Conditional likelihood for use-of-force

Greg Ridgeway

Department of Criminology
Department of Statistics and Data Science
University of Pennsylvania

November 9, 2023

Laquan McDonald shooting, October 20, 2014

- CPD Officer Van Dyke fired 16 rounds
- Officer Walsh fired no rounds, holstering his firearm
- Van Dyke sentenced in January 2019 to 7 years in prison for murder
- Walsh found not guilty of conspiracy to obstruct, left CPD

Force depends on officer and environment

- officer with characteristics $\mathbf{x}$ (e.g., age, race, sex, experience, prior involvement in shootings)
- environment $\mathbf{z}$, shared situational, organizational, community, and legal factors (e.g., time, place, lighting, suspect, policies and laws)
- $Y = y$ indicates characteristics of force

$$g(P(Y = y|\mathbf{x}, \mathbf{z})) = \alpha'\mathbf{z} + \beta'\mathbf{x} \quad (1)$$

Force depends on officer and environment

- officer with characteristics $\mathbf{x}$ (e.g., age, race, sex, experience, prior involvement in shootings)
- environment $\mathbf{z}$, shared situational, organizational, community, and legal factors (e.g., time, place, lighting, suspect, policies and laws)
- $Y = y$ indicates characteristics of force

$$g(P(Y = y|\mathbf{x}, \mathbf{z})) = \alpha'\mathbf{z} + \beta'\mathbf{x} \tag{1}$$

Force depends on officer and environment

- officer with characteristics $\mathbf{x}$ (e.g., age, race, sex, experience, prior involvement in shootings)
- environment $\mathbf{z}$, shared situational, organizational, community, and legal factors (e.g., time, place, lighting, suspect, policies and laws)
- $Y = y$ indicates characteristics of force

$$g(P(Y = y|\mathbf{x}, \mathbf{z})) = \alpha'\mathbf{z} + \beta'\mathbf{x} \tag{1}$$

A range of force models

- If $Y \in 0, 1$, shooter/non-shooter

$$\log \frac{P(Y = 1|\mathbf{x}, \mathbf{z})}{P(Y = 0|\mathbf{x}, \mathbf{z})} = \alpha'\mathbf{z} + \beta'\mathbf{x} \tag{2}$$

- If $Y \in 0, 1, 2, \ldots$, rounds fired

$$\log(P(Y = y|\mathbf{x}, \mathbf{z})) = y \log\left(\alpha'\mathbf{z} + \beta'\mathbf{x}\right) - e^{\alpha'\mathbf{z}+\beta'\mathbf{x}} - \log(y!) \tag{3}$$

- If $Y \in 0, 1, 2, \ldots, 5$, use-of-force options

$$P(Y = y_j|\mathbf{x}, \mathbf{z}) = \frac{\exp\left(\alpha_j'\mathbf{z} + \beta_j'\mathbf{x}\right)}{\sum_m \exp(\alpha_m'\mathbf{z} + \beta_m'\mathbf{x})} \tag{4}$$

Condition on a sufficient statistic to remove $\mathbf{z}$

- Let $S(\mathbf{y}_i)$ be a sufficient statistic for α (number of shooters in incident, number of rounds fired)

$$
\begin{aligned}
L_i(\alpha,\beta) &= P(Y_1=y_1,\ldots,Y_{n_i}=y_{n_i}|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\
&= P(Y_1=y_1,\ldots,Y_{n_i}=y_{n_i}|S_i,\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\not{\mathbf{z}},\not{\alpha},\beta) \\
&\qquad P(S_i|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\
&= \text{individual officer likelihood} \times \\
&\qquad \text{collective group likelihood}
\end{aligned}
$$

- Manski & Lerman (1977) and Prentice & Pyke (1979) showed that $\hat{\beta}$ using any or all of these terms are consistent for β

! The only moments and places with information on β are incidents with multiple officers varying in their actions

Condition on a sufficient statistic to remove $\mathbf{z}$

- Let $S(\mathbf{y}_i)$ be a sufficient statistic for α (number of shooters in incident, number of rounds fired)

$$
\begin{aligned}
L_i(\alpha,\beta) &= P(Y_1 = y_1,\ldots,Y_{n_i} = y_{n_i}|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\
&= P(Y_1 = y_1,\ldots,Y_{n_i} = y_{n_i}|S_i,\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\cancel{\mathbf{z}},\cancel{\alpha},\beta) \\
&\qquad P(S_i|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\
&= \text{individual officer likelihood} \times \\
&\qquad \text{collective group likelihood}
\end{aligned}
$$

- Manski & Lerman (1977) and Prentice & Pyke (1979) showed that $\hat{\beta}$ using any or all of these terms are consistent for β

> **!** The only moments and places with information on β are incidents with multiple officers varying in their actions

Condition on a sufficient statistic to remove $\mathbf{z}$

- Let $S(\mathbf{y}_i)$ be a sufficient statistic for α (number of shooters in incident, number of rounds fired)

$$
\begin{aligned}
L_i(\alpha,\beta) &= P(Y_1=y_1,\ldots,Y_{n_i}=y_{n_i}|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\
&= P(Y_1=y_1,\ldots,Y_{n_i}=y_{n_i}|S_i,\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\cancel{\mathbf{z}},\cancel{\alpha},\beta) \\
&\qquad P(S_i|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\
&= \text{individual officer likelihood} \times \\
&\qquad \text{collective group likelihood}
\end{aligned}
$$

- Manski & Lerman (1977) and Prentice & Pyke (1979) showed that $\hat{\beta}$ using any or all of these terms are consistent for β

! The only moments and places with information on β are incidents with multiple officers varying in their actions

Condition on a sufficient statistic to remove $\mathbf{z}$

- Let $S(\mathbf{y}_i)$ be a sufficient statistic for α (number of shooters in incident, number of rounds fired)

$$\begin{aligned} L_i(\alpha,\beta) &= P(Y_1=y_1,\ldots,Y_{n_i}=y_{n_i}|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\ &= P(Y_1=y_1,\ldots,Y_{n_i}=y_{n_i}|S_i,\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\cancel{\mathbf{z}},\cancel{\alpha},\beta) \\ &\qquad P(S_i|\mathbf{x}_1,\ldots,\mathbf{x}_{n_i},\mathbf{z},\alpha,\beta) \\ &= \text{individual officer likelihood} \times \\ &\qquad \text{collective group likelihood} \end{aligned}$$

- Manski & Lerman (1977) and Prentice & Pyke (1979) showed that $\hat{\beta}$ using any or all of these terms are consistent for β

The only moments and places with information on β are incidents with multiple officers varying in their actions

Condition on a sufficient statistic to remove $\mathbf{z}$

- Let $S(\mathbf{y}_i)$ be a sufficient statistic for α (number of shooters in incident, number of rounds fired)

$$
\begin{aligned}
L_i(\alpha,\beta) &= P(Y_1 = y_1, \ldots, Y_{n_i} = y_{n_i} | \mathbf{x}_1, \ldots, \mathbf{x}_{n_i}, \mathbf{z}, \alpha, \beta) \\
&= P(Y_1 = y_1, \ldots, Y_{n_i} = y_{n_i} | S_i, \mathbf{x}_1, \ldots, \mathbf{x}_{n_i}, \cancel{\mathbf{z}}, \cancel{\alpha}, \beta) \\
&\qquad\qquad P(S_i | \mathbf{x}_1, \ldots, \mathbf{x}_{n_i}, \mathbf{z}, \alpha, \beta) \\
&= \text{individual officer likelihood} \times \\
&\qquad\qquad \text{collective group likelihood}
\end{aligned}
$$

- Manski & Lerman (1977) and Prentice & Pyke (1979) showed that $\hat{\beta}$ using any or all of these terms are consistent for β

The only moments and places with information on β are incidents with multiple officers varying in their actions

Conditional likelihood offers consistent $\hat{\beta}$

- Match shooters and non-shooters on the same scene and use conditional logistic regression to estimate β

  G. Ridgeway (2016). "Officer Risk Factors Associated with Police Shootings: A Matched Case-Control Study," *Statistics and Public Policy* 3(1):1-6

- Count rounds fired by each officer on the scene and use conditional Poisson regression to estimate β

  G. Ridgeway, B. Cave, J. Grieco, and C.E. Loeffler (2021). "A Conditional Likelihood Model of the Relationship Between Officer Features and Rounds Discharged in Police Shootings," *Journal of Quantitative Criminology* 37(1):303-326.

NYPD analysis, 239 shooters, 155 non-shooters, 175 incidents

Variable	Odds-ratio	95% interval	Permutation p-value
Rank			
Police officer (reference)			
Detective	1.2	(0.2,6.3)	0.78
Sergeant	*0.2	(0.1,0.7)	0.006
Lieutenant	*0.0	(0.0,0.4)	0.003
Captain	0.1	(0.0,0.8)	0.16
Years at NYPD	1.0	(0.9,1.1)	0.89
Age when recruited	*0.9	(0.8,1.0)	0.03
Race			
White (reference)			
Black	*3.3	(1.2,8.9)	0.01
Other	1.2	(0.5,2.8)	0.71
Male	2.1	(0.5,8.9)	0.29
Education			
High school (reference)			
High school+some college	1.3	(0.5,3.0)	0.60
College	1.9	(0.6,6.1)	0.26
College+some graduate	1.8	(0.1,22.7)	0.68
CPI points $>$ 3.1	*3.1	(1.0,8.9)	0.03
Misdemeanor arrests $>$ 10.0	*0.2	(0.1,0.6)	0.002

NYPD analysis, 239 shooters, 155 non-shooters, 175 incidents

Variable	Odds-ratio	95% interval	Permutation p-value
Rank			
Police officer (reference)			
Detective	1.2	(0.2,6.3)	0.78
Sergeant	*0.2	(0.1,0.7)	0.006
Lieutenant	*0.0	(0.0,0.4)	0.003
Captain	0.1	(0.0,0.8)	0.16
Years at NYPD	1.0	(0.9,1.1)	0.89
Age when recruited	*0.9	(0.8,1.0)	0.03
Race			
White (reference)			
Black	*3.3	(1.2,8.9)	0.01
Other	1.2	(0.5,2.8)	0.71
Male	2.1	(0.5,8.9)	0.29
Education			
High school (reference)			
High school+some college	1.3	(0.5,3.0)	0.60
College	1.9	(0.6,6.1)	0.26
College+some graduate	1.8	(0.1,22.7)	0.68
CPI points > 3.1	*3.1	(1.0,8.9)	0.03
Misdemeanor arrests > 10.0	*0.2	(0.1,0.6)	0.002

46 agencies, 317 shootings, 849 officers, 5026 rounds

Variable	Rate ratio	95% interval	Permutation p-value
Age at recruitment	1.01	(0.99,1.02)	0.31
Years of experience	1.00	(0.98,1.01)	0.58
Female	0.86	(0.64,1.14)	0.27
Race (relative to white)			
Black	1.05	(0.86,1.28)	0.64
Hispanic	1.09	(0.89,1.32)	0.39
Other	0.76	(0.57,1.01)	0.05
Prior OIS (relative to 0)			
1 or more	1.02	(0.84,1.24)	0.85
2 or more	1.23	(0.88,1.72)	0.21
Prior force complaint	1.25	(0.92,1.69)	0.14
Role			
Detective	1.09	(0.78,1.52)	0.61
Sergeant or more senior	1.03	(0.87,1.22)	0.75
Other	0.66	(0.32,1.37)	0.26
Special assignment	1.28	(0.97,1.68)	0.07
Long gun (relative to pistol)	1.01	(0.78,1.30)	0.97

46 agencies, 317 shootings, 849 officers, 5026 rounds

Variable	Rate ratio	95% interval	Permutation p-value
Age at recruitment	1.01	(0.99,1.02)	0.31
Years of experience	1.00	(0.98,1.01)	0.58
Female	0.86	(0.64,1.14)	0.27
Race (relative to white)			
Black	1.05	(0.86,1.28)	0.64
Hispanic	1.09	(0.89,1.32)	0.39
Other	0.76	(0.57,1.01)	0.05
Prior OIS (relative to 0)			
1 or more	1.02	(0.84,1.24)	0.85
2 or more	1.23	(0.88,1.72)	0.21
Prior force complaint	1.25	(0.92,1.69)	0.14
Role			
Detective	1.09	(0.78,1.52)	0.61
Sergeant or more senior	1.03	(0.87,1.22)	0.75
Other	0.66	(0.32,1.37)	0.26
Special assignment	1.28	(0.97,1.68)	0.07
Long gun (relative to pistol)	1.01	(0.78,1.30)	0.97

Ordered stereotype model for force escalation

- Expand beyond shootings to other modes of force
- Identify individual officer effects, not officer features
- Let $Y = 0, \ldots, K$ order severity of use-of-force from $Y = 0$ representing no force to $Y = K$ representing lethal force

$$P(Y = y|\mathbf{z}, \mathbf{x}) = \frac{\exp(\theta_y + s_y(\alpha'\mathbf{z} + \beta'\mathbf{x}))}{\sum_{k=1}^{K} \exp(\theta_k + s_k(\alpha'\mathbf{z} + \beta'\mathbf{x}))}$$

- $\theta_0 = 0$, $s_0 = 0$, $s_1 = 1$ for identifiability
- s_k effectively quantify the "distance" between force levels

Ordered stereotype model for force escalation

- Expand beyond shootings to other modes of force
- Identify individual officer effects, not officer features
- Let $Y = 0, \ldots, K$ order severity of use-of-force from $Y = 0$ representing no force to $Y = K$ representing lethal force

$$P(Y = y|\mathbf{z}, \mathbf{x}) = \frac{\exp(\theta_y + s_y(\alpha'\mathbf{z} + \beta'\mathbf{x}))}{\sum_{k=1}^{K} \exp(\theta_k + s_k(\alpha'\mathbf{z} + \beta'\mathbf{x}))}$$

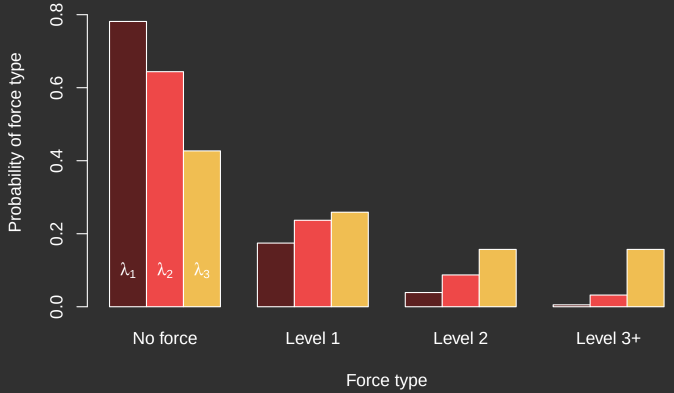
- $\theta_0 = 0$, $s_0 = 0$, $s_1 = 1$ for identifiability
- s_k effectively quantify the "distance" between force levels

Which specific officers most likely to escalate?

$$P(Y_j = y|\mathbf{z}) \quad = \quad \frac{\exp(\theta_y + s_y(\alpha'\mathbf{z} + \lambda_j))}{\sum_{k=1}^{K} \exp(\theta_k + s_k(\alpha'\mathbf{z} + \lambda_j))}$$

Ordered stereotype can model force type selection

- $\theta = \{0, -1, -2, -3\}$
- $s = \{0, 1, 2, 4\}$
- $\lambda_1 = -\frac{1}{2},\ \lambda_2 = 0,\ \lambda_3 = \frac{1}{2}$



Conditional likelihood for ordered stereotype model

- First officer does nothing, $Y_1 = 0$
- Second officer holds and pushes, $Y_2 = 1$
- Third officer strikes with a baton, $Y_3 = 3$

Conditional likelihood terms look like

$$P(Y_1 = 0, Y_2 = 1, Y_3 = 3) = \frac{e^{s_0\lambda_1+s_1\lambda_2+s_3\lambda_3}}{e^{s_0\lambda_1+s_1\lambda_2+s_3\lambda_3} + \ldots + e^{s_3\lambda_1+s_1\lambda_2+s_0\lambda_3}}$$

Unconstrained conditional multinomial logistic regression model has this same structure

Conditional likelihood for ordered stereotype model

- First officer does nothing, $Y_1 = 0$
- Second officer holds and pushes, $Y_2 = 1$
- Third officer strikes with a baton, $Y_3 = 3$

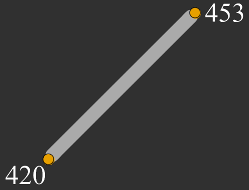
Conditional likelihood terms look like

$$\begin{aligned} & P(Y_1 = 0, Y_2 = 1, Y_3 = 3) \\ & \quad = \frac{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3}}{e^{s_0\lambda_1 + s_1\lambda_2 + s_3\lambda_3} + \ldots + e^{s_3\lambda_1 + s_1\lambda_2 + s_0\lambda_3}} \end{aligned}$$

Unconstrained conditional multinomial logistic regression model has this same structure

Seattle PD data

- 1,386 unique officers, 635 have fewer than 10 force incidents
- 3,701 force incidents with information
  - More than one officer
  - Variation in force type used
- An example subnetwork

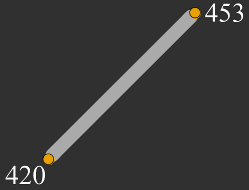


Incident	420	453
620	1	2
870	1	2
941	2	1
942	1	2
1316	2	1
1542	1	2
1747	2	1
1813	2	1
2036	2	1
2112	2	1
2303	2	1
2639	1	3
2723	2	1
2945	1	2

Officer ID	Escalation λ
453	-0.22
420	0.22

Seattle PD data

- 1,386 unique officers, 635 have fewer than 10 force incidents
- 3,701 force incidents with information
  - More than one officer
  - Variation in force type used
- An example subnetwork

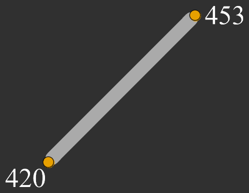


Incident	420	453
620	1	2
870	1	2
941	2	1
942	1	2
1316	2	1
1542	1	2
1747	2	1
1813	2	1
2036	2	1
2112	2	1
2303	2	1
2639	1	3
2723	2	1
2945	1	2

Officer ID	Escalation λ
453	-0.22
420	0.22

Seattle PD data

- 1,386 unique officers, 635 have fewer than 10 force incidents
- 3,701 force incidents with information
  - More than one officer
  - Variation in force type used
- An example subnetwork

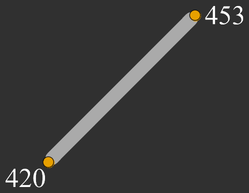


Incident	420	453
620	1	2
870	1	2
941	2	1
942	1	2
1316	2	1
1542	1	2
1747	2	1
1813	2	1
2036	2	1
2112	2	1
2303	2	1
2639	1	3
2723	2	1
2945	1	2

Officer ID	Escalation λ
453	-0.22
420	0.22

Seattle PD data

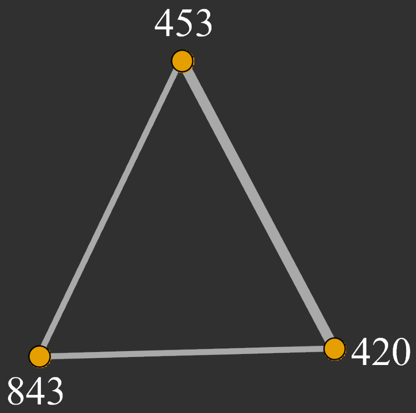
- 1,386 unique officers, 635 have fewer than 10 force incidents
- 3,701 force incidents with information
  - More than one officer
  - Variation in force type used
- An example subnetwork



Incident	420	453
620	1	2
870	1	2
941	2	1
942	1	2
1316	2	1
1542	1	2
1747	2	1
1813	2	1
2036	2	1
2112	2	1
2303	2	1
2639	1	3
2723	2	1
2945	1	2

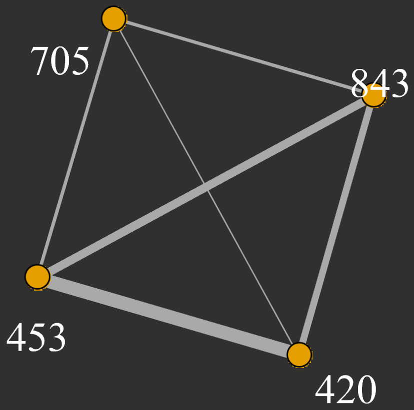
Officer ID	Escalation λ
453	-0.22
420	0.22

Example with three officer network



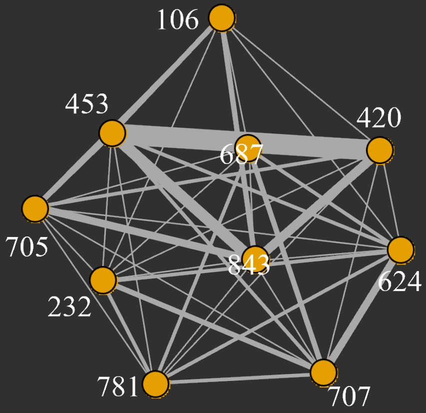
Officer ID	Escalation λ
453	-0.34
420	0.12
843	0.21
17 incidents	

Example with four officer network



Officer ID	Escalation λ
453	-0.49
420	0.04
705	0.22
843	0.22
54 incidents	

Example with ten officer network



Officer ID	Escalation λ
232	-1.78
453	-0.32
420	-0.26
106	-0.19
781	-0.16
705	-0.12
687	0.30
624	0.50
843	0.54
707	1.32
149 incidents	

Identification becomes difficult for weakly connected officers, effecting everyone's λs

Conclusion

- Conditional likelihood solves the long-standing problem of confounding by assignment
- The approach has potential beyond shooting decisions to force severity
- 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
- Opportunities
  - apply to new departments
  - assess other police performance measures
  - check robustness to contagion/anti-contagion

Conclusion

- Conditional likelihood solves the long-standing problem of confounding by assignment
- The approach has potential beyond shooting decisions to force severity
- 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
- Opportunities
  - apply to new departments
  - assess other police performance measures
  - check robustness to contagion/anti-contagion

Conclusion

- Conditional likelihood solves the long-standing problem of confounding by assignment
- The approach has potential beyond shooting decisions to force severity
- 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
- Opportunities
  - apply to new departments
  - assess other police performance measures
  - check robustness to contagion/anti-contagion

Conclusion

- Conditional likelihood solves the long-standing problem of confounding by assignment
- The approach has potential beyond shooting decisions to force severity
- 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
- Opportunities
  - apply to new departments
  - assess other police performance measures
  - check robustness to contagion/anti-contagion

Conclusion

- Conditional likelihood solves the long-standing problem of confounding by assignment
- The approach has potential beyond shooting decisions to force severity
- 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
- Opportunities
  - apply to new departments
  - assess other police performance measures
  - check robustness to contagion/anti-contagion

Conditional likelihood for use-of-force

Greg Ridgeway

Department of Criminology
Department of Statistics and Data Science
University of Pennsylvania

November 9, 2023