject and additional intelligence.

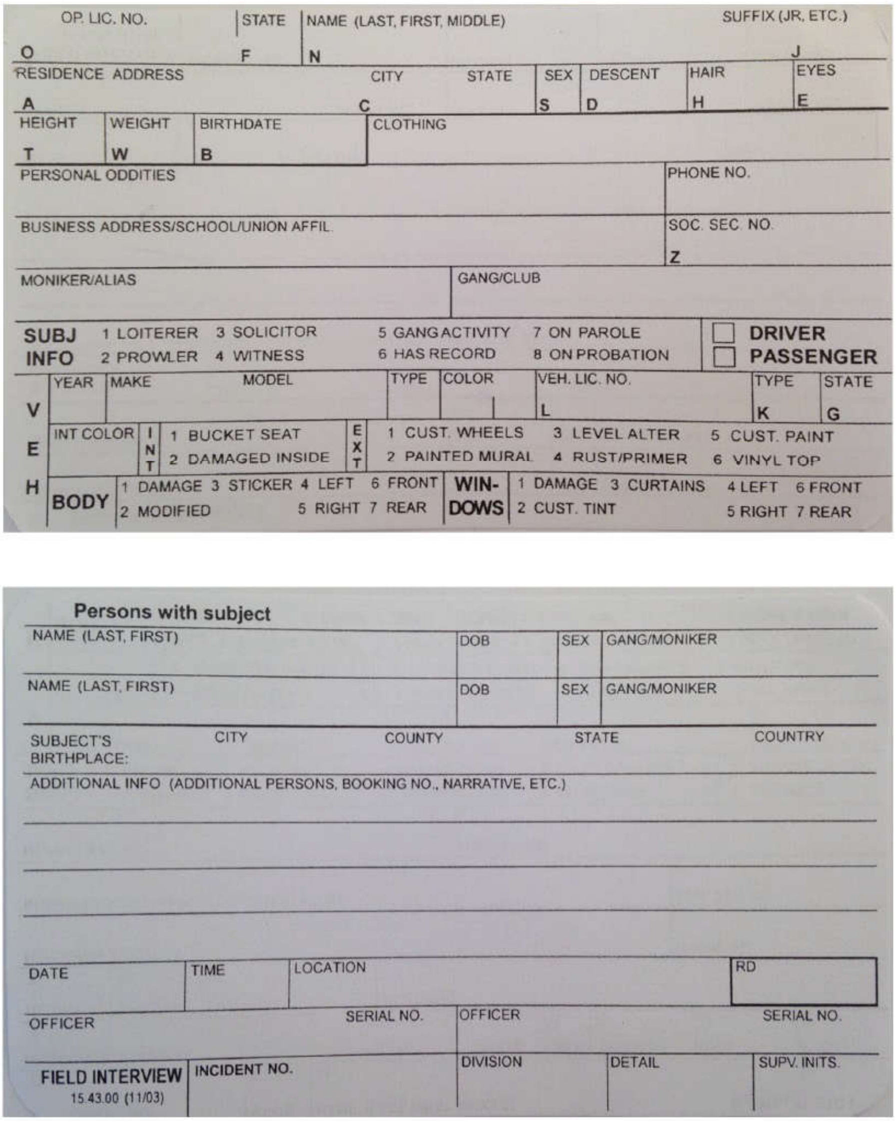
FIs are key intelligence tools for law enforcement and were one of the first data sources integrated into Palantir. When entered into the system, every FI is tagged with the time, date, and geo-coordinates. Officers are trained to pull out an FI card and "get a shake" as soon as they interact with someone in the field. One supervisor described how he uses it "to tag all the personal information I can get . . . these things come into play later on in ways you could never even imagine." Similarly, a software engineer explained how little pieces of data that might seem unsuspicious at the time of collection can eventually be pulled together to create useful intelligence: "It's a law enforcement system where that citation can, the sum of all information can build out what is needed." In addition to inputting the *content* of the cards, a captain explained there is an incentive to simply "get them in the system" as entities that future data points can be linked to.

Because point values are largely based on police contact, an important question emerges: What are grounds for police contact? Merely being identified as a chronic offender does not constitute reasonable suspicion or probable cause. However, used in conjunction with Palantir, FIs represent a proliferation of data from police–civilian interactions that law enforcement does not need a warrant to collect. When I asked an officer to provide examples of why he stops people with high point values, he replied:

> Yesterday this individual might have got stopped because he jaywalked. Today he mighta got stopped because he didn't use his turn signal or whatever the case might be. So that's two points . . . you could conduct an investigation or if something seems out of place you have your consensual stops.[7] So a pedestrian stop, this individual's walking, "Hey, can I talk to you for a moment?" "Yeah what's up?" You know, and then you just start filling out your card as he answers questions or whatever. And what it was telling us is who is out on the street, you know, who's out there not necessarily maybe committing a crime but who's active on the streets. You put the activity of . . . being in a street with maybe their violent background and one and one might create the next crime that's gonna occur.

The point system is path dependent; it generates a feedback loop by which FIs are both causes and consequences of high point values. An individual having a high point value is predictive of future police contact, and that police contact further increases the individual's point value.

The CID also creates work-ups, referred to as "Chronic Violent Crime Offender Bulletins." Individuals on these bulletins are not necessarily "wanted" nor do they have



OP. LIC. NO. | STATE | NAME (LAST, FIRST, MIDDLE) | SUFFIX (JR, ETC.)
O F N J
RESIDENCE ADDRESS | CITY | STATE | SEX | DESCENT | HAIR | EYES
A C S D H E
HEIGHT | WEIGHT | BIRTHDATE | CLOTHING
T W B
PERSONAL ODDITIES | PHONE NO.
BUSINESS ADDRESS/SCHOOL/UNION AFFIL. | SOC. SEC. NO.
Z
MONIKER/ALIAS | GANG/CLUB
SUBJ INFO 1 LOITERER 3 SOLICITOR 5 GANG ACTIVITY 7 ON PAROLE
2 PROWLER 4 WITNESS 6 HAS RECORD 8 ON PROBATION
DRIVER
PASSENGER
VEH: YEAR | MAKE | MODEL | TYPE | COLOR | VEH. LIC. NO. | TYPE | STATE
V L K G
INT COLOR | INT 1 BUCKET SEAT 2 DAMAGED INSIDE | EXT 1 CUST. WHEELS 2 PAINTED MURAL 3 LEVEL ALTER 4 RUST/PRIMER 5 CUST. PAINT 6 VINYL TOP
BODY 1 DAMAGE 3 STICKER 4 LEFT 6 FRONT 2 MODIFIED 5 RIGHT 7 REAR
WIN-DOWS 1 DAMAGE 3 CURTAINS 4 LEFT 6 FRONT 2 CUST. TINT 5 RIGHT 7 REAR

Persons with subject
NAME (LAST, FIRST) | DOB | SEX | GANG/MONIKER
NAME (LAST, FIRST) | DOB | SEX | GANG/MONIKER
SUBJECT'S BIRTHPLACE: CITY | COUNTY | STATE | COUNTRY
ADDITIONAL INFO (ADDITIONAL PERSONS, BOOKING NO., NARRATIVE, ETC.)
DATE | TIME | LOCATION | RD
OFFICER | SERIAL NO. | OFFICER | SERIAL NO.
FIELD INTERVIEW 15.43.00 (11/03) | INCIDENT NO. | DIVISION | DETAIL | SUPV. INITS.

Figure 3. Field Interview (FI) Cards
Source: LAPD.

outstanding warrants for their arrest. Rather, it is an "information only" bulletin that includes physical descriptors and oddities, gang affiliation, criminal history, parole/probation status, vehicles, frequented areas, and law enforcement contacts. The goal of these bulletins is to give officers what they refer to as "situational awareness." Officers previously had to rely exclusively on their direct knowledge of a case and specific criminal networks, but by creating a list and disseminating bulletins, the point system and associated bulletins broadens previously particularized police familiarity of individuals on the street.

Ideally, one officer explained, they could put one officer on every individual on the list and "odds are [you're] probably going to find them committing another crime." However, the police operate in an organizational context with resource constraints; respondents frequently referenced budget cuts and personnel shortages. Therefore, instead of having one officer on every chronic offender, officers engage in what I term "stratified surveillance": differentially surveilling individuals according to their risk score. An officer explained:

> [We] utilize undercover operations, or undercover units and . . . then sit our surveillance on some of the higher point offenders and just watch them on a daily basis. . . . And you start building either, you know, there's two ways of looking at it. Either kind of conducting your investigation to see if maybe there was a crime that had just been committed. Or, "We know who you are, you know, I just called you Johnny, I've never really met you before, but I know who you are now," so maybe it's put in his mind, "Oh, they're on to me, they know who I am."

This excerpt sheds light on the multiple purposes of stratified surveillance, including ongoing intelligence gathering and deterrence through signaling to individuals on the street that they are being tracked by law enforcement.

Why did the police turn to the point system in this division? In the words of one officer,

> The code of federal regulations. They say you shouldn't create a—you can't target individuals especially for any race or I forget how you say that. But then we didn't want to make it look like we're creating a gang depository of just gang affiliates or gang associates. . . . We were just trying to cover and make sure everything is right on the front end.

Other respondents echoed this sentiment, explaining the strategy was adopted, in part, as a legal compliance mechanism.

The point system is a form of quantified policing, but it is not dramatically different from its discretionary predecessor. As indicated by the low degree of transformation in Figure 2, it is largely a quantified recapitulation of traditional surveillance practices.

Shift from Reactive to Predictive Analytics

Historically, policing was mostly reactive. Patrol officers used to spend much of their time "chasing the radio." In the early 1980s, faced with evidence that reactive strategies were ineffective at reducing crime, there was a paradigm shift toward more proactive, problem-oriented policing strategies, including hot spots policing. Predictive policing is an extension of hot spots policing, made possible by the temporal density of big data (i.e., high-frequency observations). In 2012, the LAPD began using software designed by PredPol, a predictive policing company. PredPol uses a proprietary algorithm[8] predicated on the near-repeat model, which suggests once a crime occurs in a location, the immediate surrounding area is at increased risk for subsequent crime. PredPol uses three types of inputs—past type, place, and time of crime—to identify areas where future crime is most likely to occur. Predictive policing is expanding rapidly within the Department; as of March 2015, it had disseminated to 10 divisions.[9]

Officers receive printouts at the beginning of their shift that show 500 by 500 square-foot boxes overlaying small areas of division maps. Patrol officers are encouraged to spend time in predictive boxes, a strategy referred to as "risk-based deployment." Deployment is based on available time, such as when officers are not responding to calls or "booking a body." Officers record their self-reported minutes in the predictive boxes on their in-car computers. Although "data drives deployment," what the police do once in the predictive box, and how long they stay there, remains within their discretion.

One supervisor explained that by relying on data, rather than human interpretation of crime patterns, it helps him deploy his resources more efficiently:[10]

> There's an emotional element to it, and you think right now with crime being this low, a cluster could be three or four crimes. Clusters used to be 10, 12 crimes. Now three or four and they jump on it, you know. So,

> there could be overreaction. Because, there's, you know, I mean it's a human doing it. And they cannot sort out what's noise.

Officers were quick to emphasize the continued importance of their own expertise. When discussing predictive policing, most patrol officers said some version of the following statement made by a sergeant on a ride-along: "I already know where the crime's at." Part of this sentiment may stem from officers' concern that the use of algorithms represents a form of deskilling, devaluing their local and experiential knowledge and threatening their professional autonomy. In that vein, one captain described a typical exchange with his officers:

> They're like, "You know what, I know where the crime's occurring." . . . And I show them the forecast and they say, "Okay, so [at intersection], I know there are crimes, I could have told you that. I've been working here 10 years! There's always crime there." I go, "Okay, you're working here 10 years on that car, why is there still crime there if you're so knowledgeable?"

Despite some within-department conflict over their efficacy, PredPol outputs still informed where some officers drove during their uncommitted time. For example, when driving back from booking an individual at the station, a sergeant I was with decided to drive to an area not known to him for being high-crime, because he thought the location of the PredPol box was odd and he wanted to see what was going on there. In other words, predictive policing outputs sometimes—but not always—acted as a substitute for localized experiential knowledge.

A related but distinct reason why officers contest predictive policing is because they believe it places officers themselves under greater surveillance. For example, when we arrived at a crime scene on my first ride-along, I was surprised to see an officer manually type our location on his laptop. Considering how technologically advanced the Department was in other ways, I assumed cars' locations would be tracked automatically. When I asked the officer why he manually placed himself at the scene, he explained that although every police unit was equipped with an automatic vehicle locator (AVL) that pings the vehicle's location every five seconds, they were not turned on because of resistance from LAPD union representatives.[11]

Shift from Query-Based to Alert-Based Systems

The shift from query-based to alert-based systems represents, in part, an extension of existing practices and, in part, a fundamental transformation in surveillance activities. By "query-based systems," I mean databases to which users submit requests for information in the form of a search. A familiar example of a query is when a police officer runs a license plate during a traffic stop. In alert-based systems, by contrast, users receive real-time notifications (alerts) when certain variables or configurations of variables are present in the data. The shift from query-based to alert-based systems—which is made possible by high frequency data collection—has implications for the relational structure of surveillance.

Consider the following example: all warrants in L.A. County can be translated into object representations spatially, temporally, and topically in Palantir. Through tagging, users can add every known association that warrant has to people, vehicles, addresses, phone numbers, documents, incidents, citations, calls for service, ALPR readings, FIs, and the like. Officers and analysts can then set up alerts by putting a geo-fence around an area and requesting an alert every time a new warrant is issued within the area. Warrants are but one example; users can request alerts for any data points related to the entity they are interested in (e.g., calls for service, involvement in or witnesses to a traffic accident, ALPR [automatic license plate reader] readings, FIs, and so on). Using a mechanism in Palantir similar to an RSS feed, officers can be automatically notified of warrants or events involving specific individuals (or matching descriptions of individuals), addresses, or cars directly on their cell phone. Prior to automated alerts, law enforcement would know individuals' real-time location only if they were conducting 1:1 surveillance, received a tip, or encountered them in person.

Real-time notifications can be useful in operational planning. An interviewee who worked at the fusion center described how if he is about to conduct a search of a house, he can draw a fence around the house and receive notifications about risks such as whether a known gang associate lived in the home, if there was a gun registered in the house next door, or if there was a warrant for assault with a deadly weapon issued down the street.

Alerts can also be used to break down information silos within the Department. LAPD's jurisdiction is almost 500 square miles. Therefore, individual detectives may not able to connect crime series that occur across different divisions. One captain explained:

> Let's say I have something going on with the medical marijuana clinics where they're getting robbed. Okay? And it happens all over, right? But I'm a detective here in [division], I can put in an alert to Palantir that says anything that has to do with medical marijuana plus robbery plus male, black, six foot.

He continued, "I like throwing the net out there, you know? Throw it out there, let it work on it while you're doing your other stuff, you know?" Relatedly, an interviewee in Robbery-Homicide Division described a pilot project in which automated data grazing can flag potential crime series that span jurisdictional boundaries and are therefore difficult for any one person to identify. He said, "You could get an alert that would say, you know what, your case is pretty similar to this case over in Miami." If the case reaches a "merit score" (i.e., a threshold at which a certain configuration of variables is present), the system flags the cases as similar. The system matches on fields such as suspect description, license plate, type of weapon, cause of death, motive, type of crime, and M.O., such as "what kind of bindings were used . . . or was there torture involved? What type of trauma has occurred? Was there, you know, was there some type of symbolic activity?" Although the matching process is automated, decisions about what parameters the system matches on remain within the discretion of individuals at ViCAP, a unit of the FBI.

That said, the use of alerts represents not just a scaling up of existing police practices, but also a fundamental transformation in how patrol officers and investigators generate case knowledge. Under the traditional surveillance model, alerts about hot incidents and suspects are sent out from dispatch centers. However, by exploiting variation in divisions that did and did not use place- and person-based predictive policing—and divisions that started using big data during the course of my fieldwork—I was able to observe the automation of alerts and relative lack of human intermediation in broadcasting out these alerts or conducting data grazing.

It is worth noting that alert-based systems are supplementing, rather than replacing, query-based systems. Searches are still critical features of law enforcement information systems. In fact, one of the transformative features of big data systems is that *queries themselves are becoming data*. One detective explained:

> I queried the system a certain way and then another person queried the system a certain way . . . we were looking for something very similar in our query, and so even though the data may not have connected the two, the queries were similar. Yeah, so then it will be able to say hey, listen, there's an analyst in San Francisco PD that ran a very similar query as to yours and so you guys might be looking for the same thing.

A different detective explained how when he searches someone's name in one national system, he can see the number of times that name has been queried by other people. When I asked why he would want to know how many times someone's name has been queried, he replied that "if you aren't doing anything wrong," the cops are not going to be looking you up very many times over the course of your life. He continued: "Just because you haven't been arrested doesn't mean you haven't been caught." In other words, in auditable big data systems, queries can serve as quantified proxies for suspiciousness.

Lower Database Inclusion Thresholds

The last two shifts in practice represent the most fundamental transformations in surveillance activities. Law enforcement databases have long included information on individuals who have been arrested or convicted of crimes. More recently, they also include information on people who have been stopped, as evidenced by the proliferation of stop-and-frisk databases. However, as new data sensors and analytic platforms are incorporated into law enforcement operations, the police increasingly utilize data on individuals who have not had any police contact at all. Quotidian activities are being codified by law enforcement organizations. One way this is occurring is through network analysis. Figure 4 is a de-identified mockup I asked an employee at Palantir to create based on a real network diagram I obtained from an officer in the LAPD. The person of interest, "Guy Cross," is an individual with a high point value. An LAPD officer explained, "with the [Palantir] system . . . I click on him and then [a] web would spread out and show me the phones that he's associated with and the cars."

Radiating out from "Guy Cross," who has direct police contact, are all the entities he is related to, including people, cars, addresses, and phone numbers. Each line indicates how they are connected, such as by being a sibling, lover, cohabiter, co-worker, co-arrestee, or listed on a vehicle registration. The network diagram illustrates only one degree of separation, but networks can expand outward to as many degrees of separation as users have information and can tie in with other networks. To be in what I call the "secondary surveillance network," individuals do not need to have direct law enforcement contact; they simply need to have a link to the central person of interest. Once individual relationships are inputted and social networks are built into the system, individuals can be "autotracked," meaning officers can receive real-time alerts if individuals in the network come into contact with the police or other government agencies again.

The Automatic License Plate Reader (ALPR) is another example of a low-threshold "trigger mechanism" (Tracy and Morgan 2000) that results in more widespread inclusion in a database. ALPRs are dragnet surveillance tools; they take readings on everyone, not merely those under suspicion. Cameras mounted on police cars and static ALPRs at intersections take two photos of every car that passes through their line of vision—one of the license plate and one of the car—and records the time, date, and GPS coordinates (see Figure 5). Law enforcement–collected ALPR data can be supplemented with privately collected ALPRs, such as those used by repossession agents. ALPR data give the police a map of the distribution of vehicles throughout the city and, in some cases, may enable law enforcement to see an individual's typical travel patterns. For example, an analyst used ALPR data to see that a person of interest was frequently parked near a particular intersection at night, explaining to me that this intersection is likely near that person's residence or "honeycomb" (hideout).

There are several ways to use ALPR data. One is to compare them against "heat lists" of outstanding warrants or stolen cars. Another strategy is to place a geo-fence around a location of interest in order to track cars near the location. For example, after a series of copper wire thefts in the city, the police found the car involved by drawing a radius in Palantir around the three places the wire was stolen from, setting up time bounds around the time they knew the thefts occurred at each site, and querying the system for any license plates captured by ALPRs in all three locations during those time periods.

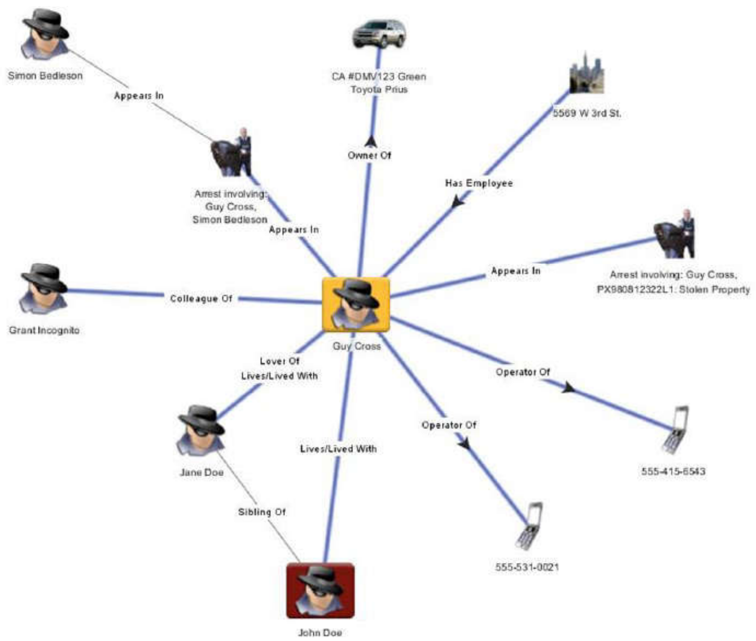


Figure 4. Network in Palantir
Source: Palantir Technologies.

However, the most common use of ALPRs is simply to store data for potential use during a future investigation. For example, one sergeant described a "body dump" (the disposal of a dead body) that occurred in a remote location near a tourist attraction where there was an ALPR. By searching ALPR readings within the time frame that police determined the body was disposed, they captured three plates—one from Utah, one from New Mexico, and one from Compton. The sergeant explained that assuming the Compton car was most likely to be involved, they ran the plate, saw the name it was registered under, searched the name in CalGang (gang database), saw that the individual was affiliated with a gang currently at war with the victim's gang, and used that information to establish probable cause to obtain a search warrant, go to the address, find the car, search the car for trace evidence, and arrest the suspect.

Although LAPD and Palantir employees frequently told me that to be "in the system," a person needed to have had criminal justice contact, the use of network diagrams and the inclusion of ALPR data in the Palantir platform offer clear examples in which individuals with no criminal justice contact are included in law enforcement databases.

Institutional Data Systems Are Integrated

Finally, the proliferation of digitized records makes it possible to merge data from previously separate institutional sources into an integrated, structural system in which disparate data points are displayed and searchable in relation to one another, and individuals can be cross-referenced across databases. This integration facilitates one of the most transformative features of the big data landscape: the creep of criminal justice surveillance into other, non–criminal justice institutions. Function creep—the phenomenon of data originally collected for one purpose being used for another—contributes to a substantial increase in the data police have access to. Indeed, law enforcement is following an institutional data imperative (Fourcade and Healy 2017),

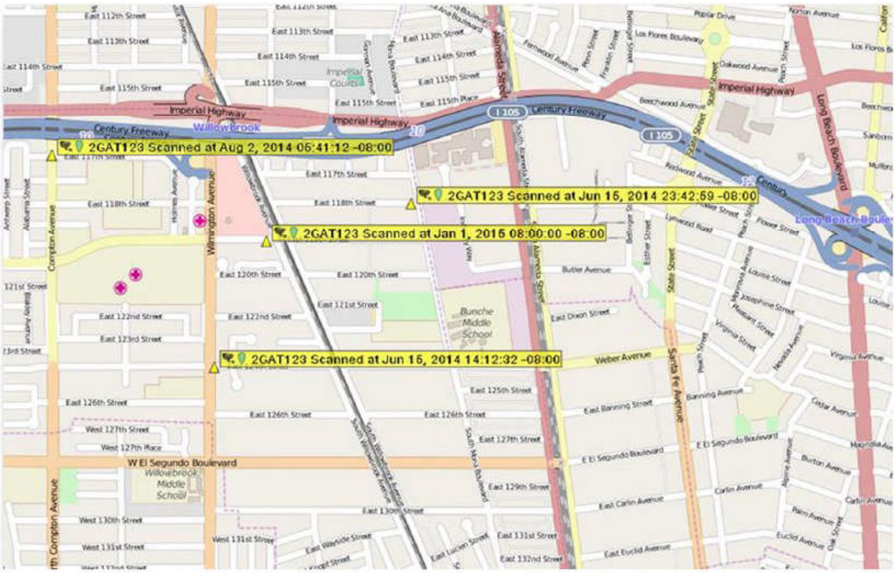


Figure 5. Plotted ALPR Readings
Source: Palantir Technologies.

securing routine access to a wide range of data on everyday activities from non-police databases. Before Palantir, officers and analysts conducted predominantly one-off searches in "siloed" systems: one to look up a rap sheet, another to search a license plate, another to search for traffic citations, and so on. The Palantir platform integrates disparate data sources and makes it possible to quickly search across databases.

Expressing his faith in the Department's investment in the platform, one captain told me, "We've dumped hundreds of thousands into that [Palantir]. . . . They're gonna take over the world. . . . I promise you they're gonna take over the world." During my fieldwork, there were more than 1,300 trained Palantir users in the region. New data sources are incorporated regularly, including information collected by the Department, external data collected by other government agencies, and privately collected data the Department purchases. Remarking on the growth, one captain said:

> I'm so happy with how big Palantir got. . . . I mean it's just every time I see the entry screen [Figure 6] where you log on there's another icon about another database that's been added . . . they now have been working with Palantir to develop a database of all the foreclosure properties . . . they just went out and found some public data on foreclosures, dragged it in, and now they're mapping it where it would be relative to our crime data and stuff.

The Palantir platform allows users to organize and visualize structured and unstructured data content (e.g., e-mails, PDFs, and photos) through "tagging," the process of labeling and linking objects and entities to identify emerging relationships. By tagging objects and entities—including, but not limited to, persons, phone numbers, addresses, documents such as law enforcement reports or tips and leads, and calls for service—and displaying the data spatially, temporally, or topically, users can see data points in context and make new connections.

Another important interagency data integration effort is the initiative to create an Enterprise Master Person Index (EMPI) in

Figure 6. Palantir Homepage
Source: LAPD.

L.A. County. L.A. EMPI would create a single view of a client across all government systems and agencies; all of an individual's interactions with law enforcement, social services, health services, mental health services, and child and family services would be merged onto one unique ID. Although interviewees working in the county's information technology office stated the explicit motivation behind the initiative was to improve service delivery, such initiatives effectively serve the latent function of extending the governance and social control capacities of the criminal justice system into other institutions.

I encountered several other examples of law enforcement using external data originally collected for non–criminal justice purposes, including data from repossession and collections agencies; social media, foreclosure, and electronic toll pass data; and address and usage information from utility bills. Respondents also indicated they were working on integrating hospital, pay parking lot, and university camera feeds; rebate data such as address information from contact lens rebates; and call data from pizza chains, including names, addresses, and phone numbers from Papa Johns and Pizza Hut. In some instances, it is simply easier for law enforcement to purchase privately collected data than to rely on in-house data because there are fewer constitutional protections, reporting requirements, and appellate checks on private sector surveillance and data collection (Pasquale 2014). Moreover, respondents explained, privately collected data is sometimes more up-to-date.

It is worth noting that such data integration is not seamless. Merging data from different sources and creating interoperable systems is part of the invisible labor that makes big data analytics possible. One civilian employee lamented, "You always forget about the data guy . . . [the] guy that does all the dirty work is usually forgotten. I'm that guy." Moreover, efforts at acquiring new data were not received evenly throughout the Department. A vocal minority of interviewees explained they did not believe leadership was fully thinking through the implications of collecting such a wide range of new data. A civilian employee complained about the seduction of new technology, saying: "We tend to just say, 'Let's just go for the sexy tool,' right? . . . We just never think about to what end." He added,

> Maybe we shouldn't collect this information. Maybe we shouldn't add consumer information. Maybe we shouldn't get everybody's Twitter feed in. . . . All we're doing right now is, "Let's just collect more and more and more data and something good will just happen." And that's I think that's kind of wishful thinking.

Law enforcement's adoption of data and analytic tools without a specific technical purpose also surfaced during my time at surveillance industry conferences. When I first observed software representatives interact with potential law enforcement customers, I assumed law enforcement would tell software representatives their needs and ask how the products could help them achieve their operational goals. However, the inverse pattern was more frequently the case: software representatives demonstrated the use of their platform in a non-law enforcement—usually military—context, and then asked local law enforcement whether they would be interested in a similar application in their local context. In other words, instead of filling analytic gaps or technical voids identified by law enforcement, software representatives helped create new kinds of institutional demand.

DISCUSSION: BIG DATA AS SOCIAL

This article draws on unique data to offer an on-the-ground account of big data surveillance. Providing a case study of the Los Angeles Police Department (LAPD), it offers insight into the reasons why the use of big data analytics spread throughout the organization, including factors particular to the LAPD such as consent decree mandates, but also broader isomorphic shifts (DiMaggio and Powell 1983) toward use of predictive analytics across organizational fields. In analyzing how the LAPD uses big data in their surveillance activities, I argue it is both *continuous* and *transformative*: the adoption of advanced analytics facilitates amplifications of existing surveillance practices, but also fundamentally changes daily operations. I describe five key shifts in practice associated with adoption of big data analytics, each of which falls on different points on the continuum between law enforcement and intelligence activities. Whereas the person-based point system and place-based predictive algorithms are largely quantified recapitulations of "traditional" (Marx 2016) surveillance, the inter-institutional integration of data and proliferation of dragnet surveillance practices—including the use of data on individuals with no direct police contact and data gathered from institutions typically not associated with crime control—represent fundamental transformations in the very nature of surveillance.

Big data and associated new technological tools permit unprecedentedly broad and deep surveillance. By broad, I mean surveillance capable of passively tracking a large number of people. Information that would previously have been unknown to law enforcement because it was too labor intensive to retrieve is more readily available, and individuals previously unknown to law enforcement are now part of the corpus through dragnet surveillance and data collection by non–criminal justice organizations. By deep, I mean able to track one individual more intensively over time, including across different institutional settings. The intended and unintended social consequences of new surveillance practices have implications for social inequality, law, and future research on big data surveillance in other fields.

Implications for Social Inequality

The role of the criminal justice system in the reproduction of inequality has received considerable attention in the literature (for a review, see Laub 2014). However, the impact of the use of big data surveillance on inequality remains an open empirical question. The use of new surveillant technologies could either reduce or reinforce existing inequalities. By contributing new insights into how big data plays out on the ground in policing, this research helps adjudicate between the two possibilities.

On the one hand, big data analytics may be a means by which to ameliorate persistent inequalities in policing. Data can be marshaled to replace unparticularized suspicion of racial minorities and human exaggeration of patterns with less biased predictions of risk. Social psychological research demonstrates that humans are "cognitive misers" (Fiske and Taylor 1991) who rely on shortcuts—such as the conflation of blackness and criminality (Quillian and Pager 2001)—to understand the world. Because stereotypes have the most cognitive utility in the face of incomplete information, if big data can be utilized to provide more complete information, it may lead officers to rely less on stereotypes about race and class. In that sense, the use of big data may serve to reduce hyper-surveillance of minority neighborhoods and the consequent erosion of community trust (Sampson and Bartusch 1998). Big data may also be used to "police the police." Digital trails are susceptible to oversight. Therefore, aggregating data on police practices may shed light on systematic patterns and institutional practices previously dismissed as individual-level bias, ultimately providing an opportunity to increase transparency and accountability. However, transparency and accountability do not flow automatically from big data policing. Data-based surveillance is less visible than traditional street policing methods (Joh 2016) and is embedded in power structures. The outcomes of struggles between law enforcement, civilians, and information technology companies—who increasingly own the storage platforms and proprietary algorithms used in data analysis—will play a role in determining whether big data policing will ameliorate or exacerbate inequalities.

On the other hand, this research highlights how data-driven surveillance practices may be implicated in the reproduction of inequality in at least three ways: by deepening the surveillance of individuals already under suspicion; widening the criminal justice dragnet unequally; and leading people to avoid "surveilling" institutions that are fundamental to social integration. First, mathematized police practices serve to place individuals already under suspicion under new and deeper forms of surveillance, *while appearing to be objective*, or, in the words of one captain, "just math." Despite the stated intent of the point system to avoid legally contestable bias in police practices, it hides both intentional and unintentional bias in policing and creates a self-perpetuating cycle: if individuals have a high point value, they are under heightened surveillance and therefore have a greater likelihood of being stopped, further increasing their point value. Such practices hinder the ability of individuals already in the criminal justice system from being further drawn into the surveillance net, while obscuring the role of enforcement in shaping risk scores. Moreover, individuals living in low-income, minority areas have a higher probability of their "risk" being quantified than those in more advantaged neighborhoods where the police are not conducting point-driven surveillance. Importantly, this quantified modality of social control has consequences that reach beyond individuals with high point values. Field interview cards record information not only about the individual in question, but also information on people the individual is with. The exponential capture of personal data beyond the primary individuals involved in the police encounter is a strategic means of channeling more individuals into the system, thus facilitating future tracking.

Whereas the point system is consequential for racial and class inequality, if not implemented effectively,[12] place-based algorithms may exacerbate neighborhood inequalities. Historical crime data are incomplete; estimates of unreported crime range from less than 17 percent to over 68 percent, depending on the offense (Langton et al. 2012). Moreover, crime data are not missing at random. Therefore, there is systematic bias in the training data: crimes that take place in public places are more visible to police and therefore more likely to be recorded (Duster 1997); individuals and groups who do not trust the police are less likely to report crimes (Sampson and Bartusch 1998); and police focus their attention and resources on black communities at a disproportionately high rate relative to drug use and crime rates (Beckett et al. 2005). These social dynamics inform the historical crime data that are fed into the predictive policing algorithm. However, once they are inputted as data, the predictions appear impartial; human judgment is hidden in the black box (Pasquale 2014) under a patina of objectivity.

Unchecked predictions may lead to an algorithmic form of confirmation bias, and subsequently, a misallocation of resources. They may justify the over-policing of minority communities and potentially take away resources from individuals and areas invisible to data collection sensors or subject to systematic underreporting. Put differently, the mechanisms for inclusion in criminal justice databases determine the surveillance patterns themselves. Predictive models are performative, creating a feedback loop in which they not only predict events such as crime or police contact, but also contribute to their future occurrence.[13]

Second, new digitized surveillance practices broaden the scope of people law enforcement can track. This can be understood as a new form of "net widening" (Cohen 1985), effectively widening the criminal justice dragnet, and doing so unequally. Consider ALPRs, one of the primary means of tracking people without police contact. Even though ALPRs are dragnet surveillance tools that collect information on everyone, rather than merely those under suspicion, the likelihood of being inputted into the system is not randomly distributed. Crime and enforcement patterns lead to unequal data capture across individuals, groups, and the city. ALPRs are deployed based on department crime statistics (i.e., to higher crime areas), raising similar questions to those posed earlier about unequal enforcement and reporting practices along lines of race, class, and neighborhood. In that sense, ALPR datasets are investigatory tools for law enforcement, but they are disproportionately "populated by the movements of particular groups" (Renan 2016:1059). Similarly, the ability to build out secondary surveillance networks in Palantir has implications for inequality, as minority individuals and individuals in poor neighborhoods have a higher probability of being in the primary (and thus secondary) surveillance net than do people in neighborhoods where the police are not conducting point-driven or other data-intensive forms of policing.

How are unequal mechanisms for inclusion in the surveillance net consequential for social inequality? Recall the operative theory from a detective that if people are not doing anything wrong, the police should not be looking them up many times over the course of their lives. However, queries are not raw data (Gitelman 2013); rather, they are, in part, a product of enforcement practices. Empirical research consistently demonstrates that stop-and-query patterns are unequally distributed by race, class, and neighborhood (Epp, Maynard-Moody, and Haider-Markel 2014). Quantified practices may thus serve to exacerbate inequalities in stop patterns, create arrest statistics needed to justify stereotypes, and ultimately lead to self-fulfilling statistical prophecies (Merton 1948). Moreover, as police contact is the entry point into the criminal justice system, the digital feedback loops associated with predictive policing may ultimately justify the growth and perpetuation of the carceral state.

One might argue that if you have nothing to hide, being included in police databases is nothing to fear. However, once individuals are in the primary or secondary surveillance net, they can become intelligence targets and linked to future data points. By virtue of being in the system, individuals are more likely—correctly or incorrectly—to be identified as suspicious. Consider how the quantification of previous stops in the point system serves as justification for future stops, or how the detective suggested a man was suspicious because his name had been queried multiple times. Using a series of data points to reconstruct an individual's intentions and behaviors, whether incriminating or exculpatory, rests on the assumption of an infallible state and of actors who run searches without error or prejudice. Much like in DNA databases (Duster 2005; Hindmarsh and Prainsack 2010; Lynch et al. 2008), in order to be a hit, one has to be in the database in the first place. Unequal rates of database inclusion can have real consequences—African Americans are seven times more likely than whites to be wrongly convicted of murder (Gross, Possley, and Stephens 2017). Therefore, analyzing the feeder mechanisms by which individuals are channeled into criminal justice databases helps us better understand how inequalities produced by differential surveillance may be magnified as individuals are processed through the criminal justice system.

Third, integrating external, non-police data into the law enforcement corpus has unanticipated consequences. Although integrated systems create new opportunities for service delivery, they also make surveillance possible across formerly discrete institutional boundaries. By using other institutions' data, criminal justice surveillance practices may have a chilling effect, deterring people from using such institutions and thereby subverting their original mandates. For example, individuals wary of criminal justice surveillance may avoid interacting with important institutions where they would leave a digital trace. Previous research demonstrates that individuals involved in the criminal justice system (i.e., who have been stopped by police, arrested, convicted, or incarcerated) engage in "system avoidance," systematically avoiding surveilling institutions such as medical, financial, educational, and labor market institutions that keep formal records (i.e., put them "in the system") (Brayne 2014). Given that involvement with the criminal justice system is highly stratified, the negative consequences of system avoidance—for future health outcomes, financial self-sufficiency, acquisition of human capital, and upward economic mobility—will be similarly disproportionately distributed, thus exacerbating any preexisting inequalities for an expanding group of already disadvantaged individuals.

This research builds on work on labeling theory, extending the relationship between the stigma of criminal justice contact and inequality into the digital age (Becker 1963; Brayne 2014; Goffman 2014; Goffman 1963; Kohler-Hausmann 2013; Lyon 2006; Pager 2007; Rios 2011; Stuart 2016; Wakefield and Wildeman 2013; Western and Pettit 2005). The integration of records may effectively extend the mark of a criminal record (Pager 2007)—or merely the mark of criminal justice *contact*—into other institutions. This creep of data across institutional contexts can lead to "cascading disadvantages" (Pasquale 2014:218; see also Gandy 2009). As individuals leave more digital traces, a "new economy of moral judgement" (Fourcade and Healy 2017:24) becomes possible. Building on Weber's concept of class situation, Fourcade and Healy (2013) argue that institutions now use actuarial techniques to track, sort, and categorize individuals into "classification situations" with different rewards and punishments attached. These classification situations differentially shape life chances (see also Bowker and Star 2000). For example, classifying individuals as low or high risk for crime, terrorist activity, loan default, or medical conditions structures not only if and how they will be surveilled, but also their life chances more generally. This research begins to account for how the marking process may be changing in the age of digitized policing, and how the big data environment creates potentially farther-reaching digitized collateral consequences of involvement in the criminal justice system.

In summary, the burden of new surveillance practices is not borne equally, nor is the error they produce (Guzik 2009). That said, this research does not necessarily suggest the police intentionally use big data maliciously. Rather, as Barocas and Selbst (2016) argue, discrimination may be, at least in part, an artifact of the data collection and analysis process itself. Algorithmic decision procedures can "reproduce existing patterns of discrimination, inherit the prejudice of prior decision makers, or simply reflect the widespread biases that persist in society" (Barocas and Selbst 2016:674). Understanding each step of data collection and analysis is crucial for understanding how data systems—despite being thought of as objective, quantified, and unbiased—may inherit the bias of their creators and users. As an institution historically implicated in the reproduction of inequality, understanding the intended and unintended consequences of machine-learned decisions and new surveillant technologies in the criminal justice system is of paramount importance.

Implications for Law

Technological tools for surveillance are far outpacing legal and regulatory responses to the new surveillant landscape. Therefore, the findings from this project have important implications for law. First, current privacy laws—such as the Privacy Act of 1974—are anachronistic because they largely concern controls at the point of data collection. With the increased capacity to store vast amounts of data for significant periods of time, privacy laws now must also account for function creep, protecting individuals from potential future secondary uses of their data. Relatedly, law enforcement routinely purchases privately collected data, blurring the lines between public and private and highlighting the importance of revisiting third-party doctrine in the digital age.[14]

Second, use of big data for predictive analytics challenges the traditional paradigm of Fourth Amendment law, which is *transactional*: it focuses on one-off interactions between law enforcement and a suspect. However, police surveillance is increasingly *programmatic*: it is ongoing, cumulative, and sometimes suspicionless (Renan 2016). In terms of ALPRs, for example, "what begins as more generalized collection can morph into something quite different when the government runs individuated searches in its datasets" (Renan 2016:1053). Therefore, it is an open question whether cumulative surveillance should require different legal frameworks, such as administrative law, "from those that govern each isolated step" (Renan 2016:1058), namely criminal procedure.

Third, when big data—such as predictive policing forecasts—are combined with small data—such as traditional individualized suspicion based on particularized facts about a suspect—it effectively makes it easier for law enforcement to meet the reasonable suspicion standard in practice. In the words of one captain, "Some officer somewhere if this [predictive policing] gets big enough is going to say, 'okay, everybody in the box is open season,' you know? And that's not the case." Therefore, legal scholars such as Ferguson (2015:336) suggest the courts should "require a higher level of detail and correlation using the insights and capabilities of big data."

Fourth, dragnet surveillance tools such as ALPRs represent a proliferation of pre-warrant surveillance and make everyday mass surveillance possible at an unprecedented scale. Pre-crime data can be mined for links once criminal suspicion comes into play. Once in a database, a suspect can repeatedly be surveilled; law enforcement can retroactively search ALPR data and identify individuals, vehicles, times, and places, rather than starting to gather information on them only once they come under suspicion. The retroactive nature of policing in an era of dragnet data collection means information is routinely accumulated and files are lying in wait. In that sense, *individuals lead incriminating lives*—daily activities, now codified as data, can be marshaled as evidence ex post facto. The proliferation of pre-warrant surveillance tools also creates new opportunities for parallel construction, the process of building a separate evidentiary base for a criminal investigation to conceal how the investigation began, if it involved warrantless surveillance or other inadmissible evidence.

Finally, previous practical constraints, which placed natural limits on the scope of surveillance, are less relevant in light of new dragnet tools. One new analytic technique or surveillant technology on its own may not be consequential, but the combined power of using, for example, the person-based point system in conjunction with ALPR data in conjunction with network diagrams in Palantir grants authorities a level of insight into an individual's life that historically would have constituted a Fourth Amendment search and thus required a warrant. However, because no one of those surveillance practices falls outside the parameters of the law in isolation, neither does their combination.[15] In that sense, intelligence is essentially pre-warrant surveillance. Detectives and prosecutors rarely find a "smoking gun," a member of Palantir's legal counsel explained, but they can now build up a sequence of events that they were previously unable to. By "integrating data into a single ontology," he continued, users can draw connections between actors and depict a coherent scheme. Hunches that would be insufficient grounds for obtaining a warrant can be retroactively backed up using existing data, and queries can be justified in hindsight after data confirm officer suspicions. Instead of needing to justify to a judge why they require a warrant, law enforcement can first take advantage of the surveillance opportunities new technologies provide.

Implications for Research in Other Fields

Big data is being utilized for surveillance practices in a wide range of institutional domains beyond policing, including but not limited to health, finance, credit, marketing, insurance, education, immigration, defense, and activism. Although this study focused on law enforcement, big data surveillance is not something over which the LAPD has exclusive domain. Rather, this case reflects broader institutional shifts toward the use of emergent technologies and advanced analytics. Thinking beyond policing, future research may consider how big data surveillance practices identified in this study may operate in similar or different ways across fields.

Drawing from work in surveillance studies helps us systematically compare use of the newest modality of surveillance—big data—across domains. Informed by Lyon's (2003) theory of surveillance as social sorting and Marx's (2016) discussion of surveillance means, goals, and data attributes, I offer three concrete questions for future research that could help us better understand the use of big data for surveillance across institutional domains: Why was big data surveillance adopted (goals)? How is big data surveillance conducted (means)? What interventions are made based on big data surveillance, and to what consequence (ends)? Table 1 summarizes these questions.

First, why was big data surveillance adopted? What institutional goals was it intended to achieve? Lyon (2003:1) argues that the organizational imperative for surveillance is one of social sorting: "Surveillance today sorts people into categories, assigning worth or risk, in ways that have real effects on their life-chances . . . it is a vital means of sorting populations for discriminatory treatment" (see also Ericson and Haggerty 1997; Rule 2007). He identifies different categories of surveillance, each of which have different mandates and classificatory goals. Actors in the criminal justice system, for example, engage in "categorical suspicion" (Lyon 2003), collecting information to classify individuals according to risk and to identify threats to law and order. The purpose of surveillance in other institutions, however, may not be categorical suspicion but rather "categorical seduction"—classifying customers for targeted marketing, financial services, or credit (Lyon 2003; see also Gandy [1993] on the "panoptic sort")—or "categorical care"—surveillance in health and welfare organizations aimed at improving services through better coordination of personal data (Ball and Webster 2007). Analyzing changes and continuities in the structure of relationships

Table 1. Framework for Analyzing Big Data Surveillance across Institutional Contexts

	Goals		*Means*	*Ends*	
Types of Surveillance	*Institutional Field*	*Relationship between Individual and Institution*	*Shifts in Surveillance Practices Associated with Big Data*	*Institutional Interventions*	*Consequences for Inequality*
Categorical Suspicion	Criminal justice, intelligence	Classifying individuals according to risk; potential as criminals/ terrorists	1) Discretionary to quantified risk assessment 2) Explanatory to predictive analytics	Marking, apprehension, social control	Stigma, spillover into other institutions
Categorical Seduction	Finance, marketing, credit	Classifying individuals according to their value to companies; potential as customers	3) Query-based to alert-based systems 4) Moderate to low inclusion thresholds 5) Disparate to integrated data	Different products, perks, access to credit, opportunities, constraints	Upward or downward economic mobility; reproducing current patterns
Categorical Care	Medical care, public assistance	Classifying individuals according to their need; potential as clients		Personalized medicine, welfarist service delivery	May reduce inequality except when intersects with suspicion or seduction

between agents of surveillance and those who are surveilled (Marx 2016) may help us more fully understand the social process of big data surveillance and its consequences for social stratification.

Whereas surveillance goals have not fundamentally changed much over the past century, surveillance means have transformed considerably. By providing a detailed analysis of how the means of surveillance have changed in the age of big data, this study may inform future research on big data surveillance in other fields. Are the five shifts in practice identified in this study—discretionary to quantified risk assessments, explanatory to predictive analytics, query-based to alert-based systems, moderate to low inclusion thresholds, and disparate to integrated databases—occurring in other institutional contexts? For example, to what extent are data collected beyond their proximate institutional environments being used in healthcare or finance?

The goals and means of big data surveillance inform the final question posed for future research: what are the ends of big data surveillance? What do institutional actors *do* based on insights gleaned from big data surveillance, and with what consequence? Surveillance involves extracting information from different flows (Deleuze and Guattari 1987; Haggerty and Ericson 2000; Marx 2016). Distinct information flows are then reassembled into a "data double"—a digital approximation of individuals based on the electronic traces they leave (Poster 1990:97)—which is used to decide on differential treatment. Digital scores and ranks can be understood as a form of capital (Fourcade and Healy 2017), used to determine who the police stop, who credit bureaus determine as credit-worthy, and who public assistance agencies deem eligible for benefits. Simply put, one's surveillance profile structures the types of communications, opportunities, constraints, and care one receives. Big data surveillance could therefore have stratifying effects if individuals in positions of structural disadvantage are more likely to be subject to harmful forms of surveillance, and those in positions of structural advantage are more likely to be targeted by advantageous surveillance and classification schemes (Fourcade and Healy 2017). Therefore, categorical suspicion, seduction, and care may have very different implications for social inequality, depending on what institutional actors *do* based on the intelligence acquired through big data surveillance.

This article demonstrates that in the digital age, individuals leave data traces hundreds of times throughout the day, each of which contributes to the corpus of big data that a growing number of institutions use for decision-making. Institutional actors making decisions based on big data may assume that data doubles are more accurate, or unbiased, representations of a person's profile than are those gleaned from "small" data, such as personal observations. However, this perspective obscures the social side of big data surveillance. Systematic bias—whether intentional or unintentional—exists in training data used for machine learning algorithms, and it may be an artifact of human discretion or the data mining process itself. Moreover, the implications of false positives and false negatives associated with big data surveillance vary widely across domains. The stakes for being wrongly arrested for a crime you did not commit are very different from receiving a movie recommendation not to your taste. Furthermore, categories of surveillance are not mutually exclusive in the age of big data, as information can be shared across previously separate institutional boundaries. For example, electronic medical records were originally created to improve prescription drug and care coordination, but they are increasingly used to police the illicit use and sale of prescription drugs. The places where categories of surveillance intersect, and therefore have ambiguous implications for inequality, may be particularly fruitful sites for future sociological inquiry.

Finally, future research may examine the political economy underpinning the procurement of analytic software that organizations use for big data surveillance. Examining the genealogies of surveillance technologies, for example, reveals that many of the resources for developing big data analytics come from federal funds. In the law enforcement context, those grants quickly become subsumed into police organizations' operating budgets. Therefore, departments have an incentive to continue using big data—or appear to be using it—even if it is not an effective means of solving the organization's first-order problems, such as reducing crime.

Understanding the implications of big data surveillance is more complex than simply knowing who is surveilled more or less. Instead, we need to understand who is surveilled by whom, in what way, and for what purpose. How surveillance structures life chances may differ according to the goals, means, and ends involved. Although surveillance is a generalizable organizational imperative, big data is changing the means of surveillance. Accordingly, this article helps us better understand *how* big data surveillance is conducted, and calls for systematic research on the relationship between the goals, means, and ends of big data surveillance across institutional domains.

CONCLUSIONS

Through a case study of the Los Angeles Police Department, this article analyzed the role of big data in surveillance practices. By socially situating big data, I examined why it was adopted, how it is used, and what the implications of its use are. Focusing on the interplay between surveillance practices, law, and technology offers new insights into social control and inequality. I argued that big data participates in and reflects existing social structures. Far from eliminating human discretion and bias, big data represents a new form of capital that is both a social product and a social resource. What data law enforcement collects, their methods for analyzing and interpreting it, and the way it informs their practice are all part of a fundamentally social process. Characterizing predictive models as "just math," and fetishizing computation as an objective process, obscures the social side of algorithmic decision-making. Individuals' interpretation of data occurs in preexisting institutional, legal, and social settings, and it is through that interpretive process that power dynamics come into play.

Use of big data has the potential to ameliorate discriminatory practices, but these findings suggest implementation is of paramount importance. As organizational theory and literature from science and technology studies suggests, when new technology is overlaid onto an old organizational structure, longstanding problems shape themselves to the contours of the new technology, and new unintended consequences are generated. The process of transforming individual actions into "objective" data raises fundamentally sociological questions that this research only begins to address. In many ways, it transposes classic concerns from the sociology of quantification about simplification, decontextualization, and the privileging of measurable complex social phenomena onto the big data landscape.

Surveillance is always ambiguous; it is implicated in both social inclusion and exclusion, and it creates both opportunities and constraints. The way in which surveillance helps achieve organizational goals and structure life chances may differ according to the individuals and institutions involved. Examining the means of big data surveillance across institutional domains is an open and timely line of inquiry, because once a new technology is disseminated in an institutional setting, it is difficult to scale back.

Acknowledgments

I wish to thank Devah Pager for her invaluable support and guidance. I am also grateful for the many helpful comments I received from Paul DiMaggio, Janet Vertesi, Kim Lane Scheppele, Maria Abascal, David Pedulla, Becky Pettit, the Social Media Collective at Microsoft Research, and the *ASR* editors and anonymous reviewers. Finally, I wish to thank anonymous individuals within the Los Angeles Police Department for making this research possible.

Funding

This research was funded by the Horowitz Foundation for Social Policy. Additional support was provided by grant, 5 R24 HD042849, awarded to the Population Research Center at The University of Texas at Austin by the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Notes

1. For example, one division in the valley began using PredPol (predictive policing) in 2012, and nine other divisions followed suit between then and March 2015.
2. More than 25 officers in Rampart Division's special operations anti-gang unit, C.R.A.S.H., were investigated or charged, and over 100 criminal cases were overturned due to police misconduct.
3. A consent decree is a binding court order memorializing an agreement between parties in exchange for an end to a civil litigation or a withdrawal of a criminal charge.
4. *Brown v. Plata* is a 2011 U.S. Supreme Court decision holding that the overcrowding of California prisons and lack of access to adequate healthcare violated prisoners' Eighth Amendment constitutional rights.
5. To date, one study has evaluated the efficacy of Operation LASER (Uchida and Swatt 2015). It found a reduction in crime in reporting districts that adopted both person-based and location-based approaches. The program has not yet been subject to external evaluation; the authors are the President of and Senior Research Associate at Justice and Security Strategies, who designed Operation LASER.
6. Block quotes are drawn from audiotaped, transcribed interviews.
7. Consensual stops may be conducted at any time when the police lack the "specific and articulable facts" (*Terry v. Ohio* 392 U.S. at 21) that justify detention or arrest.
8. PredPol's algorithm was published by Mohler and colleagues in 2015.
9. In line with other law enforcement agencies, the LAPD is a hierarchical organization and employees are usually subject to tight managerial control. However, there was more between-division variation in the use of algorithms than I originally expected. Big data technologies were not adopted in the 21 area divisions and the specialized divisions at the same time, largely due to the autonomy granted to captains in each division during the early stages of algorithmic policing. Entrepreneurial

captains who were early adopters used algorithmic techniques largely of their own volition, but as predictive policing was piloted in more divisions, one captain explained to me that he was starting to feel pressure to use it in his division, because he did not want to be the last to sign on.

10. In a randomized controlled field trial, Mohler and colleagues (2015) found PredPol's algorithm outperforms crime analysts predicting crime, and that police patrols using algorithmic forecasting led to significant reductions in crime volume. The algorithm has not yet been subject to external evaluation; the authors include co-founders and stockholders of PredPol.
11. After protracted negotiations, AVLs were turned on in Central Bureau in March 2015.
12. Place-based algorithms are most effective (and least biased) when predicting crimes with high reporting rates, such as motor vehicle theft.
13. For related work on performativity in a different field—finance—see MacKenzie, Muniesa, and Siu (2007).
14. According to *United States v. Miller* (1939) and *Smith v. Maryland* (1979), the third-party doctrine maintains that "when an individual voluntarily shares information with third parties, like telephone companies, banks, or even other individuals, the government can acquire that information from the third-party absent a warrant" (Executive Office of the President 2014).
15. See Justice Sotomayor's concurring opinion in *United States v. Jones* (2012) and Joh (2016).

References

Angwin, Julia. 2014. *Dragnet Nation: A Quest for Privacy, Security, and Freedom in a World of Relentless Surveillance*. New York: Times Books.

Ball, Kirstie, and Frank Webster, eds. 2007. *The Intensification of Surveillance: Crime, Terrorism & Warfare in the Information Age*. London, UK: Pluto Press.

Barley, Stephen R. 1986. "Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments." *Administrative Science Quarterly* 31(1):78–108.

Barley, Stephen R. 1996. "Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organization Studies." *Administrative Science Quarterly* 41(3):404–441.

Barocas, Solon, and Andrew D. Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104:671–732.

Becker, Howard S. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.

Beckett, Katherine, Kris Nyrop, Lori Pfingst, and Melissa Bowen. 2005. "Drug Use, Drug Possession Arrests, and the Question of Race: Lessons from Seattle." *Social Problems* (52)3:419–41.

Bittner, Egon. 1967. "The Police on Skid-Row: A Study of Peace Keeping." *American Sociological Review* 32(5):699–715.

Bonczar, Thomas P., and Erinn J. Herberman. 2014. "Probation and Parole in the United States, 2013." Washington, DC: Bureau of Justice Statistics.

Bowker, Geoffrey C., and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15(5):662–79.

Braga, Anthony A., and David L. Weisburd. 2010. *Policing Problem Places: Crime Hot Spots and Effective Prevention*. New York: Oxford University Press.

Braverman, Harry. 1974. *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*. New York: Monthly Review Press.

Brayne, Sarah. 2014. "Surveillance and System Avoidance: Criminal Justice Contact and Institutional Attachment." *American Sociological Review* 79(3): 367–91.

Brown v. Plata. 2011. 563 U.S. 493.

Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Durham, NC: Duke University Press.

Carson, E. Ann. 2015. "Prisoners in 2014." Washington, DC: Bureau of Justice Statistics.

Christin, Angèle. 2016. "From Daguerreotypes to Algorithms: Machines, Expertise, and Three Forms of Objectivity." *ACM Computers & Society* 46(1): 27–32.

Cohen, Stanley. 1985. *Visions of Social Control: Crime, Punishment and Classification*. Malden, MA: Polity Press.

Deleuze, Gilles, and Felix Guattari. 1987. *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: The University of Minnesota Press.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48(2):147–60.

Duster, Troy. 1997. "Pattern, Purpose and Race in the Drug War." Pp. 206–287 in *Crack in America: Demon Drugs and Social Justice*, edited by C. Reinarman and H. G. Levine. Berkeley: University of California Press.

Duster, Troy. 2005. "Race and Reification in Science." *Science* 307:1050–51.

Epp, Charles R., Steven Maynard-Moody, and Donald P. Haider-Markel. 2014. *Pulled Over: How Police Stops Define Race and Citizenship*. Chicago: University of Chicago Press.

Ericson, Richard V., and Kevin D. Haggerty. 1997. *Policing the Risk Society*. Toronto: University of Toronto Press.

Ericson, Richard V., and Kevin D. Haggerty. 2006. *The New Politics of Surveillance and Visibility*. Toronto: University of Toronto Press.

Espeland, Wendy N., and Berit I. Vannebo. 2007. "Accountability, Quantification and Law." *Annual Review of Law and Society* 3:21–43.

Executive Office of the President. 2014. "Big Data: Seizing Opportunities, Preserving Values." Washington, DC: The White House.

Feeley, Malcolm M., and Jonathan Simon. 1992. "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications." *Criminology* 30(4):449–74.

Ferguson, Andrew G. 2015. "Big Data and Predictive Reasonable Suspicion." *University of Pennsylvania Law Review* 63(2):327–410.

Fiske, John. 1998. "Surveilling the City: Whiteness, the Black Man and Democratic Totalitarianism." *Theory, Culture and Society* 15(2):67–88.

Fiske, Susan T., and Shelley E. Taylor. 1991. *Social Cognition*, 2nd ed. New York: McGraw-Hill.

Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison*. New York: Random House.

Fourcade, Marion, and Kieran Healy. 2013. "Classification Situations: Life-Chances in the Neoliberal Era." *Accounting, Organizations and Society* 38:559–72.

Fourcade, Marion, and Kieran Healy. 2017. "Seeing like a Market." *Socioeconomic Review* 15(1):9–29.

Gandy, Oscar H. 1993. *The Panoptic Sort: A Political Economy of Personal Information*. Boulder, CO: Westview Press.

Gandy, Oscar H. 2002. "Data Mining and Surveillance in the Post-9.11 Environment." Presentation to the Political Economy Section, International Association for Media and Communication Research.

Gandy, Oscar H. 2009. *Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage*. Farnham, UK: Ashgate Publishing.

Garland, David. 2001. *The Culture of Control: Crime and Social Order in Contemporary Society*. New York: Oxford University Press.

Giddens, Anthony. 1990. *The Consequences of Modernity*. Stanford, CA: Stanford University Press.

Gilliom, John. 2001. *Overseers of the Poor: Surveillance, Resistance, and the Limits of Privacy*. Chicago: University of Chicago Press.

Gitelman, Lisa, ed. 2013. *Raw Data Is an Oxymoron*. Cambridge, MA: MIT Press.

Goffman, Alice. 2014. *On the Run: Fugitive Life in an American City*. Chicago: University of Chicago Press.

Goffman, Erving. 1963. *Stigma: Notes on the Management of Spoiled Identity*. Englewood Cliffs, NJ: Prentice-Hall.

Gross, Samuel R., Maurice Possley, and Kalara Stephens. 2017. "Race and Wrongful Convictions in the United States." National Registry of Exonerations, Newkirk Center for Science and Society, University of California-Irvine.

Gustafson, Kaaryn S. 2011. *Cheating Welfare: Public Assistance and the Criminalization of Poverty*. New York: NYU Press.

Guzik, Keith. 2009. "Discrimination by Design: Predictive Data Mining as Security Practice in the United States' 'War on Terrorism.'" *Surveillance and Society* 7(1):1–17.

Hacking, Ian. 1990. *The Taming of Chance*. Cambridge, UK: Cambridge University Press.

Haggerty, Kevin D., and Richard V. Ericson. 2000. "The Surveillant Assemblage." *British Journal of Sociology* 51(4):605–622.

Harcourt, Bernard E. 2006. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.

Hindmarsh, Richard, and Barbara Prainsack, eds. 2010. *Genetic Suspects: Global Governance of Forensic DNA Profiling and Databasing*. Cambridge, UK: Cambridge University Press.

Innes, Martin. 2001. "Control Creep." *Sociological Research Online* 6:1–10.

Joh, Elizabeth E. 2016. "The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing." *Harvard Law and Policy Review* 10(1):15–42.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London, UK: SAGE.

Kling, Rob. 1991. "Computerization and Social Transformations." *Science, Technology & Human Values* 16(3):342–67.

Kohler-Hausmann, Issa. 2013. "Misdemeanor Justice: Control without Conviction." *American Journal of Sociology* 119(2):351–93.

Laney, Doug. 2001. "3D Data Management: Controlling Data Volume, Velocity, and Variety." Stamford, CT: META Group.

Langton, Lynn, Marcus Berzofsky, Christopher Krebs, and Hope Smiley-McDonald. 2012. "Victimizations Not Reported to the Police, 2006-2010." Washington, DC: Bureau of Justice Statistics.

Laub, John H. 2014. "Understanding Inequality and the Justice System Response: Charting a New Way Forward." New York: William T. Grant Foundation.

Lazer, David, and Jason Radford. 2017. "Data ex Machina: Introduction to Big Data." *Annual Review of Sociology* 43:19–39.

Los Angeles Police Department. 2017. "Sworn Personnel by Rank, Gender, and Ethnicity (SPRGE) Report" (http://www.lapdonline.org/sworn_and_civilian_report).

Lynch, Michael, Simon A. Cole, Ruth McNally, and Kathleen Jordan. 2008. *Truth Machine: The Contentious History of DNA Fingerprinting*. Chicago: University of Chicago Press.

Lyon, David. 1994. *The Electronic Eye: The Rise of Surveillance Society*. Minneapolis: University of Minnesota Press.

Lyon, David, ed. 2003. *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*. New York: Routledge.

Lyon, David, ed. 2006. *Theorizing Surveillance: The Panopticon and Beyond*. New York: Polity.

Lyon, David. 2015. *Surveillance after Snowden*. New York: Polity.

MacKenzie, Donald, Fabian Muniesa, and Lucia Siu, eds. 2007. *Do Economists Make Markets? On the Perfor-*

mativity of Economics. Princeton, NJ: Princeton University Press.
Manning, Peter K. 2011. *The Technology of Policing: Crime Mapping, Information Technology, and the Rationality of Crime Control*. New York: NYU Press.
Manning, Peter K., and John Van Maanen, eds. 1978. *Policing: A View from the Street*. New York: Random House.
Marx, Gary T. 1974. "Thoughts on a Neglected Category of Social Movement Participant: The Agent Provocateur and the Informant." *American Journal of Sociology* 80(2):402–442.
Marx, Gary T. 1988. *Undercover: Police Surveillance in America*. Berkeley: University of California Press.
Marx, Gary T. 1998. "Ethics for the New Surveillance." *The Information Society* 14(3):171–85.
Marx, Gary T. 2002. "What's New About the 'New Surveillance'? Classifying for Change and Continuity." *Surveillance and Society* 1(1):9–29.
Marx, Gary T. 2016. *Windows into the Soul: Surveillance and Society in an Age of High Technology*. Chicago: University of Chicago Press.
Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin Harcourt.
Merton, Robert K. 1948. "The Self-Fulfilling Prophecy." *The Antioch Review* 8(2):193–210.
Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83(2): 340–63.
Mohler, George O., Martin B. Short, Sean Malinowski, Mark Johnson, George E. Tita, Andrea L. Bertozzi, and P. Jeffrey Brantingham. 2015. "Randomized Controlled Field Trials of Predictive Policing." *Journal of the American Statistical Association* 110:1399–1411.
Monahan, Torin, and Neal A. Palmer. 2009. "The Emerging Politics of DHS Fusion Centers." *Security Dialogue* 40(6):617–36.
Moskos, Peter. 2008. *Cop in the Hood: My Year Policing Baltimore's Eastern District*. Princeton, NJ: Princeton University Press.
Oxford American Dictionary of Current English. 1999. Oxford: Oxford University Press.
Pager, Devah. 2007. *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration*. Chicago: University of Chicago Press.
Papachristos, Andrew V., David M. Hureau, and Anthony A. Braga. 2013. "The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence." *American Sociological Review* 78(3):417–47.
Pasquale, Frank. 2014. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
Perry, Walter L., Brian McInnis, Carter C. Price, Susan C. Smith, and John S. Hollywood. 2013. "Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations." RAND Safety and Justice Program. Santa Monica, CA. Retrieved July 6, 2017 (http://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf).
Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
Poster, Mark. 1990. *The Mode of Information*. Chicago: University of Chicago Press.
Quillian, Lincoln, and Devah Pager. 2001. "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime." *American Journal of Sociology* 107(3):717–67.
Ratcliffe, Jerry. 2008. *Intelligence-Led Policing*. Cullompton, UK: Willan Publishing.
Reiss, Albert J. 1971. *The Police and the Public*. New Haven, CT: Yale University Press.
Renan, Daphna. 2016. "The Fourth Amendment as Administrative Governance." *Stanford Law Review* 68(5):1039–1129.
Rios, Victor M. 2011. *Punished: Policing the Lives of Black and Latino Boys*. New York: NYU Press.
Roush, Craig R. 2012. "Quis Custodiet Ipsos Custodes? Limits on Widespread Surveillance and Intelligence Gathering by Local Law Enforcement after 9/11." *Marquette Law Review* 96(1):315–76.
Rule, James B. 1974. *Private Lives and Public Surveillance: Social Control in the Computer Age*. New York: Schocken Books.
Rule, James B. 2007. *Privacy in Peril: How We Are Sacrificing a Fundamental Right in Exchange for Security and Convenience*. New York: Oxford University Press.
Sampson, Robert J., and Dawn J. Bartusch. 1998. "Legal Cynicism and (Subcultural?) Tolerance of Deviance: The Neighborhood Context of Racial Differences." *Law and Society Review* 32:777–804.
Scott, Richard W. 1987. *Organizations: Rational, Natural, and Open Systems*. Englewood Cliffs, NJ: Prentice-Hall.
Scott, Richard W. 2004. "Reflections on a Half-Century of Organizational Sociology." *Annual Review of Sociology* 30:1–21.
Sherman, Lawrence W. 2013. "Targeting, Testing and Tracking Police Services: The Rise of Evidence-Based Policing, 1975–2025." *Crime and Justice* 42(1):377–451.
Sherman, Lawrence W., Patrick R. Gartin, and Michael E. Buerger. 1989. "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place." *Criminology* 27(1):27–55.
Skogan, Wesley G. 2006. *Police and Community in Chicago: A Tale of Three Cities*. New York: Oxford University Press.
Smith v. Maryland. 1979. 442 U.S. 735.
Soss, Joe, Richard C. Fording, and Sanford F. Schram. 2011. *Disciplining the Poor: Neoliberal Paternalism and the Persistent Power of Race*. Chicago: University of Chicago Press.

Stuart, Forrest. 2016. *Down, Out & Under Arrest: Policing and Everyday Life in Skid Row*. Chicago: University of Chicago Press.

Terry v. Ohio. 1968. 392 U.S. 1.

Tracy, Paul E., and Vincent Morgan. 2000. "Big Brother and His Science Kit: DNA Databases for 21st Century Crime Control." *Journal of Criminal Law and Criminology* 90(2):635–90.

Travis, Jeremy, Bruce Western, and Steve Redburn, eds. 2014. *Growth of Incarceration in the United States: Exploring Causes and Consequences*. Washington, DC: National Academy of Science.

Uchida, Craig D., and Marc L. Swatt. 2015. "Operation LASER and the Effectiveness of Hotspot Patrol: A Panel Analysis." *Police Quarterly* 16(3):287–304.

United States v. Jones. 2012. 132 S. Ct. 945, 565 U.S.

United States v. Miller. 1939. 307 U.S. 174.

U.S. Department of Justice. 2001 [2015]. "L.A. Consent Decree." Washington, DC: U.S. Department of Justice.

Wakefield, Sara, and Christopher Uggen. 2010. "Incarceration and Stratification." *Annual Review of Sociology* 36:387–406.

Wakefield, Sara, and Christopher Wildeman. 2013. *Children of the Prison Boom: Mass Incarceration and the Future of American Inequality*. New York: Oxford University Press.

Waxman, Matthew C. 2009. "Police and National Security: American Local Law Enforcement and Counter-Terrorism after 9/11." *Journal of National Security Law and Policy* 3:377–407.

Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology*. Berkeley: University of California Press.

Weisburd, David, Stephen D. Mastrofski, Ann Marie McNally, Rosann Greenspan, and James J. Willis. 2003. "Reforming to Preserve: COMPSTAT and Strategic Problem-Solving in American Policing." *Criminology and Public Policy* 2:421–56.

Western, Bruce. 2006. *Punishment and Inequality in America*. New York: Russell Sage Foundation.

Western, Bruce, and Becky Pettit. 2005. "Black-White Wage Inequality, Employment Rates, and Incarceration." *American Journal of Sociology* 111(2):553–78.

White House Police Data Initiative. 2015. "Launching the Police Data Initiative" (https://obamawhitehouse.archives.gov/blog/2015/05/18/launching-police-data-initiative).

Whren v. United States. 1996. 517 U.S. 806.

Willis, James J., Stephen D. Mastrofski, and David Weisburd. 2007. "Making Sense of COMPSTAT: A Theory-Based Analysis of Organizational Change in Three Police Departments." *Law and Society Review* 41(1):147–88.

Wilson, James Q. 1968. *Varieties of Police Behavior: The Management of Law and Order in Eight Communities*. Cambridge, MA: Harvard University Press.

Sarah Brayne is an Assistant Professor of Sociology and Faculty Research Associate in the Population Research Center at the University of Texas at Austin. Using qualitative and quantitative methods, her research examines the use of "big data" within the criminal justice system as well as the consequences of surveillance for law and social inequality.

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2017 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: 6-20-17

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By:
Emily W. Loughran, Sr. Director of Licensing
Date: 6.2.17

By:
Jason Xu, RIPS Academic Mentor
Date: 6-19-17

By:
Collin Cademartori, RIPS Student
Date: 6-19-17

By: Xi Chen
Xi Chen, RIPS Student
Date: 6-19-17

By: Alistair Letcher
Alistair Letcher, RIPS Student
Date: 6-19-17

By: Jelena Trisovic
Jelena Trisovic, RIPS Student
Date: 6-19-17

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2017 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: 6-20-17

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______

Emily Loughran, Director of Licensing

Date: 6.2.17

By: Matt Haberland

Matt Habeland NSF REU Academic Mentor

Date: 6/19/2017

By: ______

Hao Li, NSF REU Academic Mentor

Date: 6/19/17

By: ______

Osman Akar, NSF REU Student

Date: June 19, 2017

By: ______

Adam Lemuel Dhillon, NSF REU Student

Date: 6/19/17

By: ______

Honglin Chen, NSF REU Student

Date: 06/19/2017

By: ______

Alexander Insuk Song, NSF REU Student

Date: June 19, 2017

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2017 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______________________
Emily Loughran, Director of Licensing
Date: 6.2.17

By: ______________________
Matt Habeland NSF REU Academic Mentor
Date: ______________________

By: ______________________
Hao Li, NSF REU Academic Mentor
Date: ______________________

By: ______________________
Osman Akar, NSF REU Student
Date: ______________________

By: ______________________
Adam Lemuel Dhillon, NSF REU Student
Date: ______________________

By: ______________________
Honglin Chen, NSF REU Student
Date: ______________________

By: ______________________
Alexander Insuk Song, NSF REU Student
Date: ______________________

By: ______________________________

Honglin Chen, NSF REU Student

Date: ______________________________

By: ______________________________

Tiankuang Zhou, NSF REU Student

Date: ______________________________

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2017 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. Release from Liability. Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

Sean Malinowski, Deputy Chief & Chief of Staff

Date: ____________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ____________________
Emily W. Loughran, Sr. Director of Licensing
Date: 6·2·17

By: ____________________
Jason Xu, RIPS Academic Mentor
Date: ____________________

By: ____________________
Collin Cademartori, RIPS Student
Date: ____________________

By: ____________________
Xi Chen, RIPS Student
Date: ____________________

By: ____________________
Alistair Letcher, RIPS Student
Date: ____________________

By: ____________________
Jalena Trisovic, RIPS Student
Date: ____________________

March 8, 2017
Joe Slezak
IBM

Dear Mr. Slezak:

Thank you for your presentations and interest in response to our request for information about video analytics. Given the issues surrounding the non-disclosure agreement, however, we are unable to work with you further on this effort.

We thank you again for your interest.

Sincerely,

Craig D. Uchida	Daniel Gomez
President, JSS	OIC, ITB

Deputy Chief Bob Green, LAPD

Visit to UCLA on 22 February 2017
Location: Fowler Building Room A103B
Contact: Jeff Brantingham, 424-298-7732

Context

The course is a general introduction to criminology. Students are doing an open-ended research project where they are tasked with "rewriting the rules of police engagement". The goal is to get them to see how complex and challenging the outside world really is.

My hope is for Chief Green to provide students with a window into what policing is really like. Most of the students will have never interacted with a police officer, meaning that their knowledge about policing is filtered through Hollywood and selective media sources. An ideal outcome is for them to question what they think they know.

Potential Talking Points

- About the LAPD
  - Area policed & population size served
  - Number of sworn officers
  - Average number of patrol officers on the street at any one time
  - What are LAPD patrol officers like?
    - age, gender, race-ethnicity
- Policing from the point of view of the patrol officer
  - What are the some of the main things that police patrol units do every day?
    - Most students in the class are going to conceive of all policing as strictly hard-charging to get the bad guy. What is wrong with this picture?
  - Walk us through the general decision making process of police officers for a type of encounter (some possible examples):
    - A traffic stop
    - A domestic dispute
    - An altercation between people on the street
- Risk
  - What are the greatest sources of risk for police officers when they are on patrol?
  - How is the perception of risk by officers tied to:
    - guns in the hands of the public?
    - alcohol or substance abuse by the people they are contacting?
    - mental health issues?
  - How much does a patrol officer typically know about what is happening when they arrive at a call?

    - Are all types of calls equally likely to pose a risk to officers and the public, or are there some types of calls that are greater risk than others?
- Statistics relevant to risk?
    - How many public contacts does LAPD make in a year?
        - In 2015, the Department had 1,503,758 public contacts (from Use of Force End of Year Review).
    - How many public contacts resulted in use of force?
        - In 2015, there were 1,924 use of force incidents, 0.13% of the Department's total public contacts.
    - How many public contacts resulted in an officer involved shooting?
        - In 2015, there were 48 officer involved shootings, 0.003% of all public contacts
    - In what percentage of OIS incidents was the individual armed?
        - In 2015, 42 of 48 OIS incidents, or 87.5%
    - How many illegal guns does LAPD recover in a year?
        -
- Use of Force
    - How is use of force defined?
    - How does LAPD handle use of force?

- Trust & the Community
    - Why does community trust matter so much?
    - How is LAPD committed to community trust?

IPAM 2017 RIPS-LAPD Project

Conversational Turn-Taking in Police Body-Worn Video

Industry Sponsor: Deputy Chief Sean Malinowski (LAPD Chief of Staff); Sgt. Dan Gomez, Mr. Arnold Suzukamo (LAPD-IT Bureau).

Academic Mentor:

Academic Supervisors: Jeff Brantingham, UCLA Anthropology; Dr. Craig Uchida, Justice & Security Strategies

Introduction

Body-worn video (BWV) or on-body cameras provide a novel means to collect very fine-information about police-public interactions. The general use model requires officers to initiate recording of video whenever there is an encounter with a member of the public. During such interactions, BWV is recorded in real-time. Recording is terminated at the officer's discretion. BWV is not streamed or reviewed in real-time, but rather is uploaded to a secure cloud storage system at the end of an officer's shift.

BWV is designed to provide another line of evidence for the actions of individuals and the outcomes of interactions between police and members of the public. BWV is therefore evidence relevant to legal proceedings like any other form of evidence collected by police. In a limited number studies, BWV has been shown to reduce the likelihood that situations escalate to a point requiring use of force.

There are considerable challenges facing wide-spread use of BWV. Even small scale deployments are expected to lead to massive volumes of video data that will quickly outstrip the ability of law enforcement agencies to analyze. The resulting fallback position will be to review BWV footage only when it corresponds to adverse outcomes (e.g., use of force). Most video will go unused. Many of the potential benefits of BWV may therefore go unrealized.

The 2017 LAPD-RIPS Project

The 2017 RIPS-LAPD team will work to develop methods for the automatic discrimination and labeling of audio-video segments into the following categories: (1) the focal police officer speaking; (2) other actors speaking; and (3) overlapping speech involving the focal officer and others. The focal police officer is defined as the officer wearing the camera. The goal is not speech content recognition, or transcription. Rather we wish to identify when police officers exclusively are speaking relative to one or more other actors in a video scene and when the officer and others are trying to override one another with speech. Measures of conversational turn taking may then be computed. Conversational turn taking may provide evidence of when interactions are escalating or de-escalating without specific knowledge of the content of speech. Understanding when interactions escalate and de-escalate can be of tremendous value in helping to minimize the risk of adverse outcomes in police-public interactions.

The project will rely on a range of data types BWV metadata (e.g., time stamps), BWV audio, and the video images themselves. Computations may be done in Matlab, Mathematica, C, C++, R, Java, or other appropriate computational language.

Key Milestones:

1. Statistical assessment of LAPD BWV and other associated data.
2. Develop speech segmentation methods.
3. Measuring conversational turn taking.
4. Testing of efficacy of methods.
5. Present to LAPD.

References

Ariel, Barak, WilliamA Farrar, and Alex Sutherland. 2014. The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology*:1-27.

Kim, Samuel, Sree Harsha Yella, and Fabio Valente. 2012. "Automatic detection of conflict escalation in spoken conversations." Interspeech September, 1167-1170.

Merkurjev, Ekaterina, Justin Sunu, and Andrea L. Bertozzi. "Graph MBO method for multiclass segmentation of hyperspectral stand-off detection video." Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014.

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
UCLA
Institute for Pure and Applied Mathematics
Dr. P. Jeffrey Brantingham

(Hereafter "Requestor")

The undersigned hereby agrees to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from video using machine learning algorithms. Researchers will examine video footage from LAPD to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestor, and that material is to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestor hereby agrees that all Protected Confidential Materials to which he is given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestor agrees not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestor further agrees to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestor acknowledges the confidential nature of the Protected Confidential Material supplied by the LAPD, and agrees that disclosure by the Requestor or any individual or group of individuals at the request or direction of the Requestor to anyone not directly identified in this Agreement is strictly prohibited.

Importantly, the Requestor assures that data identified to a specific individual will not be revealed under any circumstances and that the information is being used for research and statistical purposes only.

Project findings and reports will not contain information about individuals or private persons.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestor shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestor shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestor' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestor.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestor under this Agreement, Requestor agrees to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestor.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestor agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestor and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestor will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestor acknowledges that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestor agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestor, its use by Requestor, or any and all resulting analyses or conclusions derived from the Materials. Requestor shall indemnify and hold the City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestor.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT	**FOR REQUESTOR**
______________________________	By: ______________________________
MAGGIE GOODRICH, Chief Information Officer Commanding Officer Information Technology Bureau	P. Jeffrey Brantingham, Ph.D. University of California, Los Angeles
Date: ______________________________	Date: ______________________________

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016 ("UCLA IPAM RIPS")
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from video using machine learning algorithms. Researchers will examine video footage from LAPD to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. “Dissemination” as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor’s services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
MAGGIE GOODRICH, Chief Information Officer
Commanding Officer
Information Technology Bureau

Date: ________________________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ______________________________
Emily Loughran, Director of Licensing
Date: ______________________________

By: ______________________________
Giang Tran, RIPS Academic Mentor
Date: ______________________________

By: ______________________________
Stephanie Allen, RIPS Student
Date: ______________________________

By: ______________________________
David Madras, RIPS Student
Date: ______________________________

By: ______________________________
Ye Ye, RIPS Student
Date: ______________________________

By: ______________________________
Greg Zanotti, RIPS Student
Date: ______________________________

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016 (“UCLA IPAM RIPS”)
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department (“LAPD”), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, “Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: ______________________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ____________________________
Emily Loughran, Director of Licensing
Date: ____________________________

By: ____________________________
Jason Xu, RIPS Academic Mentor
Date: ____________________________

By: ____________________________
Collin Cademartori, RIPS Student
Date: ____________________________

By: ____________________________
Xi Chen, RIPS Student
Date: ____________________________

By: ____________________________
Alistair Letcher, RIPS Student
Date: ____________________________

By: ____________________________
Jalena Trisovic, RIPS Student
Date: __________________________

CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016 (“UCLA IPAM RIPS”)
(Hereafter "Requestors")

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department (“LAPD”), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, “Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. “Dissemination” as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor’s services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: __________________________________

FOR THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, INSTITUTE FOR PURE AND APPLIED MATHEMATICS, RESEARCH IN INDUSTRIAL PROJECTS FOR STUDENTS 2016

By: ______________________________
Emily Loughran, Director of Licensing
Date: ______________________________

By: ______________________________
Jason Xu, RIPS Academic Mentor
Date: ______________________________

By: ______________________________
Collin Cademartori, RIPS Student
Date: ______________________________

By: ______________________________
Xi Chen, RIPS Student
Date: ______________________________

By: ______________________________
Alistair Letcher, RIPS Student
Date: ______________________________

By: ______________________________
Jalena Trisovic, RIPS Student
Date: ____________________________

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from video using machine learning algorithms. Researchers will examine video footage from LAPD to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control, including any and all copies (whether electronic or non-electronic), to the Los Angeles Police

Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. “Dissemination” as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor’s services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
MAGGIE GOODRICH, Chief Information Officer
Commanding Officer
Information Technology Bureau

Date: __________________________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ________________________________
Emily Loughran, Director of Licensing
Date: ________________________________

By: ________________________________
Matt Haberland, NSF REU Academic Mentor
Date: ________________________________

By: ________________________________
Alicia Figueroa, NSF REU Student
Date: ________________________________

By: ________________________________
Deborah Tonne, NSF REU Student
Date: ________________________________

By: ________________________________
Yun Liu, NSF REU Student
Date: ________________________________

By: ________________________________
Benjamin Lu, NSF REU Student
Date: ______________________________

<u>**CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT**</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. “Dissemination” as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor’s services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: __________________________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______________________________
Emily Loughran, Director of Licensing
Date: ______________________________

By: ______________________________
Matt Habeland NSF REU Academic Mentor
Date: ______________________________

By: ______________________________
Hao Li, NSF REU Academic Mentor
Date: ______________________________

By: ______________________________
Osman Akar, NSF REU Student
Date: ______________________________

By: ______________________________
Adam Lemuel Dhillon, NSF REU Student
Date: ______________________________

By: ______________________________
Honglin Chen, NSF REU Student
Date: ______________________________

By: ______________________________
Alexander Insuk Song, NSF REU Student
Date: ______________________________

By: ________________________________

Honglin Chen, NSF REU Student

Date: ______________________________

By: ________________________________

Tiankuang Zhou, NSF REU Student

Date: ______________________________

<u>CONFIDENTIALITY AND NONDISCLOSURE AGREEMENT</u>

Between

THE LOS ANGELES POLICE DEPARTMENT
And
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016 ("UCLA MATH DEPT NSF REU")
(Hereafter "Requestors)

The undersigned hereby agree to the following as conditions to the receipt and utilization of data from the Los Angeles Police Department ("LAPD"), for the purpose of assisting the LAPD with analyzing video footage. This project is titled, "Analyzing Body-Worn Camera Video in the Los Angeles Police Department". The purpose of this project is to identify specific features from body worn video footage using machine learning algorithms. Researchers will examine video and audio from LAPD body worn video footage to determine specific interactions between the police and the public.

1. Definitions

A. "Protected Confidential Material" includes all written information, whether originals or copies, including but not limited to reports, documents, notes, interviews, electronically stored data, photographs, charts or any other information supplied by the LAPD to Requestors, and it to be treated as non-public and protected from disclosure or dissemination, in accordance with the terms of this Agreement.

2. Treatment and Use of Protected Confidential Material. Requestors hereby agree that all Protected Confidential Materials to which they are given access shall remain the property of the City of Los Angeles. Such materials shall be used only for the Project and shall not be used for any other purpose not described in this Agreement. Requestors agree not to copy, disseminate, or allow access to any Protected Confidential Material.

Requestors further agree to secure any Protected Confidential Material received from the LAPD in such a way that unauthorized persons or entities cannot retrieve the information by any means, including but not limited to access via computer, remote terminal, or by any other electronic or non-electronic means.

Requestors acknowledge the confidential nature of the Protected Confidential Material supplied by the LAPD, and agree that disclosure by the Requestors or any individual or group of individuals at the request or direction of the Requestors to anyone not directly identified in this Agreement is strictly prohibited.

3. Return of Protected Confidential Materials. Upon completion of the Project, Requestors shall immediately return all Protected Confidential Material in their possession or control,

including any and all copies (whether electronic or non-electronic), to the Los Angeles Police Department. Requestors shall certify in writing that all originals and copies of the material provided under this Agreement have been returned.

4. Monitoring of Compliance and Demand for Document Return. The LAPD may monitor, audit and review the Requestors' program activities and policies to ensure compliance with the requirements and conditions of this Agreement. If the LAPD determines that the requirements and conditions of this Agreement are not being satisfactorily met, it may require the immediate return of all copies of the Protected Confidential Material obtained under this Agreement, take such action as deemed appropriate to protect the security and privacy of this material, and refuse any future requests for information or records from the Requestors.

5. Protection of Personal Identifying Information. In order to protect the identities of any persons whose records are supplied to the Requestors under this Agreement, Requestors agree to:

A. Use the Protected Confidential Material furnished under this Agreement only for the purpose described by Requestors.

B. Replace the name and other personal identifiers with an alphanumeric or other appropriate code for purpose of conducting the necessary project analyses;

C. Restrict access of all data supplied by LAPD to those individuals whose responsibilities cannot be accomplished without such access; and

D. Store all Protected Confidential Material received from LAPD in secure locked containers.

6. Project Treatment. Requestors agree to insert into the preface of any report citing data analysis conducted on any of the Protected Confidential Material, a disclaimer that the analysis and report are solely the work product of the Requestors and do not represent the position or conclusions of the Los Angeles Police Department.

At the conclusion of the Project, Requestors will provide the LAPD with a copy of any written report derived from the Project. LAPD shall retain the discretion to use the report for whatever purpose or further analysis it deems appropriate.

Requestors acknowledge that any written or oral report generated pursuant to analysis of any of the Protected Confidential Material is not to be published or circulated in any manner other than as explicitly set forth under this Agreement. The City retains sole authority to approve disseminating to individuals, agencies, organizations or entities not parties to this agreement specific information regarding the services, reports, Deliverables and other materials resulting from this Agreement. "Dissemination" as used in this section includes, but is not limited to printed and online articles, reports or publications, and public relations and advertising materials for Requestor's services or participation under this Agreement.

7. **Release from Liability.** Requestors agree that the City of Los Angeles and any of its agents or employees shall not be liable for any acts or omissions arising from the production of the Protected Confidential Material to Requestors, its use by Requestors, or any and all resulting analyses or conclusions derived from the Materials. Requestors shall indemnify and hold the

City of Los Angeles and its employees and officers harmless for any and all claims, lawsuits, causes of action, damages or costs incurred in any adjudication or settlement of claims, including attorney's fees and costs, which may arise from any alleged use or misuse of documents provided by the LAPD pursuant to this Agreement, or by any negligent or willful act or omission on the part of Requestors.

This Agreement will become effective upon signature of the parties.

I/We hereby agree to all conditions and requirements set forth in this Agreement:

FOR THE LOS ANGELES POLICE DEPARTMENT

____
Sean Malinowski, Deputy Chief & Chief of Staff

Date: __________________________________

THE REGENTS OF THE UNIVERSITY OF CALIFORNIA LOS ANGELES, DEPARTMENT OF MATHEMATICS, NATIONAL SCEINCE FOUNDATION, RESEARCH EXPERIENCE FOR UNDERGRADUATES 2016

By: ______________________________
Emily Loughran, Director of Licensing
Date: ______________________________

By: ______________________________
Matt Habeland NSF REU Academic Mentor
Date: ______________________________

By: ______________________________
Hao Li, NSF REU Academic Mentor
Date: ______________________________

By: ______________________________
Osman Akar, NSF REU Student
Date: ______________________________

By: ______________________________
Adam Lemuel Dhillon, NSF REU Student
Date: ______________________________

By: ______________________________
Honglin Chen, NSF REU Student
Date: ______________________________

By: ______________________________
Alexander Insuk Song, NSF REU Student
Date: ______________________________

By: ________________________________

Honglin Chen, NSF REU Student

Date: ______________________________

By: ________________________________

Tiankuang Zhou, NSF REU Student

Date: ______________________________

Conversational Turn-Taking in Police Body-Worn Video

Alistair Letcher, Jelena Trišović, Collin Cademartori, Xi Chen

Academic Mentor: Jason Xu; Academic Supervisor: Jeff Brantingham

Motivation

- The Los Angeles Police Department
- Police body-worn video (BWV)
  - Turned on for every interaction with the public
  - Evidence for the actions of police officers and individuals
  - Behavioural impact: reduce police use of force and assaults against officers
- High volume of data
  - Hours of video to review for each court case
  - Can we automatically detect the segments of importance?

The problem

Automatic discrimination and labelling of BWV audio segments into the following categories:

1. The focal police officer speaking.
2. Other actors speaking.
3. Overlapping speech involving the focal officer and others.

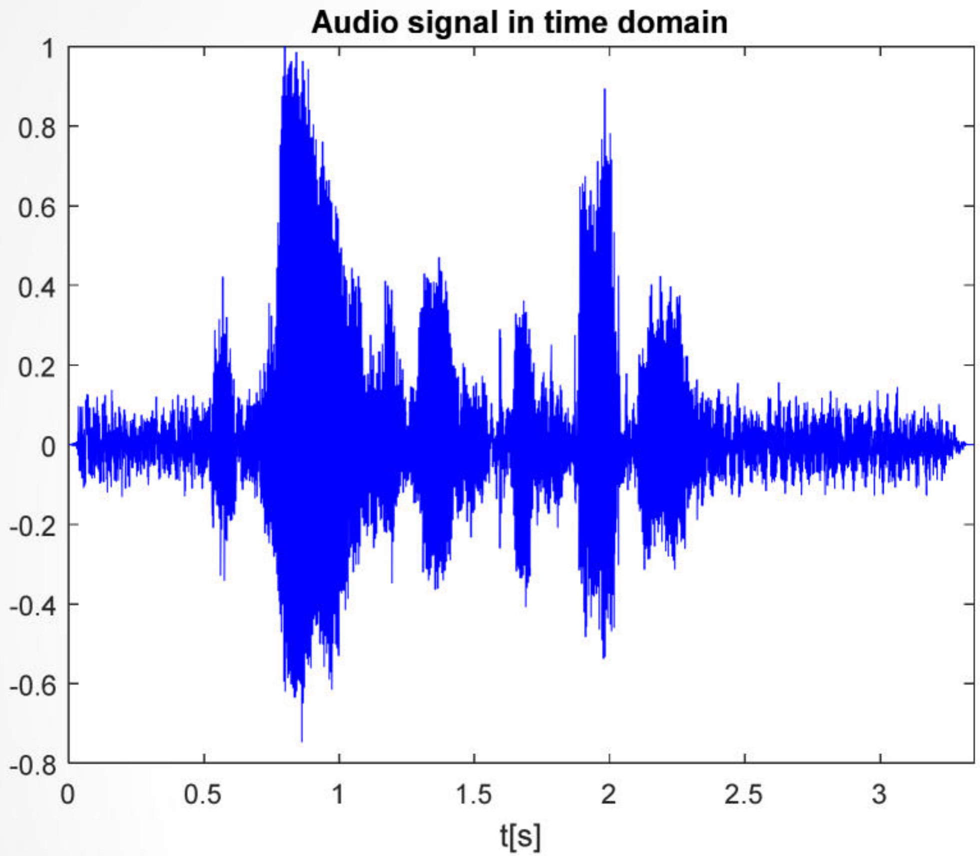
→ Automatically recognize events like escalation and conflict, based on speech overlap and turn-taking.

Overview of the project

- Adaptive denoising of the audio signal
- Segmentation into speech/nonspeech
  - Feature selection and optimisation
  - Unsupervised clustering
  - Supervised learning

- Future work
  - Isolating the police officer's voice
  - Problems with overlap detection and turn-taking
  - Other approaches to evaluate conflict escalation and detection

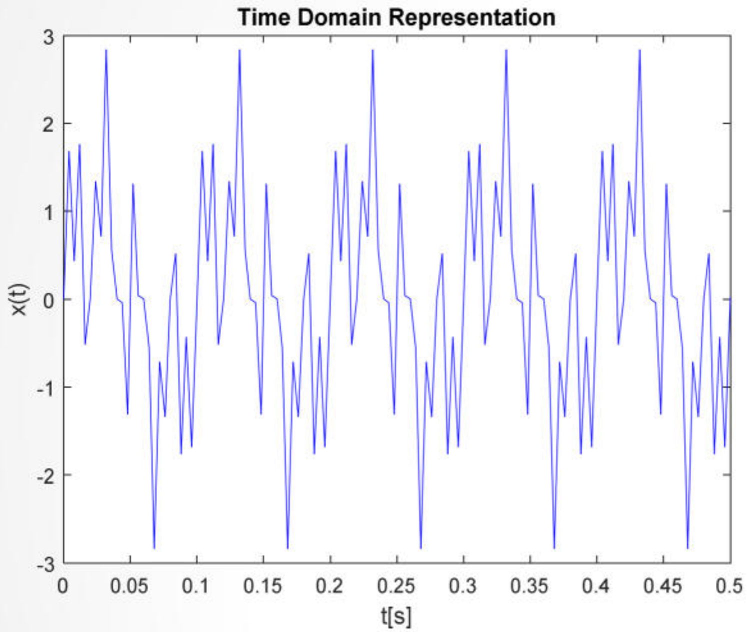
Audio Signal Representation



- A waveform which evolves over time
- Magnitude is proportional to volume

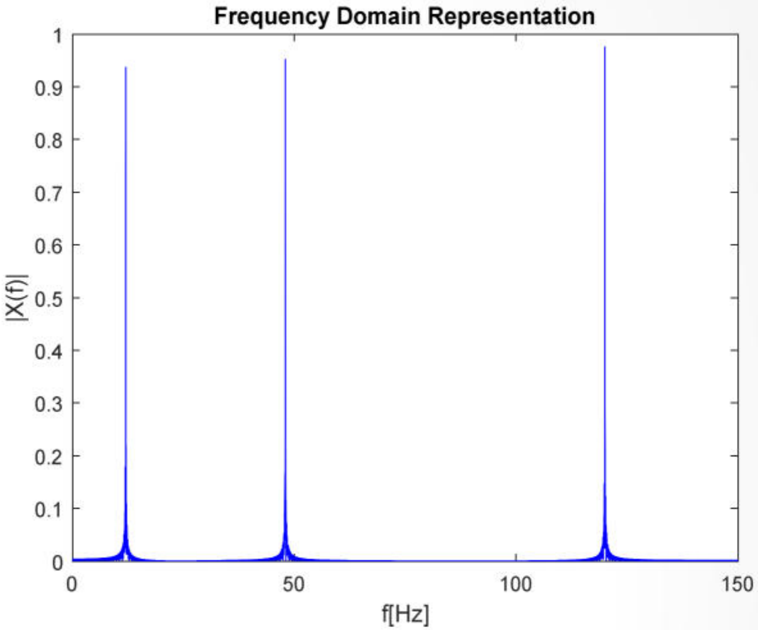
Time Domain VS Frequency Domain

Time Domain - x(t)



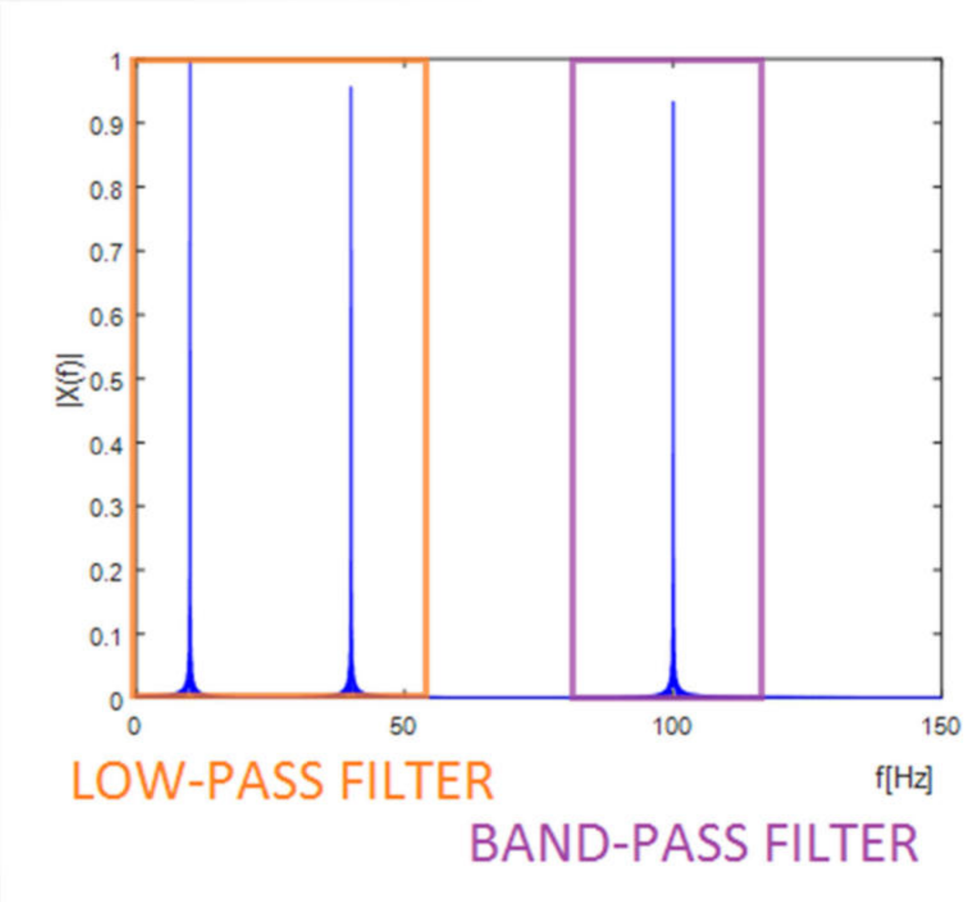
$\rightarrow Fourier\ transform \rightarrow$

Frequency Domain - X(f)



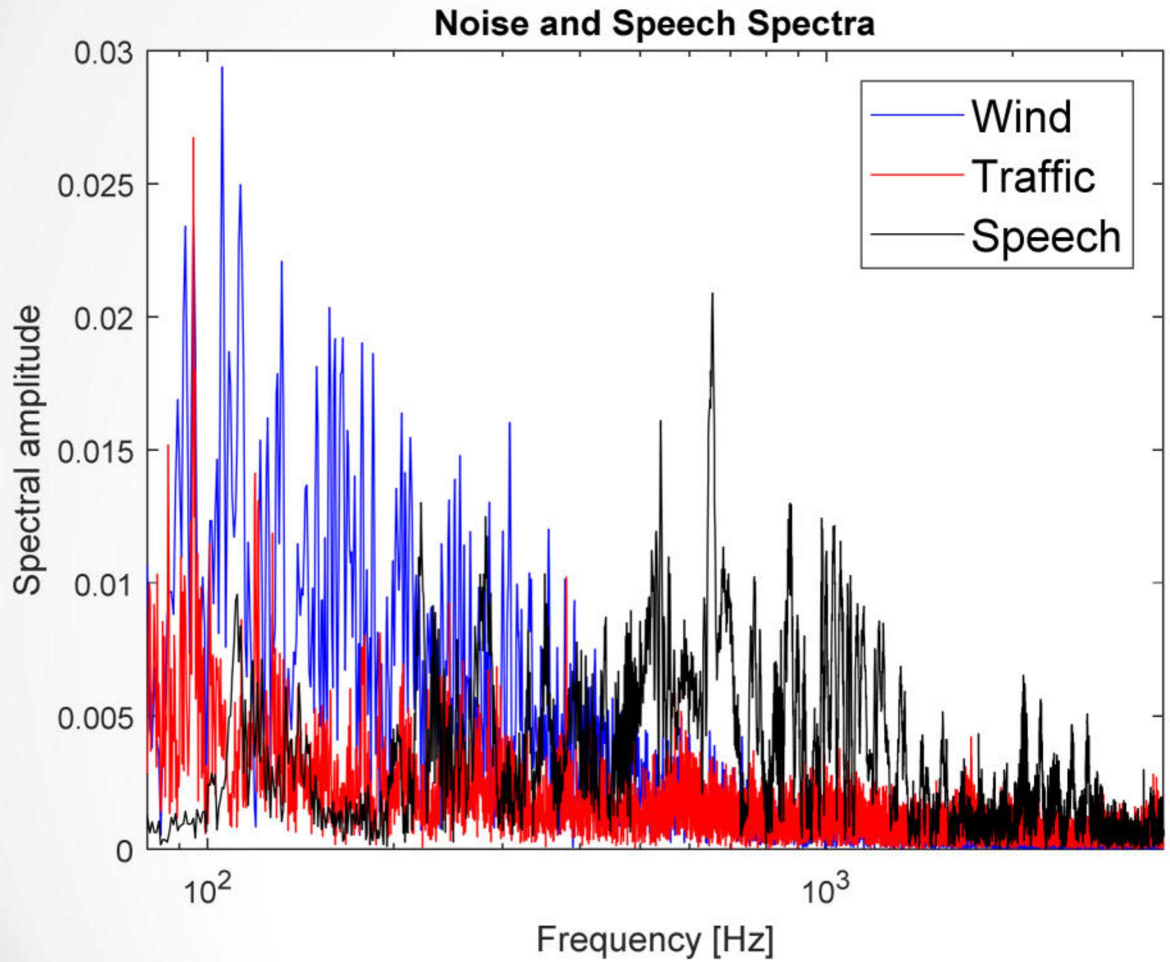
$$x(t) = \sin(2\pi 10 \cdot t) + \sin(2\pi 40 \cdot t) + \sin(2\pi 100 \cdot t)$$

Standard Filtering Methods



- Filter is a device or process which removes unwanted components or features from the signal
- Usually, filters remove whole frequency bands (low-pass filter, high-pass filter, band-stop filter…)

Why is traffic and wind noise difficult to filter?



We cannot apply an ordinary low-pass or band-pass filter!

We have to use a more complicated way of filtering – an adaptive filter!

$y[n] = y(nT_s)$

The Problem

$$y[n] = x[n] + e[n]$$

$y[n]$ – observation

$e[n]$ – noise

$x[n]$ – original signal

We know y, how do we get x?

Wiener Filter

$$\hat{x}(n) = \sum_{i=1}^{N} a_i \, y(n-i)$$

- $\hat{x}(n)$ – estimate of $x(n)$
- $a_i,\ i \in \{1,2,\dots,N\}$ – filter coefficients

Why is the filter adaptive?

The filter coefficients are those values of $a_i,\ i \in \{1,2,\dots,N\}$ that minimize the following function:

$$J = E\left(\left(x(n) - \hat{x}(n)\right)^2\right)$$

Wiener Filter

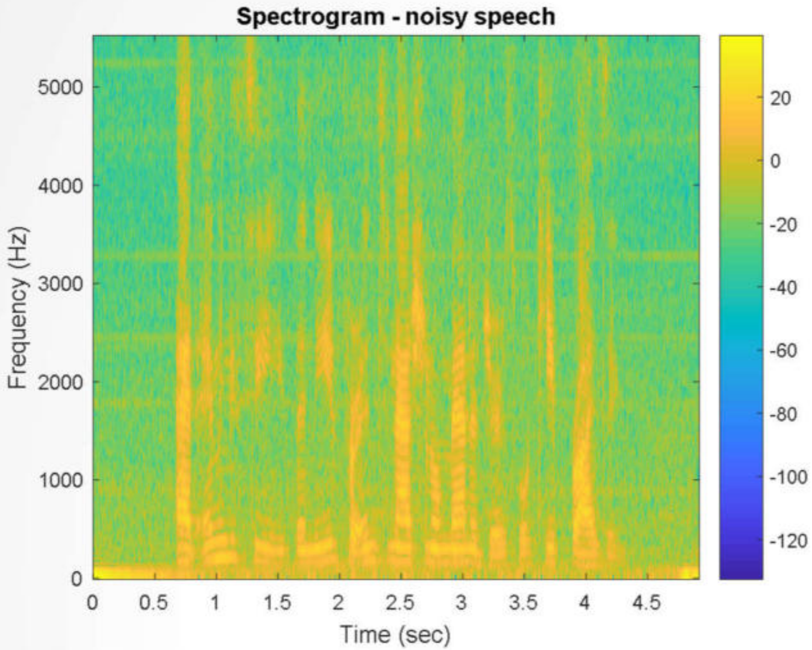
$$J = E((x(n) - \hat{x}(n))^2) = E\left(\left(y(n) - e(n)\right) - \left(\sum_{i=1}^{N} a_i\, y(n-i)\right)\right)^2$$

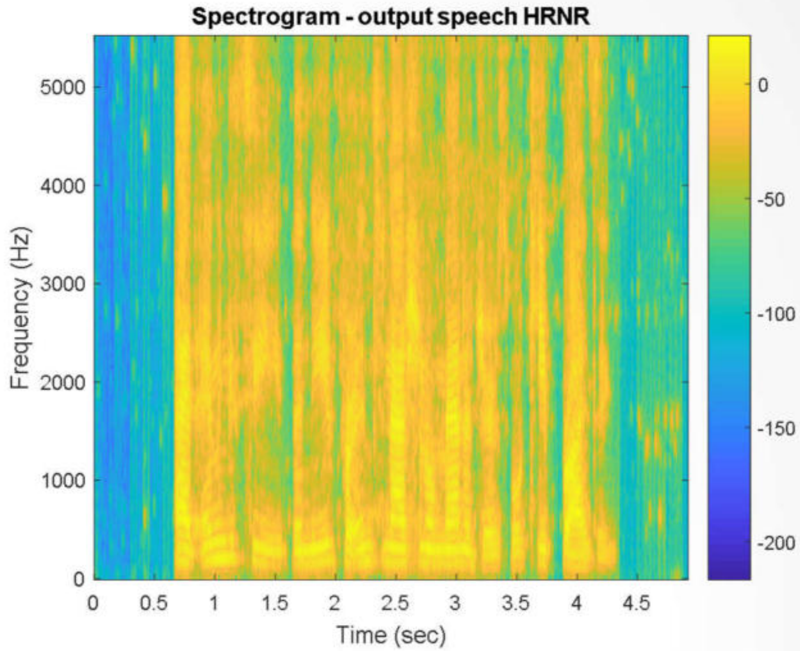
- $y(n)$ is the observation
- $a_i,\ i \in \{1,2,\dots,N\}$ are parameters
- $e(n)$ is unknown, but we can make an estimate of it!

Denoising Algorithm

- Using Gaussian Mixture Model, a short sample of noise is found in the initial audio signal and provided to the Wiener filter
- The initial noise parameters are estimated
- Filtering is performed in two phases:
  - Two-Step Noise Reduction (TSNR) – noise removal with double Wiener filtering
  - Harmonic Regeneration Noise Reduction (HRNR) – reconstruction of the speech components which were damaged in the TSNR phase

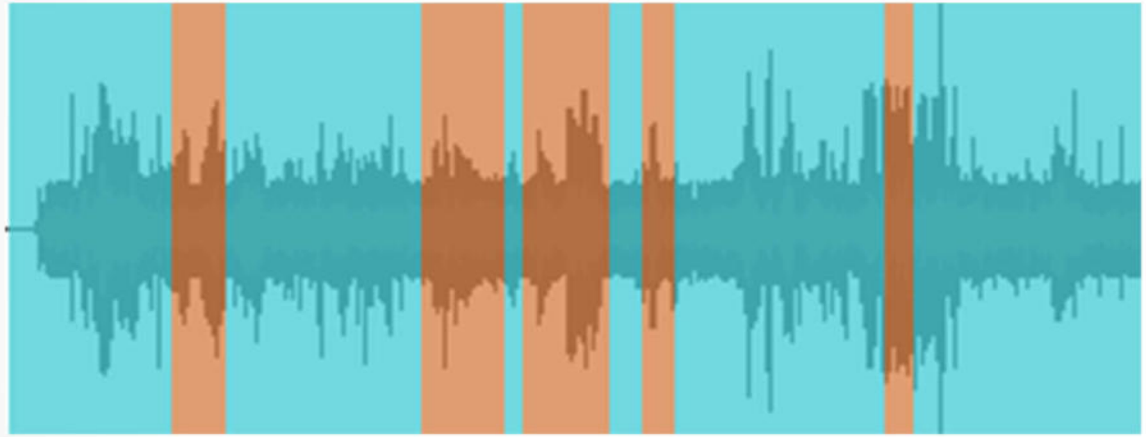
Results


Spectrogram - noisy speech
Frequency (Hz)
0
1000
2000
3000
4000
5000
Time (sec)
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
20
0
-20
-40
-60
-80
-100
-120


Spectrogram - output speech HRNR
Frequency (Hz)
0
1000
2000
3000
4000
5000
Time (sec)
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
0
-50
-100
-150
-200

Representing the Audio Signal

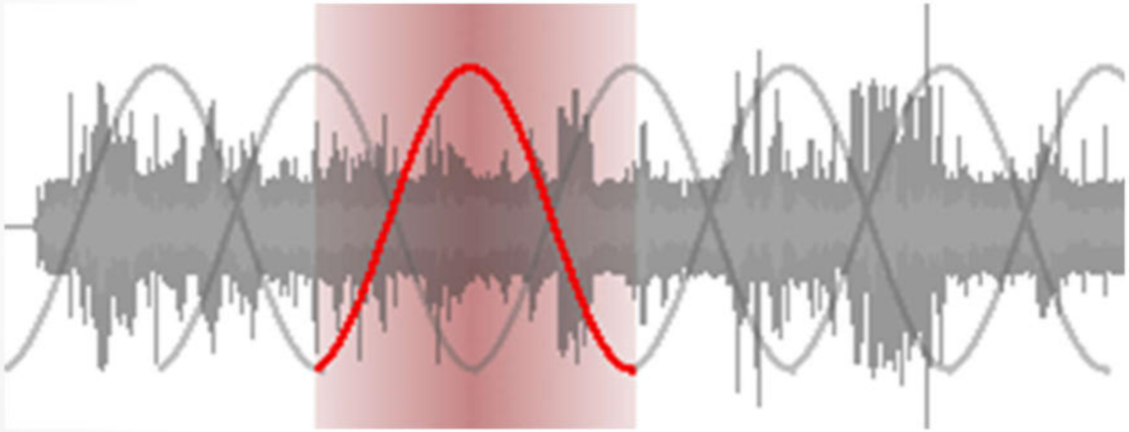
- Original audio signal just a sequence of amplitudes
- This form obscures the changing composition of the signal
- Makes distinguishing speech from non-speech impossible



Orange regions = speech
Blue regions = non-speech

Representing the Audio Signal

- Compute features that capture changing signal composition
- Break signal up into small overlapping windows
- In each window, reweight signal by a windowing function
  - Emphasizes the middle portion
  - Removes artificial jumps at the endpoints
- New features are computed from these weighted amplitudes



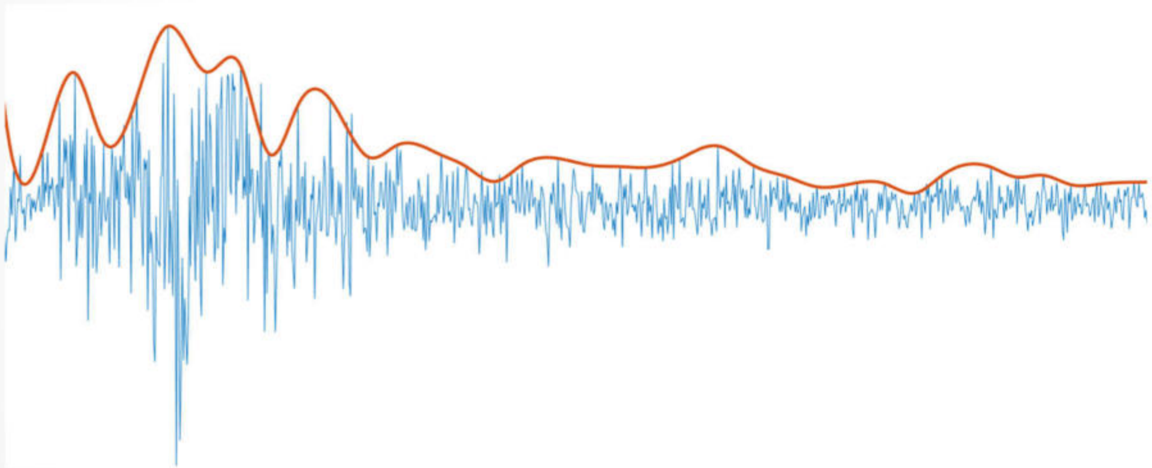
A segmentation of the audio signal into Hamming windows

Representing the Audio Signal

- Most important features are the Mel-Frequency Cepstral Coefficients
- First look at the spectrum of the signal within a window
- Want to modify spectrum so that we see the data the way the ear hears it
- The ear is especially sensitive to low frequencies
  - Perceives larger gaps between them than between higher frequencies
- Apply transformation to convert to Mel-frequencies
  - Equal sized gaps on this scale are perceived as being equal by the human ear

Representing the Audio Signal

- Need to convert the Mel spectrum into a finite number of features
- Idea: apply a discrete Fourier transform to the Mel spectrum
- Take only the low frequency coefficients in this new "spectrum"
  - Ignores components that change quickly, smoothing out spectrum
  - Summarizes the broad trends with only a few coefficients



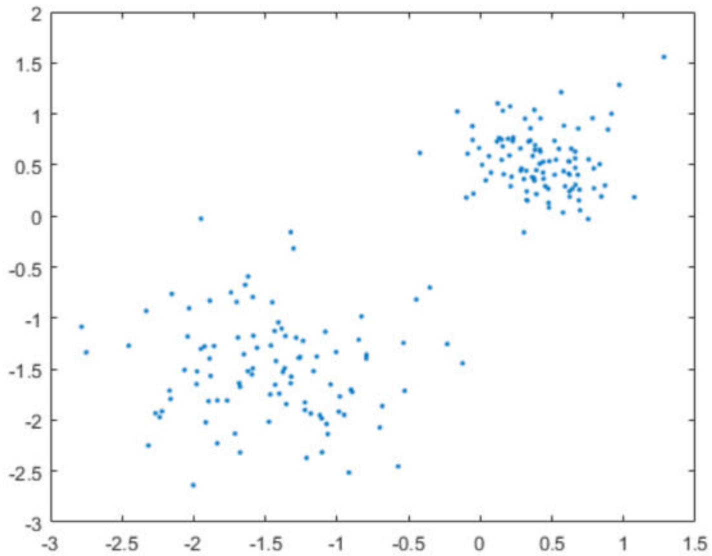
The Fourier transform of a signal and an envelope curve consisting of lower frequency components

Representing the Audio Signal

- Stable properties of human voice come out over longer time scales
- Again split up sequence into overlapping windows
- Compute mean and variance of short-term features within window
- This achieves two goals
  - Decomposes signal into pieces that can characterize human voice
  - Tracks longer term trends in the evolution of those features

Representing the Audio Signal

- Want to test these features' ability to distinguish speech and non-speech
- Idea: try to identify two natural clusters within the data
  - Check if speech and non-speech separated
- Use k-means algorithm for clustering
- Completely unsupervised
  - Algorithm had no examples of speech or non-speech
- Roughly separated speech and non-speech
- Not sufficiently robust to our noisy environments



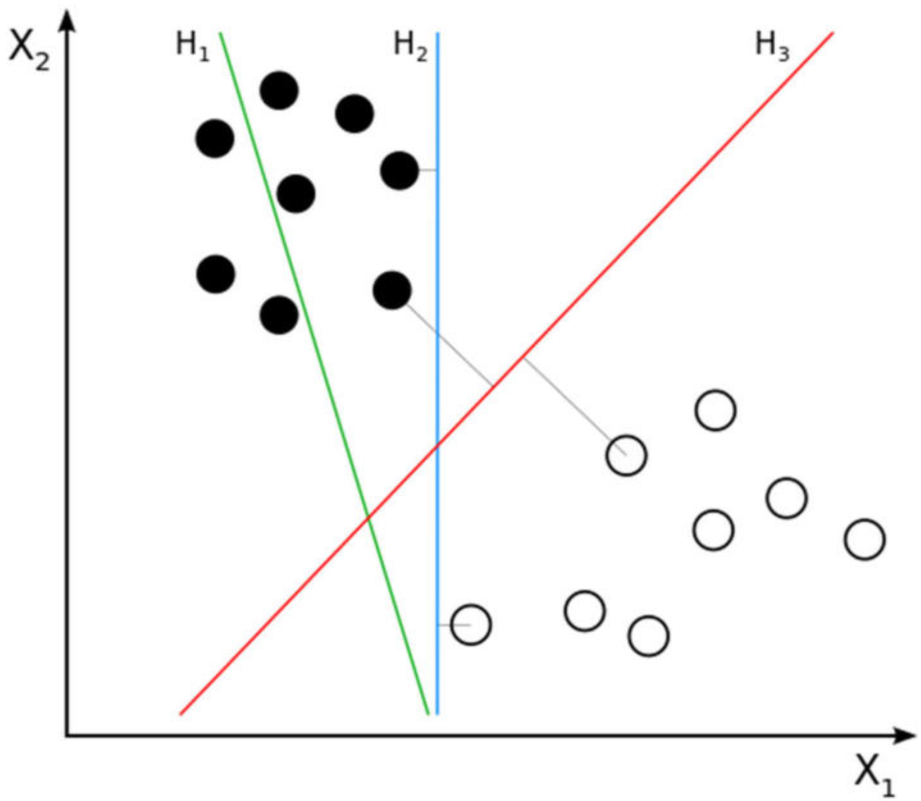
An example of data clearly separated into two clusters in this coordinate representation

Speech/Non-Speech Detection

- Supervised learning: use training (labelled) data to predict labels for unseen test data
- Methods that we used
  - Support Vector Machines
  - Gaussian Mixture Models

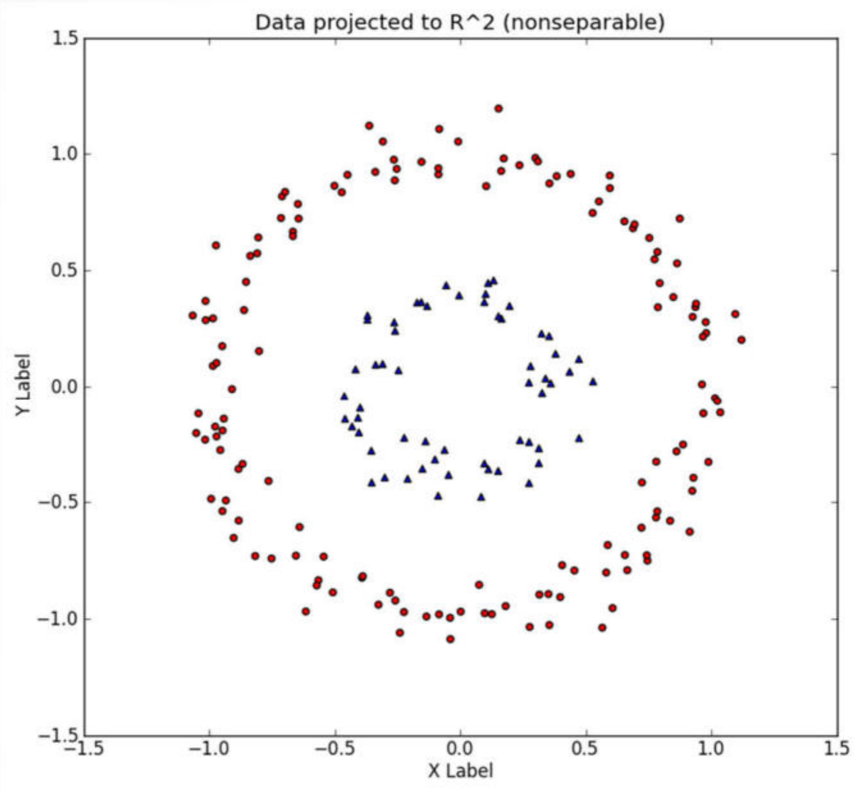
Support Vector Machine

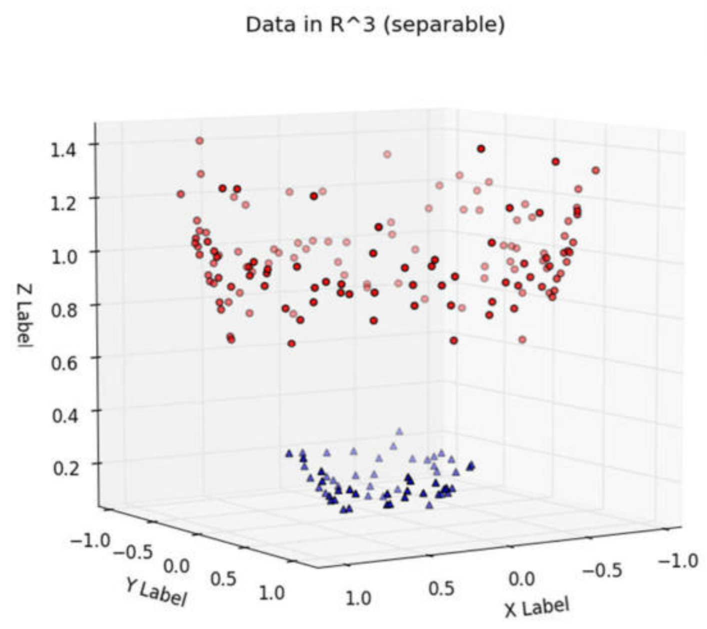
- Construct a maximum margin hyperplane that separates the two classes



Kernel Trick

- Sometimes training points are not linearly separable
- Map the input points to a higher dimension space to find a linear boundary





$$\phi((a, b)) = (a, b, a^2 + b^2)$$

Gaussian Mixture Model

- Probabilistic model that assumes all data points are generated from a mixture of Gaussians with unknown parameters
- We use the Expectation–maximization algorithm to train two Gaussian Mixture Models, one for speech and one for non-speech
- Given a new data point, compare the likelihoods that it fits the two models and label it accordingly

Cross Validation Test

Results

	False Alarm	Misdetection
SVM with linear kernel	0.0615	0.0435
SVM with polynomial kernel	0.0280	0.0286
SVM with RBF kernel	0.0305	0.0317
Gaussian Mixture Model	0.0437	0.0720

Future work

- Identify the police officer
  - Clustering and MFCC's alone are insufficient. Use prosodic features and Hidden Markov Models?
  - Use volume to our advantage.
  - Train our SVM to differentiate between radio and non-radio.
- Conflict escalation/detection
  - Issues with conversational turn-taking and overlap, both technical and based on empirical evidence.
  - Other approaches: tone, speech recognition (keywords and repetition), relative volume.

Questions?

#	Bkg # / Date / Time / Arresting Division / Investigating Unit / Arrest Detail	RD Arrested / Arrest Location / Arrest Type / Arrest Group / Arrest Charge	Last Name, First Name / AKA's / Sex Desc Hair Eyes Ht Wt Age / Arrestees Address	A - Veh (Yr Mk Mod Sty Top Bot Lic St) / Driver's License / Clothing / Personal Descriptors
	Bkg #			
	Date / Time	RD Arrested		
	Arresting Division	Arrest Location	Last Name, First Name	A - Veh (Yr Mk Mod Sty Top Bot Lic St)
	Investigating Unit	Arrest Type	AKA's	Driver's License
	Arrest Detail	Arrest Group	Sex Desc Hair Eyes Ht Wt Age	Clothing
#		Arrest Charge	Arrestees Address	Personal Descriptors
	4873230			
	01/04/2017 0910	1412		
	4214	PARK & OCEAN FRONT WK	MINNEY, JEREMY	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	FB	NARCOTIC DRUG LAWS	M W BLN BLU 602 231 39	ORG SHT,BLK PNT,BLK SNDLS
###		11377(A)HS	204 HAMPTON DR	?
	4873628			
	01/04/2017 1600	1415		
	4214	893 WARREN AV	BOYER, MICAH	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	A	MISC OTHER VIOLS	M W BRO BLU 602 180 40	BLU PNT,BLK JKT,WHT SHS
###		853.7PC	1942 TRANSIENT	?
	4214			[redacted]
	14	1443		GRY SWEATR,BLU PNTS,WHT SHOES
	X	2422 ABBOT KINNEY BL	HARRIS, LEVELL	R08 - TATTOOS - PICTURE - TORSO, BACK
		F	?	003 - COMPLEXION - DARK
	171404150	WEAPON (CARRY/POSS)	M B BLK BRO 502 130 26	024 - EARS - LARGE
###		21310PC	1942 TRANSIENT	039 - EYES - SLANTED
	4873857			
	01/04/2017 1920	1412		
	4214	OCEAN FRONT WALK & PARK AV	ELHADARY, NAGI	? ? ? ? ? ? ? ?
	CT	F	?	[redacted]
	FB	AGGRAVATED ASSAULT	M O BLK BRO 511 165 38	BLK JKT,BLK PNTS,BLU SHOES
###		245(A)(4)PC	1942 TRANSIENT	?
	4874475			
	01/05/2017 1410	1413		
	4214	3RD ST & ROSE AV	ZORN, SANDRA	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	FB	MISC OTHER VIOLS	F H BRO BLK 504 120 60	BRO SWEATER/BLU JEANS/BLK SHS
###		853.7PC	1942 TRANSIENT	?
	Bkg #			
	Date / Time	RD Arrested		
	Arresting Division	Arrest Location	Last Name, First Name	A - Veh (Yr Mk Mod Sty Top Bot Lic St)
	Investigating Unit	Arrest Type	AKA's	Driver's License
	Arrest Detail	Arrest Group	Sex Desc Hair Eyes Ht Wt Age	Clothing
#		Arrest Charge	Arrestees Address	Personal Descriptors

	4874486			
	01/05/2017 1400	1413		
	4214	3RD ST & ROSE AV	DAMBERT, DAVID	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	FB	MISC OTHER VIOLS	M H BLN GRN 602 175 57	PURPLE SWEATSHIRT,BLU PANTS
###		853.7PC	1942 TRANSIENT	?
	4874502			
	01/05/2017 1830	1457		
	4214	12123 CULVER BL	MCKINNEY, STEPHANIE	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	A	MISC OTHER VIOLS	F W BRO BRO 506 195 47	WHTSHIRT,BLUPANTS
###		853.7PC	1942 TRANSIENT	?
	4875545			
	01/06/2017 1620	1406		
	4214	3756 OVERLAND	HOSKINS, DAVID	? ? ? ? ? ? ? ?
	CT	F	?	[redacted]
	NARC	<NA>	M W BRO GRN 600 165 37	BLK SHRT,BLKPNTS,BLK SHOES
###		3454(C)PC	1942 TRANSIENT	?
	4875562			
	01/06/2017 1615	1452		
	4214	LINCOLN & WASHINGTON	BRYANT, MAURICE	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	FB	MISC OTHER VIOLS	M B BLK BRO 509 155 40	GRY SHRT,BLU PNTS,WHT SHOES
###		853.7PC	1942 TRANSIENT	?
	4214			?
	14	1415		BLK SHRT,BLK PNTS,BLK SHOES
	A	888 LINCOLN BL	ROBINSON, SPENCER	001 - COMPLEXION - LIGHT/FAIR
		F	?	024 - EARS - LARGE
	?	ROBBERY	M W BLK BRO 509 150 21	044 - NOSE - LONG
###		211PC	1942 TRANSIENT	064 - FACE - ROUND
	4876002			
	01/07/2017 0900	1441		
	4214	MILDRED ST & STRONGS ST	WILLIAMS, JARON	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	A	DRIVING UNDER INFLU	M B BLK BRO 604 215 30	BLK JACKET,RED SHIRT,BLU JNS,BLU SHS
###		23152(B)VC	1942 TRANSIENT	?
	4876020			
	01/07/2017 0930	1411		
	4214	OCEANFRONT WALK WK & OZONE ST	MCGEE, MICCA	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	Z	MISC OTHER VIOLS	M W BRO BRO 600 145 20	BLK JACKET,WHT SHORTS
###		853.7PC	1942 TRANSIENT	?

	4876048			
	01/07/2017 0910	1400		
	4214	OCEANFRONT WALK WK & OZONE ST	RAPHAEL, PATRICK	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	Z	MISC OTHER VIOLS	M B BLK BRO 600 175 40	ALL BLACK
###		853.7PC	1942 TRANSIENT	?
	4876058			
	01/07/2017 0910	1411		
	4214	OCEANFRONT WALK WK & OZONE ST	RILEY, ROBERT	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	Z	MISC OTHER VIOLS	M B BLK BRO 600 150 46	BLK HOODIE,BLU JNS,BLK PTS
###		853.7PC	1942 TRANIENT	?
	4876075			
	01/07/2017 1315	1413		
	4214	3RD ST & ROSE AV	MEYER, THOMAS	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	A	MISC OTHER VIOLS	M W BRO BRO 510 160 59	GRY JACKET, BLUE JEANS, NO SHO
###		853.7PC	1942 TRANSIENT	?
	1943			NONE
	CT	1489		PINK SHIRT,BLU JNS,WHT SHS
	LAX	5300 W 96ST ST	BURNHAM, ASHLEY	002 - COMPLEXION - MEDIUM
		F	?	020 - EARS - PIERCED
	?	MISC OTHER VIOLS	F W BLN BLU 509 200 32	044 - NOSE - LONG
###		3454(C)PC	1942 TRANSIENT	064 - FACE - ROUND
	4876945			
	01/08/2017 1150	1415		
	4214	893 WARREN AV	BINNS, DANIEL	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	A	MISC OTHER VIOLS	M B BRO BRO 506 150 41	BLK SWEATER,WHT SHIRT,BLU JNS
###		853.7PC	204 HAMTON DR	?
	4214			NONE
	87	1412		BLU PNTS,PINK SHRT
	FB	OCEAN FRONT WK & PALOMA	DAVIS, CHRISTELYN	U06 - TATTOOS - OTHER SUBJECT - NECK
		M	?	P03 - TATTOOS - NAMES / WORDS / INIT - CHEEK
	?	NARCOTIC DRUG LAWS	F B BLK BRO 504 200 42	122 - HAIR - BUSHY
###		11377(A)HS	1942 TRANSIENT	125 - HAIR - DYED/BLEACHED
	4877135			
	01/08/2017 1800	1494		
	1943	380 WORLD WAY	HUMBERT, FREDERIC	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	LAX	OTHER ASSAULTS	M W BLK BRO 603 225 44	GRY SHRT,BLU SHORTS,BLK SHOES
###		148(A)(1)PC	1942 TRANSIENT	?

	4877232			
	01/08/2017 2055	1412		
	4214	SPEEDWAY & PARK AV	FLORES, NELSON	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	FB	MISC OTHER VIOLS	M H BLK BRO 507 160 43	TAN JACKET,BLK PNTS,BRO SHOES.
###		853.7PC	1942 TRANSIENT	?
	4214			?
	14	1441		BLU SHRT,BLUPNTS,BLK SHOES
	A	1712 ABBOT KENNY AV	LAWRENCE, JACOB	001 - COMPLEXION - LIGHT/FAIR
		F	?	039 - EYES - SLANTED
	171404350	BURGLARY	M W BLK BRO 508 170 39	044 - NOSE - LONG
###		459PC	1942 TRANSIENT	062 - FACE - LONG
	4878107			
	01/09/2017 2130	1435		
	4214	3718 S BARRINGTON AV	MATIASMORALES, VICTOR	? ? ? ? ? ? ? ?
	87	M	?	NONE
	A	NARCOTIC DRUG LAWS	M H BLK BRO 506 140 32	BLK SHOES, BLK JEANS, GRY/BLU SWEATER
###		11550(A)HS	1942 TRANSIENT	?
	1943			[redacted]
	CT	1489		GRY PNT,PRPLE/GRY SHT,CAMO JKT,GRY SHS
	LAX	96TH & HINDRY	SMITH, MELENA	001 - COMPLEXION - LIGHT/FAIR
		F	?	020 - EARS - PIERCED
	?	MISC OTHER VIOLS	F W BRO GRN 502 135 43	043 - NOSE - HOOKED
###		594(A)PC	1943 TRANSIENT	066 - FACE - HEAVY MAKEUP
	4878602			
	01/10/2017 1205	1413		
	4214	SUNSET BL & 3RD ST	GARCIA, NICHOLAS	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	FB	DRUNKENNESS	M H BRO BLU 510 160 53	GRY,SHRT,JAIL,BLU,BRO,SHS
###		647(F)PC	503 OLYMPIC BL	?
	4878703			
	01/10/2017 1555	1412		
	4214	151 OCEAN FRONT WK	GONZALES, TYAIRA	? ? ? ? ? ? ? ?
	CT	M	?	NONE
	F	NARCOTIC DRUG LAWS	F B BLK BRO 509 160 21	BLU SHRT BLU PNTS BLK SHS
###		11377(A)HS	1942 TRANSIENT	?
	4878718			
	01/10/2017 1310	1413		
	4214	3RD ST & ROSE AV	SESSIONS, RACHAEL	? ? ? ? ? ? ? ?
	14	F	?	?
	Z	BURGLARY	F W BRO BLU 501 135 34	BRO SHOES, BLU PANTS, BLU SHIRT
###		459PC	1942 TRANSIENT	?

	4878898			
	01/11/2017 0150	1441		
	4214	509 WASHINGTON BL	JOHNSON, WILLIAM	? ? ? ? ? ? ? ?
	99	F	HESLER	[redacted]
	FB	MISC OTHER VIOLS	M W BRO GRN 509 180 57	BLK SHOES, GRY PANTS, BLU JACK
###		451(D)PC	1942 TRANSIENT	?
	4879131			
	01/11/2017 0820	1453		
	4214	LINCOLN BL & WASHINGTON AV	MARZANO, NICHOLAS	? ? ? ? ? ? ? ?
	CT	O	?	[redacted]
	Z	MISC OTHER VIOLS	M W BRO BRO 507 140 40	BLU SWEATER,BLU JEANS,BLK SHS
###		3455(A)PC	1942 TRANSIENT	?
	4879663			
	01/11/2017 1340	1413		
	4214	3RD AV & ROSE AV	WHITTAKER, ALEXIS	? ? ? ? ? ? ? ?
	14	F	?	[redacted]
	FB	AGGRAVATED ASSAULT	F B BLK BRO 508 130 23	RED/BLK JCKT, BLK PANTS, BRO BOOTS
###		245(A)(1)PC	1942 TRANSIENT	?
	4879725			
	01/11/2017 1800	1426		
	4214	NATIONAL & SAWTELLE	SMITH, CARLOS	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	A	OTHER ASSAULTS	M B BLK BRO 510 160 38	BLK SHS,BLU PNTS,WHT SHIRT
###		240PC	1942 TRANSIENT	?
	4880340			
	01/12/2017 1410	1431		
	4214	MARKET ST & OCEAN FRONT WK	TALLEY, SHAWN	? ? ? ? ? ? ? ?
	14	F	PECK	[redacted]
	FB	WEAPON (CARRY/POSS)	M B BLK BRO 504 135 48	BLK SHIRT,BLK PNTS,GRY SHS
###		21310PC	1942 TRANSIENT	?
	4214			NONE
	87	1415		BLK JACKET,BRO JNS,WHT SHOES
	NARC	LINCOLN & LAKE	ROMAN, ERNEST	081 - TEETH - BROKEN
		F	?	086 - TEETH - MISSING
	171400520	NARCOTIC DRUG LAWS	M H BLK BRO 507 140 55	U19 - TATTOOS - OTHER SUBJECT - EAR, RIGHT
###		11379(A)HS	1942 TRANSIENT	U20 - TATTOOS - OTHER SUBJECT - EAR, LEFT
	4881506			? ? ? ? ? ? ? ?
	01/13/2017 1605	1415		NONE
	4214	LINCOLN & LAKE	SHERROD, JOSEPH	BLK JACKET,BLU JNS,BLK SHOES
	87	F	?	002 - COMPLEXION - MEDIUM
	NARC	NARCOTIC DRUG LAWS	M B BLK BRO 509 195 54	064 - FACE - ROUND
###		11379(A)HS	1942 TRANSIENT	T07 - TATTOOS - WORDS, OTHER - TORSO, FRONT

	4881850			
	01/14/2017 0945	1431		
	4214	PACIFIC & WINDWARD	THILKING, JOHN	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	Z	MISC OTHER VIOLS	M W BLN BLU 601 185 41	BLU SHT,BLK PNT,BLK SHS
###		853.7PC	1942 TRANSIENT	?
	4881859			
	01/14/2017 1030	1495		
	1943	SEPULVEDA & WESTCHESTER	ROGERS, RYAN	? ? ? ? ? ? ? ?
	CT	M	?	[redacted]
	LAX	AGGRAVATED ASSAULT	M W BLK BLU 509 170 42	BLU JNS,BLK SHT,GRN SNDLS
###		273.5(A)PC	1942 TRANSIENT	?
	1943			NONE
	CT	1494		GRN/RED SHT,BLU SHRTS,BLK SNDLS
	LAX	380 WORLD WAY	STRINGER, JEFFREY	001 - COMPLEXION - LIGHT/FAIR
		F	?	024 - EARS - LARGE
	?	WEAPON (CARRY/POSS)	M O BLK BRO 508 170 42	046 - NOSE - SMALL
###		22210PC	1943 TRANSIENT	122 - HAIR - BUSHY