Which police officers escalate force?

Greg Ridgeway

Rebecca W. Bushnell Professor of Criminology
Professor of Statistics and Data Science
University of Pennsylvania

March 27, 2024

Penn
UNIVERSITY of PENNSYLVANIA

Some officers seem inclined to escalate

- Laquan McDonald shooting, October 20, 2014
- CPD Officer Van Dyke fired 16 rounds
- Officer Walsh fired no rounds, holstering his firearm

Untangling officer risk from environment risk

Officers who use more force differ from other officers in obvious ways

- In the field
- In particular environments
- Conducting higher risk operations

Untangling officer risk from environment risk

Officers who use more force differ from other officers in obvious ways

- In the field
- In particular environments
- Conducting higher risk operations

“the overrepresentation of minority officers among police shooters [is] closely associated with racially varying pattern of assignment, socialization, and residence” - Fyfe (1981)

Match officers on the same scene

Find *features of officers* predictive of being a shooter (Ridgeway 2016)

- Data from 175 NYPD shooting incidents (239 shooters, 155 non-shooters)
- Conditional logistic regression
- Officers frequently accumulating negative marks in their file 3x more likely to shoot

Match officers on the same scene

Find *features of officers* predictive of being a shooter (Ridgeway 2016)

- Data from 175 NYPD shooting incidents (239 shooters, 155 non-shooters)
- Conditional logistic regression
- Officers frequently accumulating negative marks in their file 3x more likely to shoot

Find *features* predictive of rounds fired (Ridgeway, Cave, Grieco, Loeffler 2021)

- Data from 317 shootings from 46 large agencies (849 officers, 5026 rounds)
- Truncated conditional Poisson regression
- No officer feature predictive of number rounds

Match officers on the same scene

Find *features of officers* predictive of being a shooter (Ridgeway 2016)

- Data from 175 NYPD shooting incidents (239 shooters, 155 non-shooters)
- Conditional logistic regression
- Officers frequently accumulating negative marks in their file 3x more likely to shoot

Find *features* predictive of rounds fired (Ridgeway, Cave, Grieco, Loeffler 2021)

- Data from 317 shootings from 46 large agencies (849 officers, 5026 rounds)
- Truncated conditional Poisson regression
- No officer feature predictive of number rounds

Analyses did not pinpoint specific officers, results not actionable

- Identify *specific officers* rather than features of officers
- Data needs demand moving beyond shootings to other force types

Type of force depends on officer and environment

- $Y = y$ indicates type of force, $y \in \{0, 1, 2, 3\}$
- Each officer has λ, latent propensity to escalate force
- Environment $\mathbf{z}$ (e.g., time, place, lighting, suspect, policies and laws)

Type of force depends on officer and environment

- $Y = y$ indicates type of force, $y \in \{0, 1, 2, 3\}$
- Each officer has λ, latent propensity to escalate force
- Environment $\mathbf{z}$ (e.g., time, place, lighting, suspect, policies and laws)

$$P(Y_i = y|\mathbf{z}) \quad = \quad \frac{\exp(\theta_y + s_y(h(\mathbf{z}) + \lambda_i))}{\sum_{k=0}^{3} \exp(\theta_k + s_k(h(\mathbf{z}) + \lambda_i))}$$

- Like a multinomial regression model with an officer fixed effect and nuisance environmental effect

Type of force depends on officer and environment

- $Y = y$ indicates type of force, $y \in \{0, 1, 2, 3\}$
- Each officer has λ, latent propensity to escalate force
- Environment $\mathbf{z}$ (e.g., time, place, lighting, suspect, policies and laws)

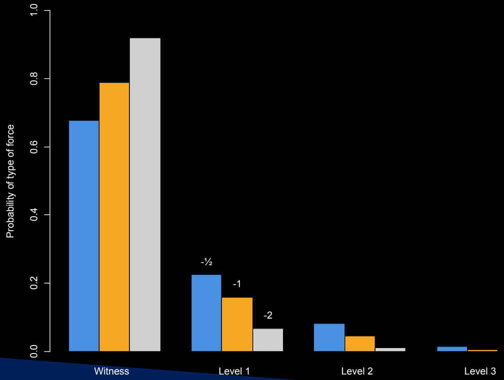
$$P(Y_i = y|\mathbf{z}) = \frac{\exp(\theta_y + s_y(h(\mathbf{z}) + \lambda_i))}{\sum_{k=0}^{3} \exp(\theta_k + s_k(h(\mathbf{z}) + \lambda_i))}$$

- Like a multinomial regression model with an officer fixed effect and nuisance environmental effect
- With an order constraint on $\mathbf{s}$, identical to Anderson (1984) ordinal stereotype model
  - Other ordinal models do not simplify when conditioning on outcomes

Penn
University of Pennsylvania

Ordinal stereotype imposes a distribution on force types

- θ set to match Seattle's force rate
- $h(\mathbf{z}) = 0$
- $\mathbf{s} = \{0, 1, \frac{3}{2}, 2\}$
- $\lambda_1 = -\frac{1}{2}, \lambda_2 = -1, \lambda_3 = -2$



Penn
UNIVERSITY of PENNSYLVANIA

Data collection and model estimation is hard

$$P(Y_i = y|\mathbf{z}) \quad = \quad \frac{\exp(\theta_y + s_y(h(\mathbf{z}) + \lambda_i))}{\sum_{k=0}^{3} \exp(\theta_k + s_k(h(\mathbf{z}) + \lambda_i))}$$

Difficult to

- randomly sample officers in time and space
- document $\mathbf{z}$
- choose $h(\mathbf{z})$

Change to a conditional likelihood question

Consider a use of force incident

- with three officers on scene
- one officer does nothing (Level 0)
- one officer physically restrains (Level 1)
- one officer strikes baton to the head (Level 3)

What's the probability $Y_1 = 0$, $Y_2 = 1$, and $Y_3 = 3$?

Change to a conditional likelihood question

Consider a use of force incident

- with three officers on scene
- one officer does nothing (Level 0)
- one officer physically restrains (Level 1)
- one officer strikes baton to the head (Level 3)

What's the probability $Y_1 = 0$, $Y_2 = 1$, and $Y_3 = 3$?

$$P(Y_1 = 0, Y_2 = 1, Y_3 = 3 | \mathbf{k} = \{1, 1, 0, 1\}, \mathbf{s}, \lambda, \theta, h(\mathbf{z}))$$
$$= \frac{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3}}{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3} + \ldots + e^{s_3\lambda_1 + s_1\lambda_2 + s_0\lambda_3}}$$

Penn UNIVERSITY of PENNSYLVANIA

Change to a conditional likelihood question

Consider a use of force incident

- with three officers on scene
- one officer does nothing (Level 0)
- one officer physically restrains (Level 1)
- one officer strikes baton to the head (Level 3)

What's the probability $Y_1 = 0$, $Y_2 = 1$, and $Y_3 = 3$?

$$P(Y_1 = 0, Y_2 = 1, Y_3 = 3 | \mathbf{k} = \{1, 1, 0, 1\}, \mathbf{s}, \lambda, \theta, h(\mathbf{z})) = \frac{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3}}{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3} + \ldots + e^{s_3\lambda_1 + s_1\lambda_2 + s_0\lambda_3}}$$

Probability does not depend on the environment. No need to measure $\mathbf{z}$ or worry about $h(\mathbf{z})$

Advantages of conditional likelihood

For a general incident with m officers

$$P(\mathbf{Y}=\mathbf{y}|\mathbf{s},\lambda,\mathbf{k}) = \frac{\exp\left(\sum_{i=1}^{m} s_{y_i}\lambda_i\right)}{\sum_{\mathbf{y}^*\in\mathcal{K}}\exp\left(\sum_{i=1}^{m} s_{y_i^*}\lambda_i\right)}$$

- No need for environmental/situational measures
- Moments when all officers have $y=0$ have no information
- Moments involving a single officer have no information
- Depends on a conditional independence $\left(Y_i, Y_j \perp\!\!\!\perp \mathbf{z}, \lambda\right)$

Advantages of conditional likelihood

For a general incident with m officers

$$P(\mathbf{Y} = \mathbf{y} | \mathbf{s}, \lambda, \mathbf{k}) = \frac{\exp\left(\sum_{i=1}^{m} s_{y_i}\lambda_i\right)}{\sum_{\mathbf{y}^* \in \mathcal{K}} \exp\left(\sum_{i=1}^{m} s_{y_i^*}\lambda_i\right)}$$

- No need for environmental/situational measures
- Moments when all officers have $y = 0$ have no information
- Moments involving a single officer have no information
- Depends on a conditional independence $\left(Y_i, Y_j \perp\!\!\!\perp \mathbf{z}, \lambda\right)$

! The only times and places with information for the conditional likelihood are those with multiple officers on the scene of a use-of-force incident

Computational challenges

For a general incident with m officers

$$P(\mathbf{Y}=\mathbf{y}|\mathbf{s},\lambda,\mathbf{k}) = \frac{\exp\left(\sum_{i=1}^{m} s_{y_i}\lambda_i\right)}{\sum_{\mathbf{y}^*\in\mathcal{K}} \exp\left(\sum_{i=1}^{m} s_{y_i^*}\lambda_i\right)}$$

Efficient computation of Poisson-Multinomial denominator

- $m=2$, easily computed
- $3 \leq m \leq 7$, no-repeat Heap (1967) recursive algorithm, $O(m!)$ time
- $m \geq 8$, discrete Fourier transform Lin et al (2023), $O(m^4)$ time
- Efficient denominator calculation plus parallelization by incident makes MCMC possible

Penn
University of Pennsylvania

λ is identifiable up to an additive constant

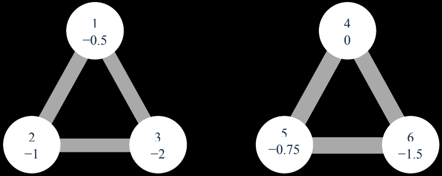
$$L_\ell(\mathbf{s}, \lambda + \mathbf{1}c) = \frac{\exp\left(\sum_{i=1}^{m_\ell} s_{y_i}(\lambda_i + c)\right)}{\sum_{\mathbf{y}^* \in \mathcal{K}_\ell} \exp\left(\sum_{i=1}^{m_\ell} s_{y_i^*}(\lambda_i + c)\right)} = L_\ell(\mathbf{s}, \lambda)$$

- Differences between two officers' λ can be identifiable

λ is identifiable up to an additive constant

$$L_\ell(\mathbf{s}, \lambda + \mathbf{1}c) = \frac{\exp\left(\sum_{i=1}^{m_\ell} s_{y_i}(\lambda_i + c)\right)}{\sum_{\mathbf{y}^* \in \mathcal{K}_\ell} \exp\left(\sum_{i=1}^{m_\ell} s_{y_i^*}(\lambda_i + c)\right)} = L_\ell(\mathbf{s}, \lambda)$$

- Differences between two officers' λ can be identifiable
- Disconnected use-of-force networks introduce additional identifiability problems

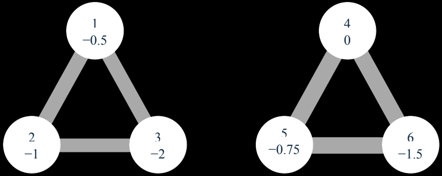


- Only officers within the same network are comparable

λ is identifiable up to an additive constant

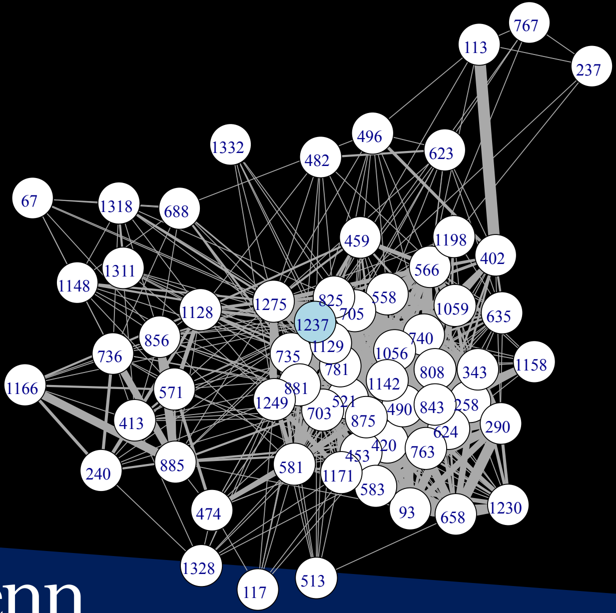
$$L_\ell(\mathbf{s}, \lambda + \mathbf{1}c) = \frac{\exp\left(\sum_{i=1}^{m_\ell} s_{y_i}(\lambda_i + c)\right)}{\sum_{\mathbf{y}^* \in \mathcal{K}_\ell} \exp\left(\sum_{i=1}^{m_\ell} s_{y_i^*}(\lambda_i + c)\right)} = L_\ell(\mathbf{s}, \lambda)$$

- Differences between two officers' λ can be identifiable
- Disconnected use-of-force networks introduce additional identifiability problems



- Only officers within the same network are comparable
- Additional complexity when comparing with officers weakly connected

Local use-of-force network for Officer 1237



Penn
UNIVERSITY of PENNSYLVANIA

Defining local network

Officer i is in Officer 1237's local network if

$$\sigma^2_{i|1237} = \mathrm{Var}(\lambda_i|\lambda_{1237}) \approx \mathrm{Var}(\lambda_i) - \frac{\mathrm{Cov}(\lambda_i, \lambda_{1237})^2}{\mathrm{Var}(\lambda_{1237})} < t$$

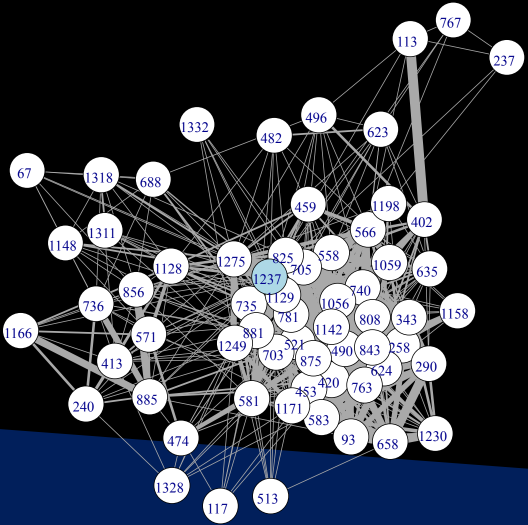
Defining local network

Officer i is in Officer 1237's local network if

$$\sigma^2_{i|1237} = \mathrm{Var}(\lambda_i|\lambda_{1237}) \approx \mathrm{Var}(\lambda_i) - \frac{\mathrm{Cov}(\lambda_i, \lambda_{1237})^2}{\mathrm{Var}(\lambda_{1237})} < t$$

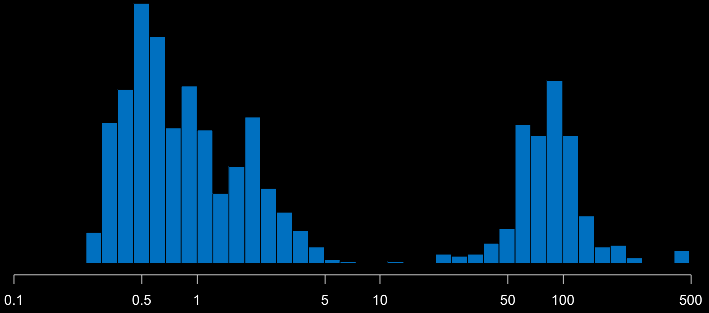
$\sigma^2_{i|1237} < 0.7$ for Officers 1275, 705, 703, 1129, and 825

$\sigma^2_{i|1237} > 70$ for Officers 67, 1311, and 117



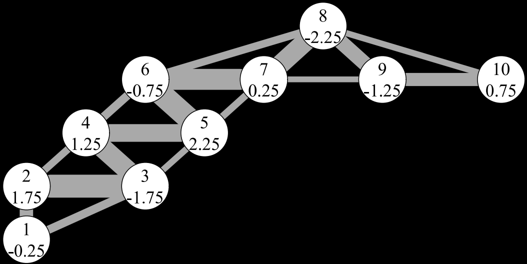
Penn
UNIVERSITY of PENNSYLVANIA

$\mathrm{Var}(\lambda_i|\lambda_{1237})$ reveals well-connected officers

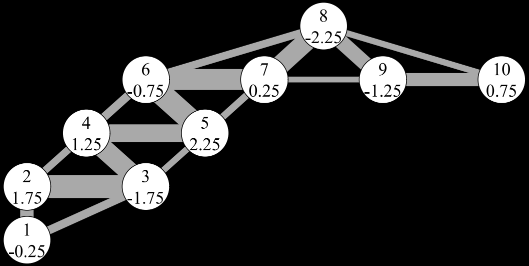


- $\lambda_{1237} - \lambda_i$ is estimable with precision only if Officer 1237's and Officer i's networks are well-connected
- Officers in disconnected subgraphs or with few shared incidents will have a large conditional variance

A ten-officer example



A ten-officer example



		posterior	
Parameter	true	mean	SD
$\lambda_1 - \lambda_2$	-2.00	-2.09	0.34
$\lambda_1 - \lambda_3$	1.50	1.85	0.31
$\lambda_1 - \lambda_4$	-1.50	-1.74	0.43
$\lambda_1 - \lambda_5$	-2.50	-3.08	0.58
$\lambda_1 - \lambda_6$	0.50	0.47	0.51
$\lambda_1 - \lambda_7$	-0.50	-0.44	0.52
$\lambda_1 - \lambda_8$	2.00	2.33	0.57
$\lambda_1 - \lambda_9$	1.00	1.04	0.57
$\lambda_1 - \lambda_{10}$	-1.00	-0.91	0.62
s_2	1.50	1.44	0.08
s_3	2.00	1.88	0.12

Analysis of full Seattle PD data

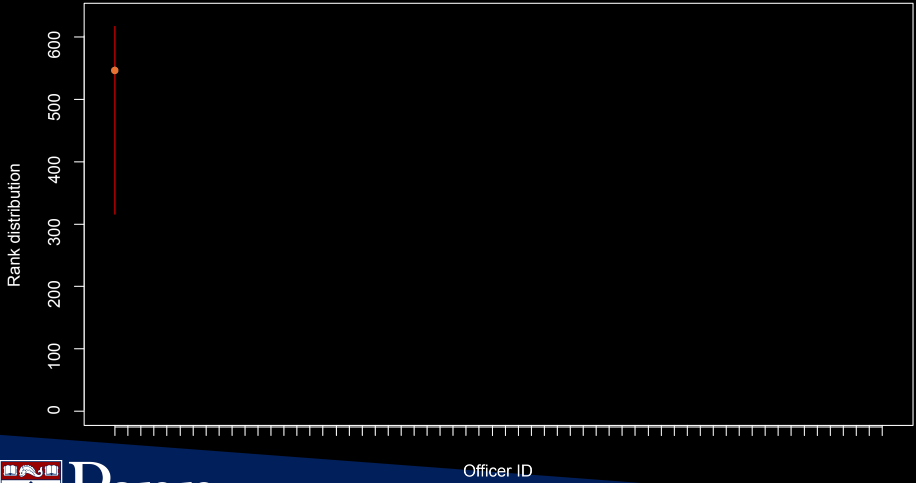
- 3,701 force incidents
- 1,386 unique officers
- 5,909 uses of force
- 15,666 officers witnessing force

Officer 1237 likely a high force escalator

- Posterior rank distribution of λs for Officer 1237's local network
- 626 officers in Officer 1237's local network

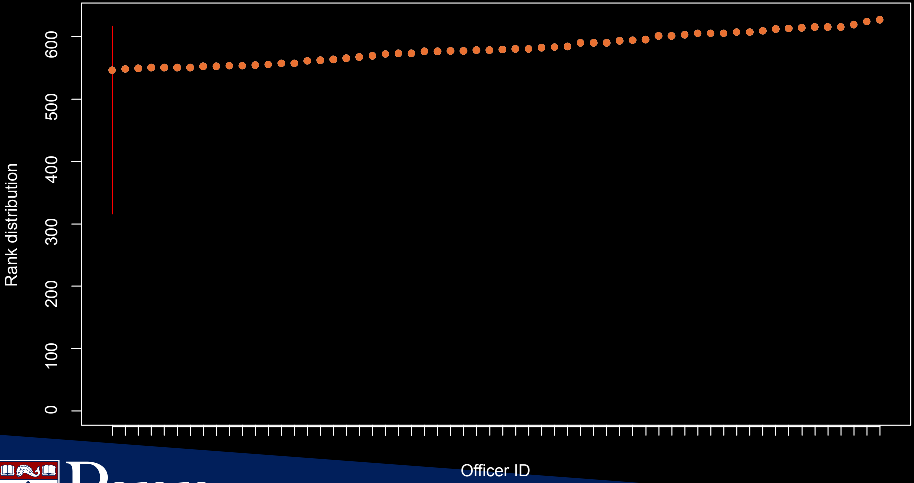
Officer 1237 likely a high force escalator

- Posterior rank distribution of λs for Officer 1237's local network
- 626 officers in Officer 1237's local network



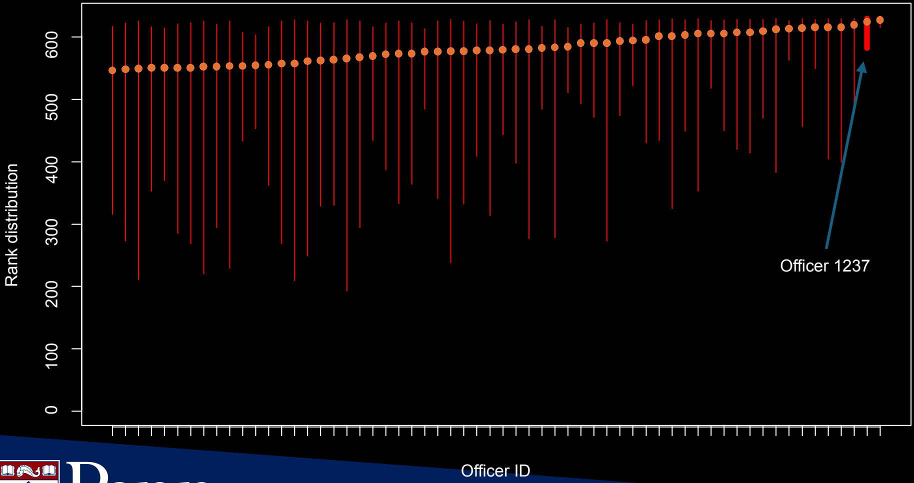
Officer 1237 likely a high force escalator

- Posterior rank distribution of λs for Officer 1237's local network
- 626 officers in Officer 1237's local network



Officer 1237 likely a high force escalator

- Posterior rank distribution of λs for Officer 1237's local network
- 626 officers in Officer 1237's local network



Flag officers with high force escalation

Officer ID	Peers	Count of force type used				Prob. rank
		Witness	Level 1	Level 2	Level 3	top 5%
#1237	626	5	11	7	0	0.94

Flag officers with high force escalation

Officer ID	Peers	Count of force type used				Prob. rank
		Witness	Level 1	Level 2	Level 3	top 5%
#0412	515	0	10	7	0	1.00
#0018	638	6	22	1	0	1.00
#0434	514	0	6	2	0	1.00
#0911	515	0	7	0	0	1.00
#0251	514	0	7	1	1	0.99
#0479	514	0	10	1	0	0.99
#0478	514	0	4	2	0	0.97
#0746	555	2	0	5	0	0.97
#1237	626	5	11	7	0	0.94

Conclusions

- Method that can be integrated into police Early Intervention Systems

Conclusions

- Method that can be integrated into police Early Intervention Systems
- Conditional likelihood solves the long-standing problem of confounding by assignment

Conclusions

- Method that can be integrated into police Early Intervention Systems
- Conditional likelihood solves the long-standing problem of confounding by assignment
- Demonstrates the value in documenting witness officers, now mandated in some consent decrees such as in Chicago

Conclusions

- Method that can be integrated into police Early Intervention Systems
- Conditional likelihood solves the long-standing problem of confounding by assignment
- Demonstrates the value in documenting witness officers, now mandated in some consent decrees such as in Chicago
- Interesting combination of policing, statistics, mathematics, and computer science
  - Poisson-Multinomial, Schur complement, discrete Fourier transform, Heap's algorithm, Markov Chain Monte Carlo, parallel computing

Which police officers escalate force?

Greg Ridgeway

Rebecca W. Bushnell Professor of Criminology
Professor of Statistics and Data Science
University of Pennsylvania

March 27, 2024