“Modern technology can make many specific contributions to criminal administration. The most significant will come from the use of computers to collect and analyze the masses of data the system needs to understand the crime control process”

- 1967 President’s Commission on Law Enforcement and Administration of Justice

Criminological Data Science

Greg Ridgeway
Department of Criminology
Department of Statistics

In 1967 data and computing too inaccessible to achieve the vision

- Programming languages with limited support
  - Michigan Algorithm Decoder
- Lack of technical expertise or rough user interfaces
  - LAPD could not adapt software to its data
- Challenges in accessing computing resources
  - NYPD connected to an MIT computer through a telephone line
- Cost of computing
  - IBM System 370 machines with 1MB of memory and 800MB of storage cost $25 million in 2017 dollars.

Commission foretold many computing innovations for justice

- "portable recording devices" to facilitate data collection
- computers that could automate the dispatch of patrol cars closest to calls for service
- networked alarms that could notify nearby officers without a dispatcher
- alteration of police deployments in real-time as data reveal emerging problems
- new wireless networks to reduce communication congestion
- …even electronic cocktail olives

Technological advances should come with increased public safety…

	1967
Violent crime clearance rate	45%
Property crime clearance rate	20%
Murder rate per 100,000	6.2

…yet basic performance measures of crime are unchanged over 50 years

	1967	2017
Violent crime clearance rate	45%	46%
Property crime clearance rate	20%	18%
Murder rate per 100,000	6.2	5.3

Canada shows the same trends

	USA		Canada	
	1967	2016	1967	2016
Murder clearance rate	91%	59%	92%	75%
Murder rate per 100,000	6.2	5.3	1.8	1.7

Demonstrate data, statistical models, and computation with four examples

Criminological Data Science

- Effect of gang injunctions on crime
- Benchmarking performance
- Officer risk factors for shooting
- Grams of marijuana in a joint

Data science to evaluate crime prevention initiatives

Criminological Data Science

- Effect of gang injunctions on crime
- Benchmarking performance
- Officer risk factors for shooting
- Grams of marijuana in a joint

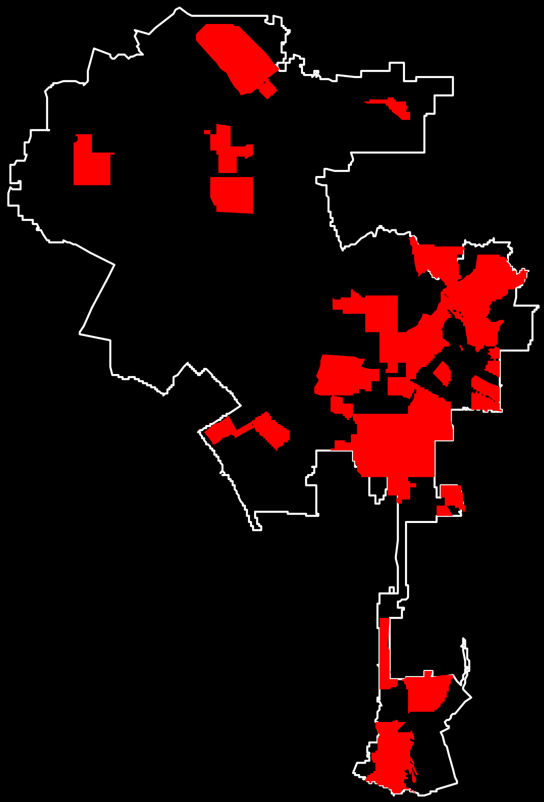
Civil Gang Injunctions (CGIs) are Neighborhood-Focused Interventions

- Designed to interfere with routine behaviors of gang members within defined “safety zones”
- Civilly enjoin otherwise-legal activities
  - publicly congregating with other gang members
  - consuming alcohol
  - being outside after curfew
  - possessing spray paint

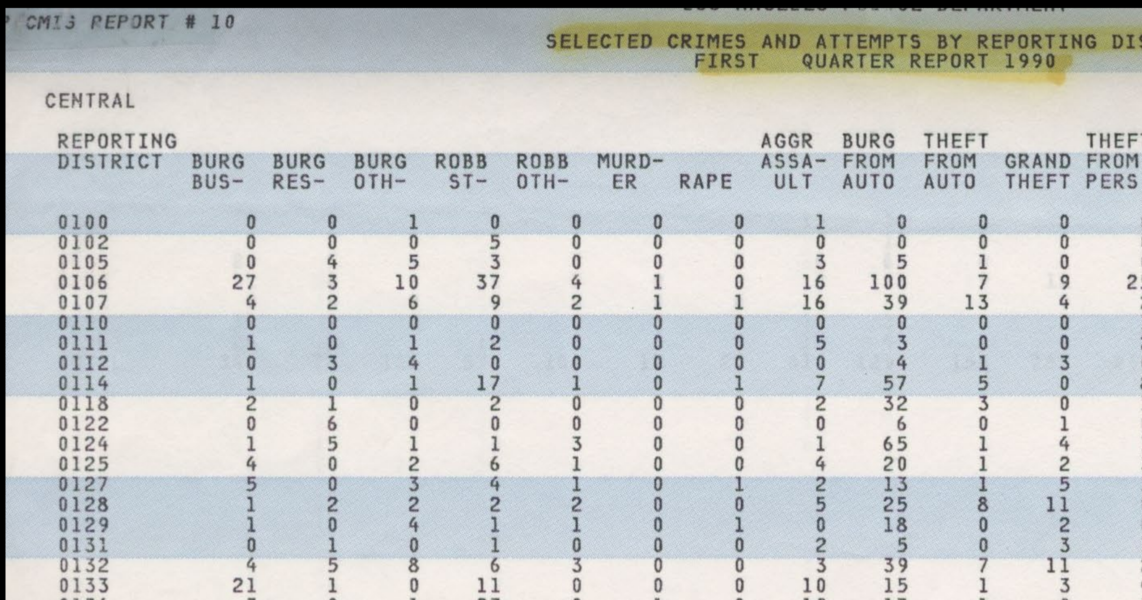
G. Ridgeway, J. Grogger, R.A. Moyer, and J.M. MacDonald. “Effect of Gang Injunctions on Crime: A Study of Los Angeles from 1988-2014.”

Los Angeles Provides Useful Framework to Assess CGI Effects

- LA historically has experienced severe gang-related crime
- 48 CGIs in effect in LA through 2014; 3 earlier CGIs terminated
- Any effects begin when complaint is served
- Our analysis uses quarterly LAPD crime reports (1988-2014)
- 939 RDs over 108 quarters



Data From 1988-2004 Collected from 2,300 Pages at LA Library

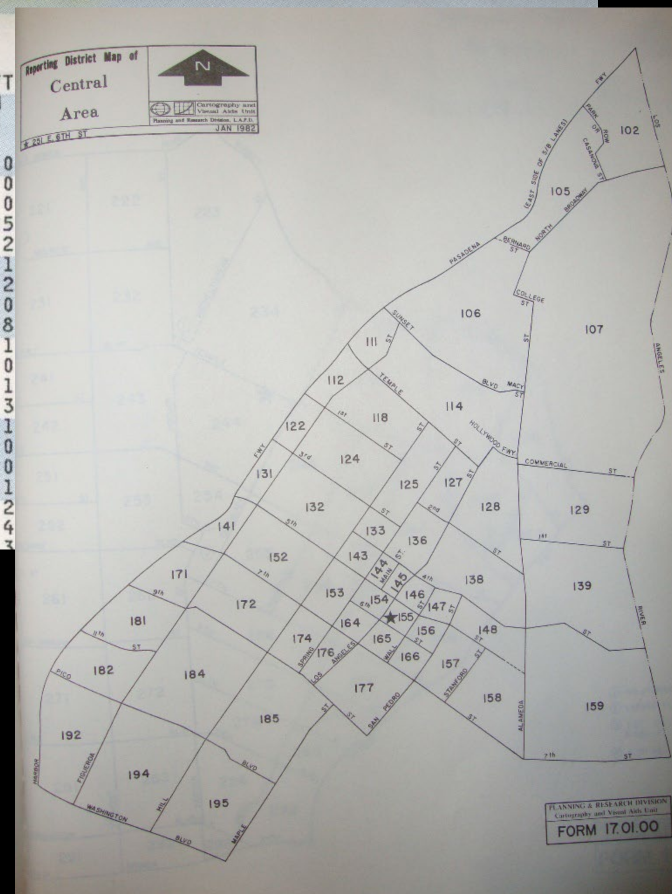


CMIS REPORT # 10

SELECTED CRIMES AND ATTEMPTS BY REPORTING DISTRICT
FIRST QUARTER REPORT 1990

CENTRAL

REPORTING DISTRICT	BURG BUS-	BURG RES-	BURG OTH-	ROBB ST-	ROBB OTH-	MURD- ER	RAPE	AGGR ASSA- ULT	BURG FROM AUTO	THEFT FROM AUTO	GRAND THEFT	THEFT FROM PERS
0100	0	0	1	0	0	0	0	1	0	0	0	0
0102	0	0	0	5	0	0	0	0	0	0	0	0
0105	0	4	5	3	0	0	0	3	5	1	0	0
0106	27	3	10	37	4	1	0	16	100	7	9	25
0107	4	2	6	9	2	1	1	16	39	13	4	2
0110	0	0	0	0	0	0	0	0	0	0	0	1
0111	1	0	1	2	0	0	0	5	3	0	0	2
0112	0	3	4	0	0	0	0	0	4	1	3	0
0114	1	0	1	17	1	0	1	7	57	5	0	8
0118	2	1	0	2	0	0	0	2	32	3	0	1
0122	0	6	0	0	0	0	0	0	6	0	1	0
0124	1	5	1	1	3	0	0	1	65	1	4	1
0125	4	0	2	6	1	0	0	4	20	1	2	3
0127	5	0	3	4	1	0	1	2	13	1	5	1
0128	1	2	2	2	2	0	0	5	25	8	11	0
0129	1	0	4	1	1	0	1	0	18	0	2	0
0131	0	1	0	1	0	0	0	2	5	0	3	1
0132	4	5	8	6	3	0	0	3	39	7	11	2
0133	21	1	0	11	0	0	0	10	15	1	3	4
0136	3	0	1	27	0	1	0	19	17	1	0	3



- Data from 2005-2014 came from LAPD incident level crime data
- All data available at github.com/gregridgeway/LAPDcrimedata

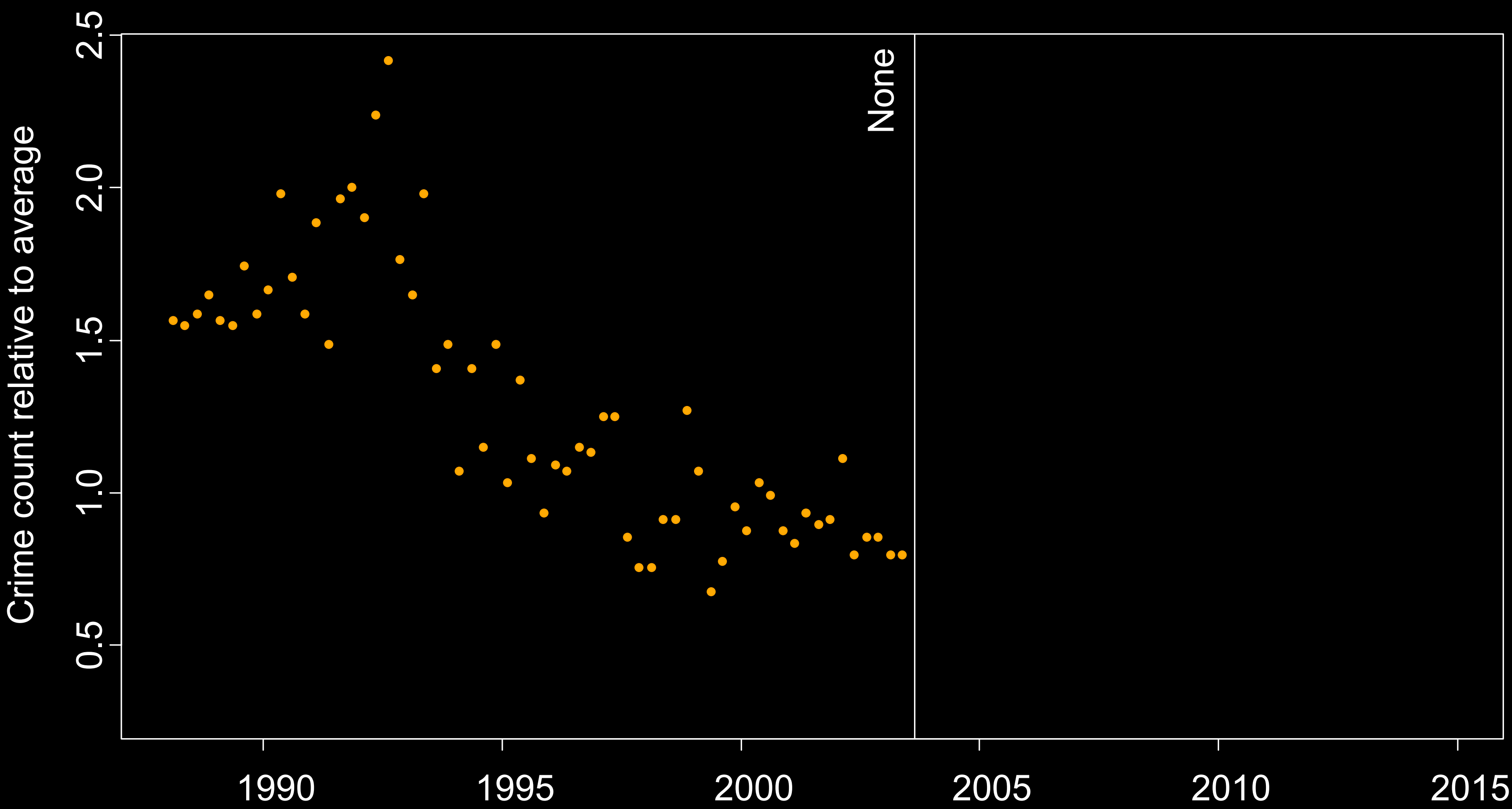
Isolate Effect of CGI Using Three Approaches

- *Stepped wedge design* – compare RDs in, adjacent to, and near CGI safety zones over time
- *Short term, CGI RDs only* – analyze only RDs with CGIs in the year before and after a CGI begins
- *Rampart scandal disruption* – use the disruption to assess crime changes before the CGI, during the CGI, when the CGI was suspended, and restarted

Isolate Effect of CGI Using Three Approaches

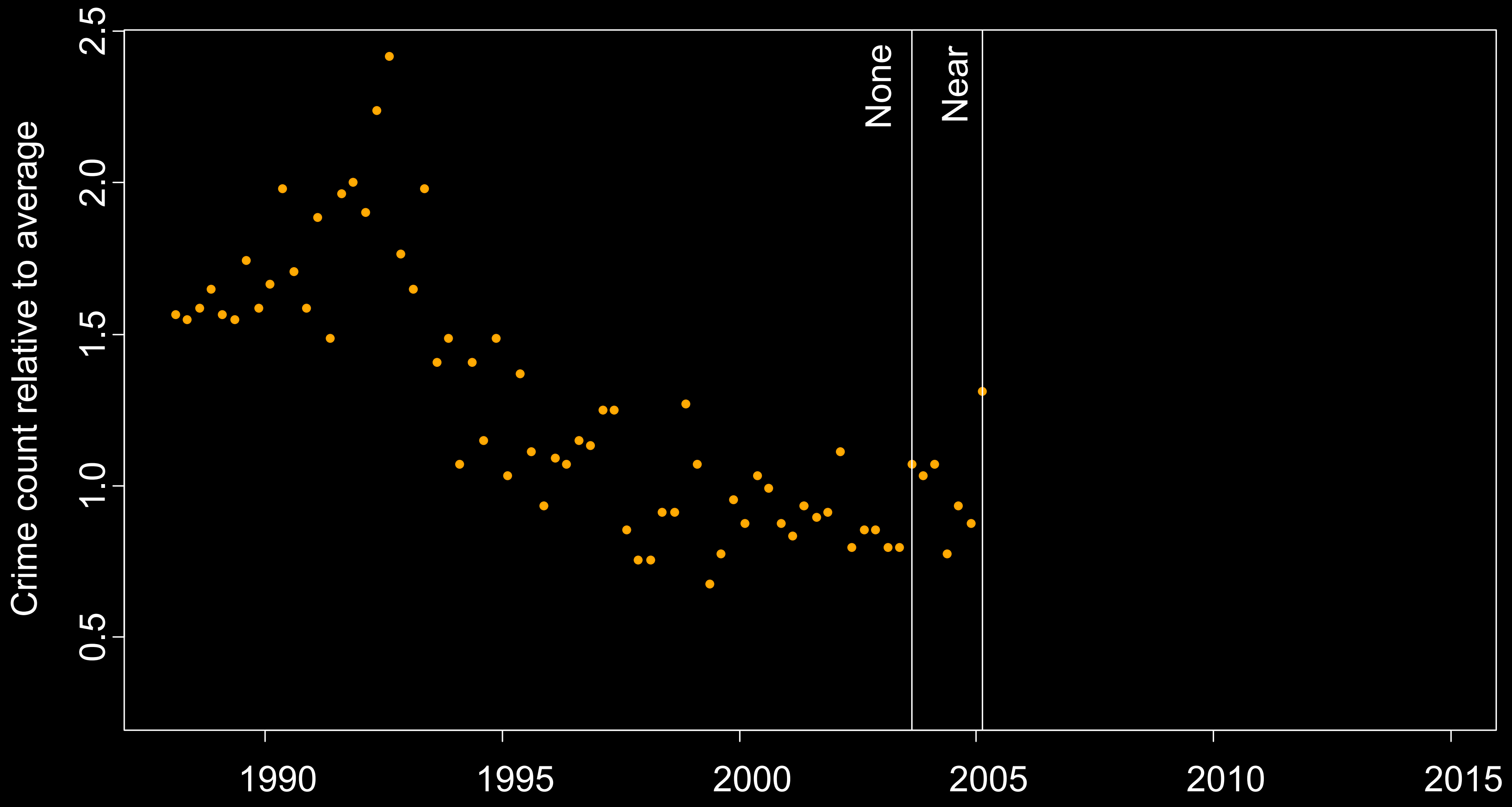
- *Stepped wedge design* – compare RDs in, adjacent to, and near CGI safety zones over time
- *Short term, CGI RDs only* – analyze only RDs with CGIs in the year before and after a CGI begins
- *Rampart scandal disruption* – use the disruption to assess crime changes before the CGI, during the CGI, when the CGI was suspended, and restarted

RD1204 Transitions from No Safety Zone, …



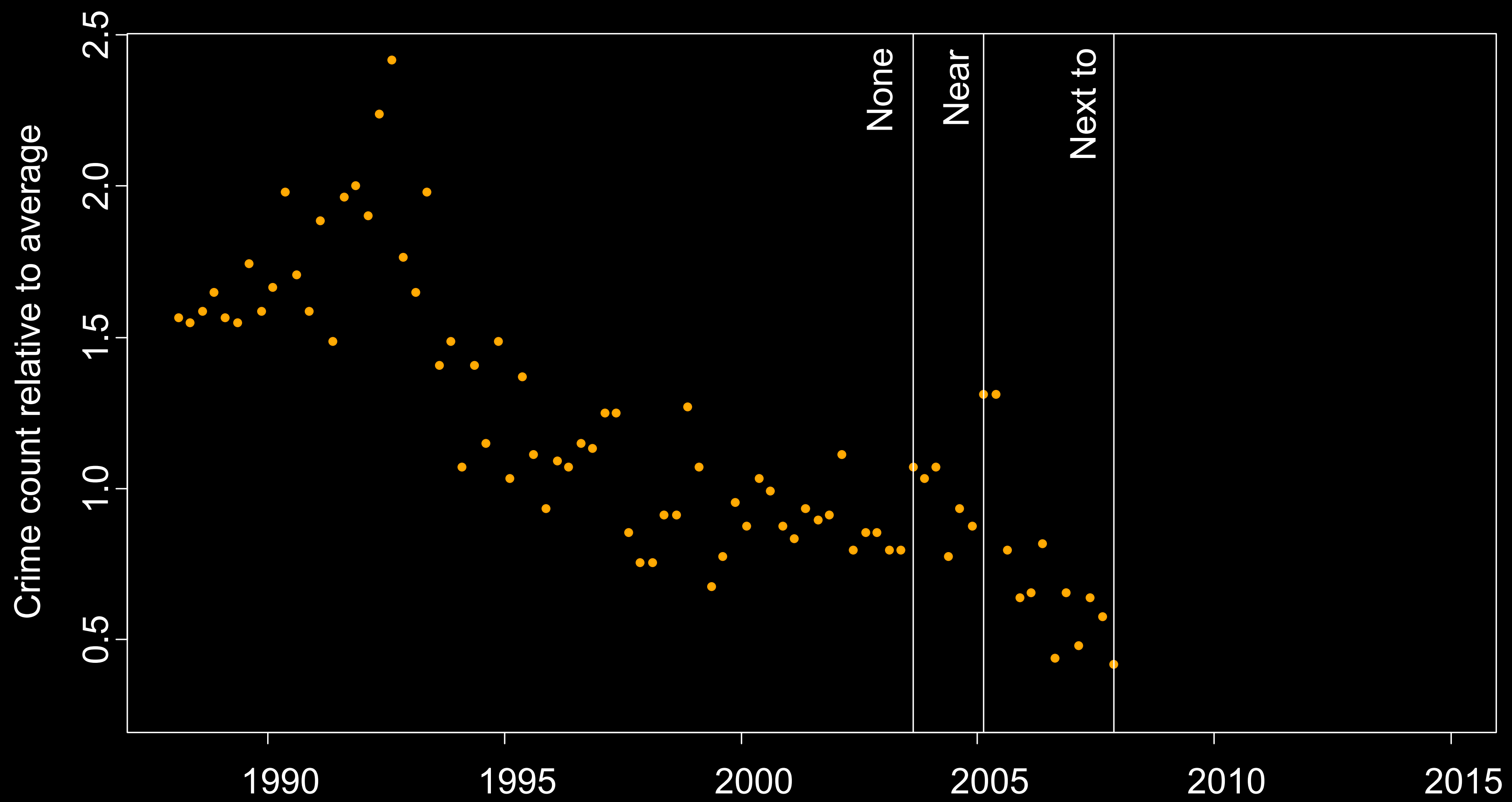
Fields Sep 2018

RD1204 Transitions from No Safety Zone, to Near, …



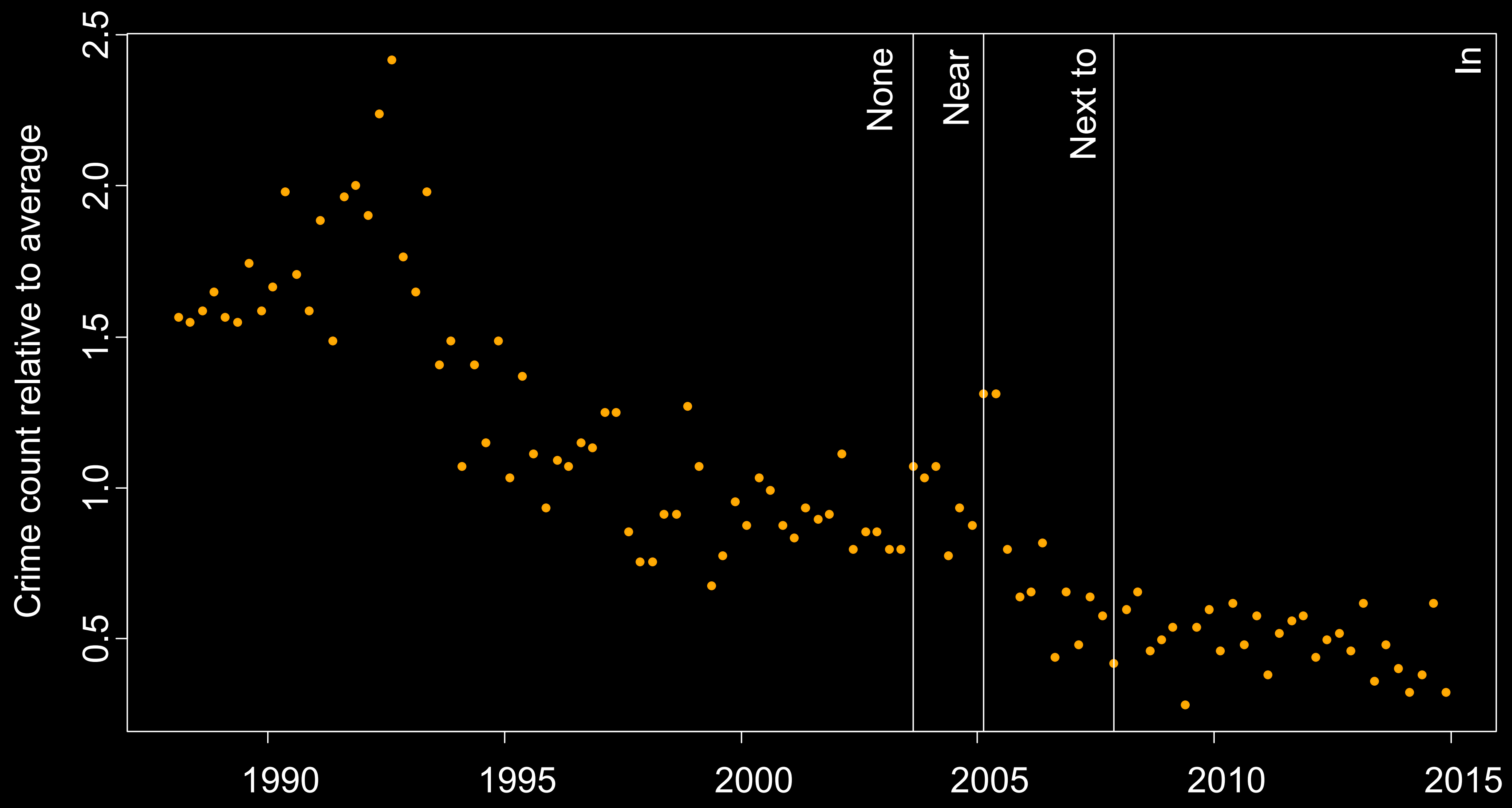
Fields Sep 2018

RD1204 Transitions from No Safety Zone, to Near, Next to, ...



Fields Sep 2018

RD1204 Transitions from No Safety Zone, to Near, Next to, in a Safety Zone



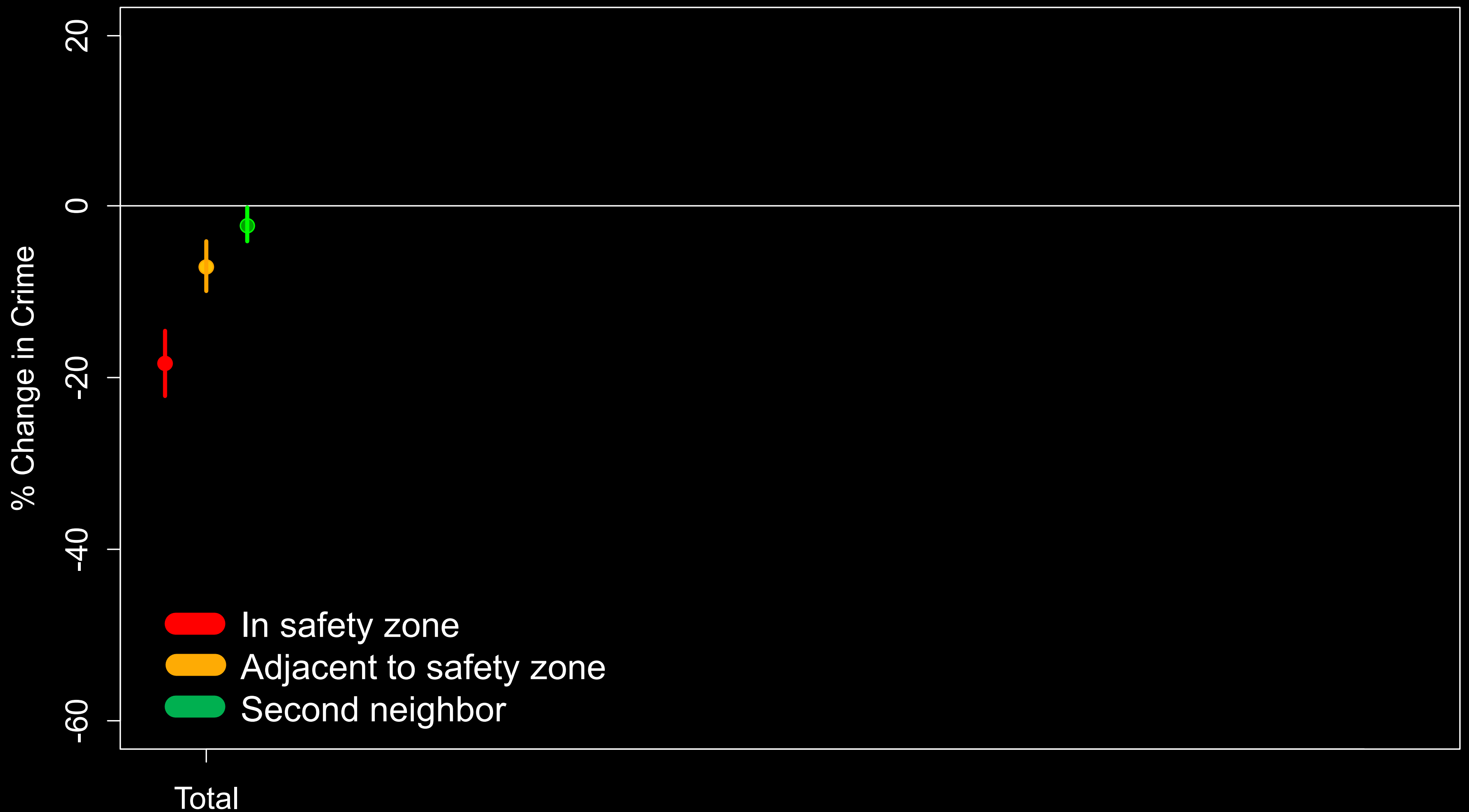
Fields Sep 2018

Stepped Wedge Design Detects Shifts in Crime Rates When CGIs Begin

$$\log(\lambda_{it}) = \beta_1 \text{InSZ}_{it} + \beta_2 \text{NextSZ}_{it} + \beta_3 \text{2ndNeighbor}_{it} + \gamma_i + s(t) + \alpha_{Q(t)}$$

- γ_i is an RD fixed effect
- $s(t)$ is a smooth 108 quarter crime trend
- $\alpha_{Q(t)}$ is a quarter fixed effect

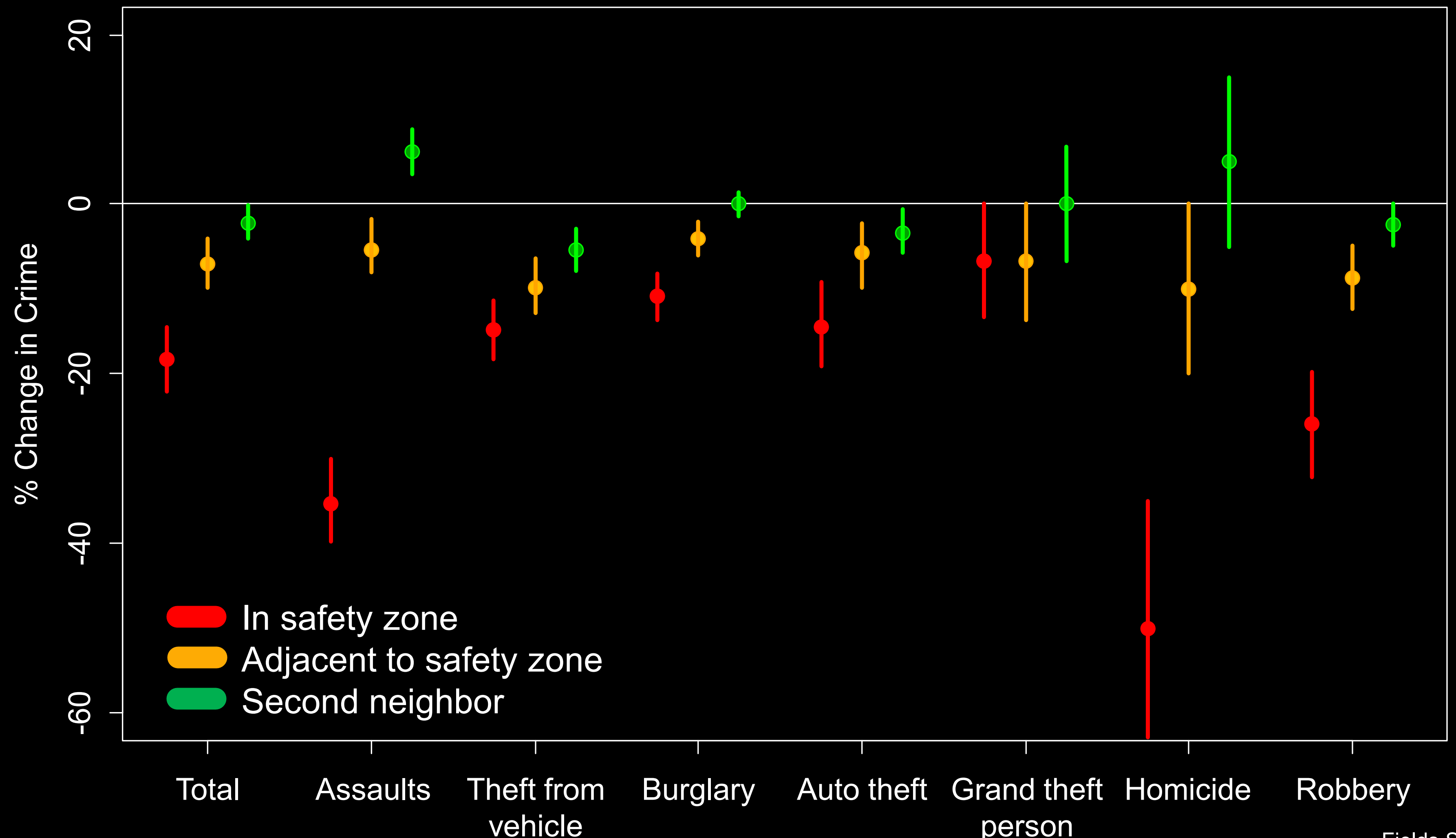
Neighborhoods Closest to Safety Zones See Largest Crime Decreases



Fields Sep 2018

Neighborhoods Closest to Safety Zones See Largest Crime Decreases

Neighborhoods Closest to Safety Zones See Largest Crime Decreases



Fields Sep 2018

Blending data from multiple sources, multiple decades aids discovery

- Data, in this case, were difficult to access
  - 2,300 printed pages
  - Old LAPD maps
- Compiling these data were essential for understanding the role of CGIs in crime prevention

- Early 2018, federal judge blocked the enforcement of the CGIs, citing due process violations

Data science to measure performance

Criminological Data Science

- Effect of gang injunctions on crime
- Benchmarking performance
- Officer risk factors for shooting
- Grams of marijuana in a joint

Create benchmarks to identify problematic system components

- Fair comparisons
  - Match an officer’s activities with activities conducted by other officers in the same time, place, and context
  - Existing systems group officers by just precinct or division and shift
- Outlier thresholds
  - Compute the probability that the data indicate an officer is a problem
  - Existing methods use fixed thresholds (more than 3 complaints) or traditional statistical choices (more than 2 or 3 standard deviations above the average)

Create benchmarks to identify problematic system components

- Fair comparisons
  - Match an officer's activities with activities conducted by other officers in the same time, place, and context
  - Existing systems group officers by just precinct or division and shift
- Outlier thresholds
  - Compute the probability that the data indicate an officer is a problem
  - Existing methods use fixed thresholds (more than 3 complaints) or traditional statistical choices (more than 2 or 3 standard deviations above the average)

We wish to determine if Officer A's stop patterns are unusual

Stop Characteristic		Officer A (%) n = 392	
Outcomes	Black	86	
	Frisk	12	

We know a lot about the time, place, and context of Officer A’s stops

Stop Characteristic		Officer A (%) n = 392	
Outcomes	Black	86	
	Frisk	12	
Month	January	3	
	February	4	
	March	8	
Day of the week	Monday	13	
	Tuesday	11	
	Wednesday	14	
Time of day	(4-6 pm]	9	
	(6-8 pm]	8	
	(8-10 pm]	23	
	(10 pm-12 am]	17	
Patrol borough	Brooklyn North	100	
Precinct	B	98	
	C	1	
Outside		96	
In uniform	Yes	99	
Radio run	Yes	1	

Find stops made by other officers occurring at the same time, place, and context

Stop Characteristic		Officer A (%) n = 392	Internal Benchmark (%) ESS = 3,676
Outcomes	Black	86	
	Frisk	12	
Month	January	3	**3**
	February	4	**4**
	March	8	**9**
Day of the week	Monday	13	**13**
	Tuesday	11	**10**
	Wednesday	14	**15**
Time of day	(4-6 pm]	9	**10**
	(6-8 pm]	8	**8**
	(8-10 pm]	23	**23**
	(10 pm-12 am]	17	**17**
Patrol borough	Brooklyn North	100	**100**
Precinct	B	98	**98**
	C	1	**1**
Outside		96	**94**
In uniform	Yes	99	**97**
Radio run	Yes	1	**3**

Benchmark also matches on fine location data

A higher percentage of people who Officer A stops are black

Stop Characteristic		Officer A (%) n = 392	Internal Benchmark (%) ESS = 3,676
Outcomes	Black	86	**55**
	Frisk	12	**11**
Month	January	3	**3**
	February	4	**4**
	March	8	**9**
Day of the week	Monday	13	**13**
	Tuesday	11	**10**
	Wednesday	14	**15**
Time of day	(4-6 pm]	9	**10**
	(6-8 pm]	8	**8**
	(8-10 pm]	23	**23**
	(10 pm-12 am]	17	**17**
Patrol borough	Brooklyn North	100	**100**
Precinct	B	98	**98**
	C	1	**1**
Outside		96	**94**
In uniform	Yes	99	**97**
Radio run	Yes	1	**3**

Idea: Reweight stops made by other officers to resemble this officer's stops

Example Officer

- Align their distributions
$$f(\mathbf{x}|t=1) = w(\mathbf{x}) f(\mathbf{x}|t=0)$$
- Solving for $w(\mathbf{x})$ yields the propensity score weight
$$w(\mathbf{x}) \propto \frac{P(t=1|\mathbf{x})}{1-P(t=1|\mathbf{x})}$$
- Need to estimate $P(t=1|\mathbf{x})$

Good benchmark construction is a computational challenge

- $\log \frac{P(t=1|\mathbf{x})}{P(t=0|\mathbf{x})} = \beta' \mathbf{x}$
- Select β to maximize

$$\sum_{i=1}^{N} t_i \beta' \mathbf{x}_i - \log(1 + \exp(\beta' \mathbf{x}_i))$$

- Easy to optimize even with large datasets
- Performs poorly with correlated features, interactions, saturation, threshold effects, …

Penalizing coefficients can produce a more stable model

- $\log \frac{P(t=1|\mathbf{x})}{P(t=0|\mathbf{x})} = \beta' \mathbf{x}$
- Instead select β to maximize

$$\sum_{i=1}^{N} t_i \beta' \mathbf{x}_i - \log(1 + \exp(\beta' \mathbf{x}_i)) - \lambda \sum_{j=1}^{d} |\beta_j|$$

- If λ=0 results in ordinary logistic regression
- If λ=∞ produces constant predictions, $\bar{t}$
- Stabilizes model for correlated features

Use a large, flexible class of basis functions

- $\log \frac{P(t=1|\mathbf{x})}{P(t=0|\mathbf{x})} = \beta' \mathbf{h}(\mathbf{x})$

$$\sum_{i=1}^{N} t_i \beta' \mathbf{h}(\mathbf{x}_i) - \log(1 + \exp(\beta' \mathbf{h}(\mathbf{x}_i))) - \lambda \sum_{j=1}^{d} |\beta_j|$$

- Let $h_1(\mathbf{x}),\ldots,h_K(\mathbf{x})$ be all piecewise constant functions of **x** and their interactions
- *K* is huge, *h* spans a large class of functional forms

Iterative, tree-structured search makes computation possible

$$\sum_{i=1}^{N} t_i \beta' \mathbf{h}(\mathbf{x}_i) - \log(1 + \exp(\beta' \mathbf{h}(\mathbf{x}_i))) - \lambda \sum_{j=1}^{d} |\beta_j|$$

- Relaxing λ from infinity to a little less than infinity is equivalent to modifying one of the βs
- β_j to modify is the one associated with the $h_j(\mathbf{x})$ most correlated with $t_i - \hat{P}(t = 1|\mathbf{x}_i)$
- Tree structured search can find $h_j(\mathbf{x})$ fast
- Then relax λ a little more…

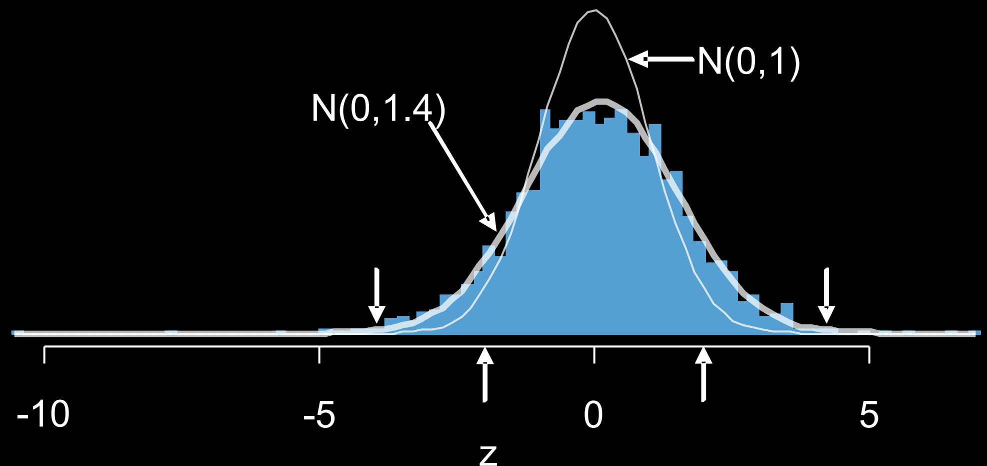
That computation produces a well-matched, fair, credible benchmark

Stop Characteristic		Officer A (%) n = 392	Internal Benchmark (%) ESS = 3,676
Outcomes	Black	86	**55**
	Frisk	12	**11**
Month	January	3	**3**
	February	4	**4**
	March	8	**9**
Day of the week	Monday	13	**13**
	Tuesday	11	**10**
	Wednesday	14	**15**
Time of day	(4-6 pm]	9	**10**
	(6-8 pm]	8	**8**
	(8-10 pm]	23	**23**
	(10 pm-12 am]	17	**17**
Patrol borough	Brooklyn North	100	**100**
Precinct	B	98	**98**
	C	1	**1**
Outside		96	**94**
In uniform	Yes	99	**97**
Radio run	Yes	1	**3**

Create benchmarks to identify problematic system components

- Fair comparisons
  - Match an officer's activities with activities conducted by other officers in the same time, place, and context
  - Existing systems group officers by just precinct or division and shift
- Outlier thresholds
  - Compute the probability that the data indicate an officer is a problem
  - Existing methods use fixed thresholds (more than 3 complaints) or traditional statistical choices (more than 2 or 3 standard deviations above the average)

Repeat for Every NYPD Officer Actively Involved in Stops

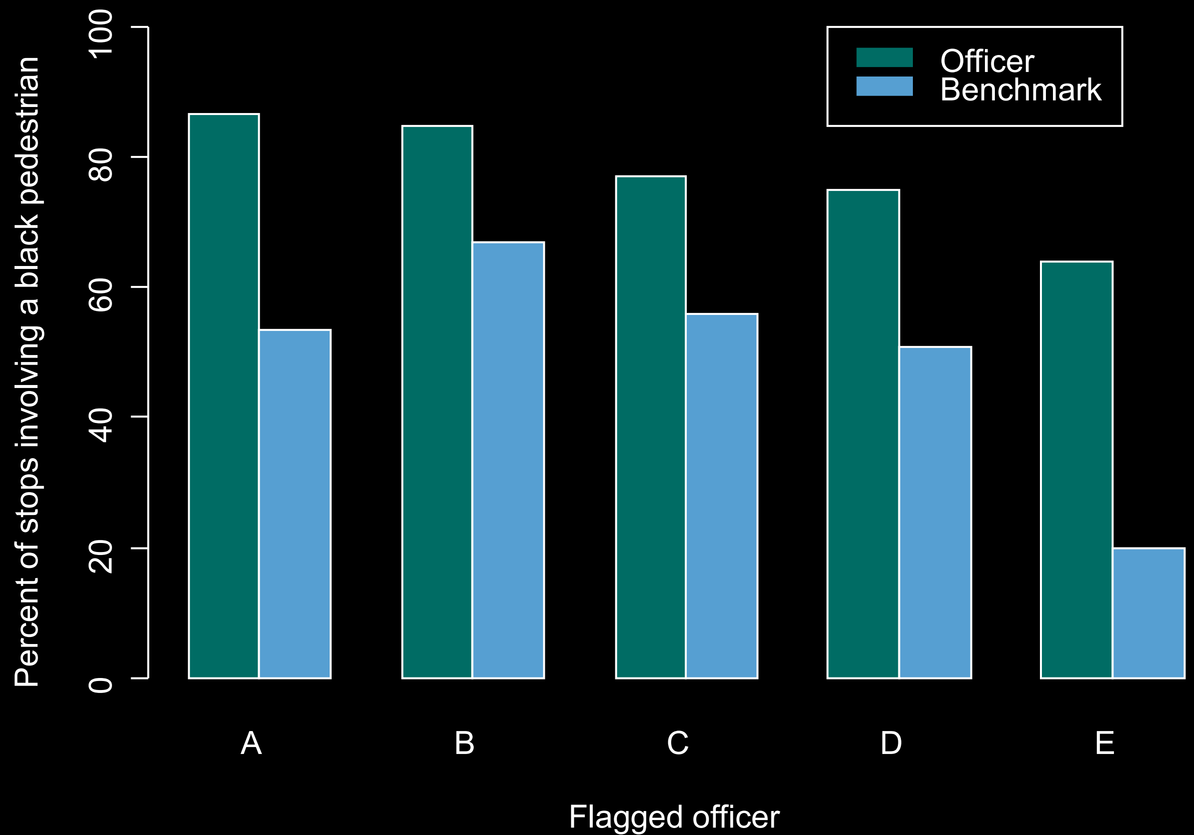


- $P(\text{problem}|z) = 1 - \frac{f(z|\text{no problem})f(\text{no problem})}{f(z)}$

$$\geq 1 - \frac{f_0(z)}{f(z)}$$

- Right tail consists of 5 officers with “problem officer” probabilities in excess of 50%
- Standard cutoff of z > 2.0 flags 242 officers, 90% of which have false discovery rate estimated to be greater than 0.999

Analysis flagged five officers overstopping black pedestrians



Benchmarking is applicable to numerous contexts… like opioid prescriptions

Patient features	Hospital X	Benchmark patients treated at other hospitals
Opioid Rx rate	63.6	
Days supply	8.5	
Male (%)	48.0	
Age at admission	49.9	
Primary admission reason (%)		
Circulatory	22.2	
Musculoskeletal	7.5	
Injury	11.2	
Diagnosis history (%)		
Arthritis	2.3	
Obesity	31.7	
Tobacco use	42.4	
Alcohol liver damage	13.6	
Prior prescriptions (%)		
Hypolipidemics	41.8	
Antihypertension	15.6	
NSAID	48.8	
Antidiabetes	30.9	
Antidepressants	45.5	

Created a set of benchmark patients who resemble Hospital X's patients…

Patient features	Hospital X	Benchmark patients treated at other hospitals
Opioid Rx rate	63.6	
Days supply	8.5	
Male (%)	48.0	47.1
Age at admission	49.9	49.5
Primary admission reason (%)		
Circulatory	22.2	21.8
Musculoskeletal	7.5	7.1
Injury	11.2	10.8
Diagnosis history (%)		
Arthritis	2.3	2.0
Obesity	31.7	32.4
Tobacco use	42.4	41.6
Alcohol liver damage	13.6	13.2
Prior prescriptions (%)		
Hypolipidemics	41.8	40.7
Antihypertension	15.6	15.1
NSAID	48.8	49.3
Antidiabetes	30.9	30.3
Antidepressants	45.5	45.3

Hospital X prescriptions are more frequent and larger than its benchmark

Patient features	Hospital X	Benchmark patients treated at other hospitals
Opioid Rx rate	63.6	36.1
Days supply	8.5	4.6
Male (%)	48.0	47.1
Age at admission	49.9	49.5
Primary admission reason (%)		
Circulatory	22.2	21.8
Musculoskeletal	7.5	7.1
Injury	11.2	10.8
Diagnosis history (%)		
Arthritis	2.3	2.0
Obesity	31.7	32.4
Tobacco use	42.4	41.6
Alcohol liver damage	13.6	13.2
Prior prescriptions (%)		
Hypolipidemics	41.8	40.7
Antihypertension	15.6	15.1
NSAID	48.8	49.3
Antidiabetes	30.9	30.3
Antidepressants	45.5	45.3

Fields Sep 2018

Data science to detect risks

Criminological Data Science

- Effect of gang injunctions on crime
- Benchmarking performance
- Officer risk factors for shooting
- Grams of marijuana in a joint

“There is virtually no empirical support for assertions that individual officer characteristics are measurably related to any type of performance in office” – Fyfe (1989)

Learn the factors affecting the probability of shooting

$$\log \frac{P(S=1|\mathbf{x},\mathbf{z})}{1-P(S=1|\mathbf{x},\mathbf{z})} = h(\mathrm{z}) + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_d x_d$$

- *S* is 1 if the officer shoots
- **x** are the officer's features (race, experience, record)
- **z** are the features of a particular scenario (kinds of suspects involved, location, and lighting)

Compare shooting and nonshooting officers on the same scene

1. Multiple officers on the scene
2. They all share the same environment features
3. Test whether officers with certain features are more likely to be the shooter

Consider the conditional likelihood of a shooting involving two officers

$$P(S_{\mathrm{v}} = 1, S_{\mathrm{w}} = 0 | S_{\mathrm{v}} + S_{\mathrm{w}} = 1, \mathbf{x}_{\mathrm{v}}, \mathbf{x}_{\mathrm{w}}, \mathbf{z}) = \frac{e^{\beta' x_{\mathrm{v}}}}{e^{\beta' x_{\mathrm{v}}} + e^{\beta' x_{\mathrm{w}}}}$$

- For shootings involving more officers the denominator becomes more complex
  - One shooting involved 12 officers, 8 shooters
  - Denominator had $\binom{12}{8} = 495$ terms

Utilized data on a review of three years of NYPD records

- Gathered data on all shooting incidents adjudicated in 2004, 2005, and 2006
- For each shooting I recorded
  - department ID numbers for shooters in the incident
  - department ID numbers for non-shooting officers that were witnesses or in the immediate vicinity of the shooting
- 106 incidents involving 150 shooting officers and 141 non-shooting officers
- Collected data on age, race, experience, education, training, and past performance

Do male officers shoot more?

Officer feature	Risk difference
Male	

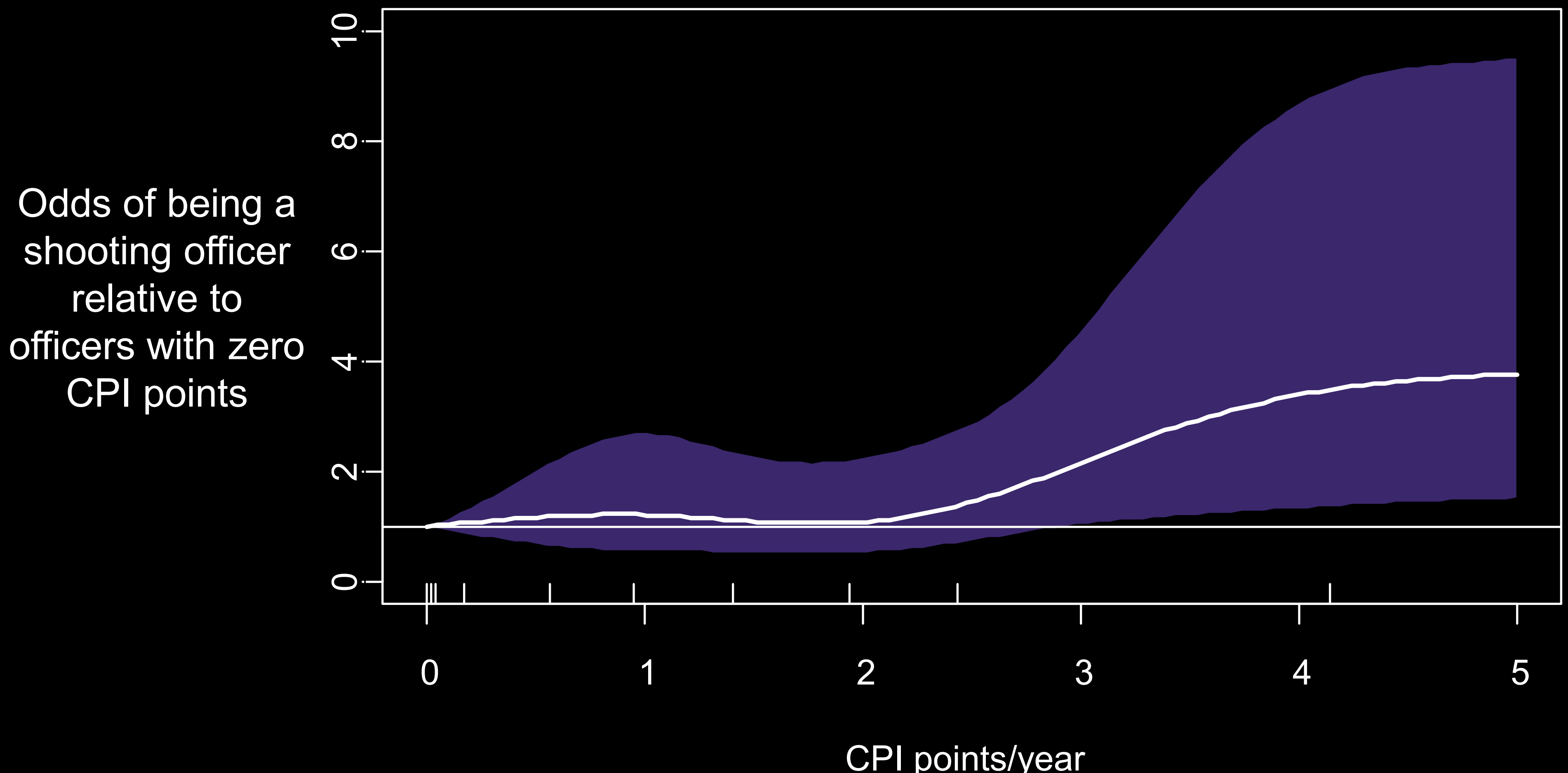
Male and female officers equally likely to shoot

Officer feature	Risk difference
Male	No difference

Race and age at recruitment associated with shooting

Officer feature	Risk difference
Male	No difference
Race	
White (reference)	
Black	+226%
Hispanic	No difference
Years at NYPD	No difference
Age when recruited	-11%
Education	No difference

Exceeding 3 CPI/year triples the shooting risk



Fields Sep 2018

Method generalizes to the number of rounds fired

- Many departments only document which officers shoot
- Model the shooting rate

$$\log \lambda(\mathbf{x}, \mathbf{z}) = h(\mathbf{z}) + \beta'\mathbf{x}$$

- Consider conditional likelihood

$$P(R_1 = r_1, R_2 = r_2 | R_1 + R_2 = r_1 + r_2, R_1 > 0, R_2 > 0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{z})$$

- The result involves some large, numerically unstable sums

$$\prod_{s=1}^{S} \frac{1}{\sum_{\sum i_p = \sum r_p, i_p > 0} \prod_{p=1}^{P_s - 1} \exp\left(\left(i_p - r_p\right)\beta'\left(\mathbf{x}_p - \mathbf{x}_{P_S}\right)\right)\left(i_1, i_2, \cdots, i_{P_s}\right)!}$$

Data science for fun… and policy

Criminological Data Science

- Effect of gang injunctions on crime
- Benchmarking performance
- Officer risk factors for shooting
- Grams of marijuana in a joint

Average weight of marijuana is an important, yet unknown, quantity

- Official estimates guess the average weight of marijuana in joint in order to
  - Project quantity of drugs imported by drug trafficking organizations
  - Project revenues from taxing marijuana
  - Estimate the effect of dose on health and behavioral outcomes
- Surveys of drug use ask about recent drug use, marijuana often reported as number of joints
- Half of marijuana users consume marijuana as joints

Commonly used estimate of 0.5g likely too large

- Brown and Silverman (1974) proposed a pricing model linking price to location, time, and volume

$$p_{ijk} \approx e^{\beta_j} e^{\alpha_k} v_{ijk}^{\gamma}$$

Arrestee Drug Abuse Monitoring collected drug market transaction data

- Data on 10,628 arrestees reporting marijuana price and quantity 2000-2010 in 43 counties
- Marijuana measured in grams (*n* = 5,845), ounces (*n* = 8,027), or joints (*n* = 2,230)

"Jointly" model loose and joint transactions, linking on price

- Brown-Silverman model suggests a log-normal model for loose marijuana purchase price

$$p_{ijk} \approx e^{\beta_j} e^{\alpha_k} v_{ijk}^{\gamma}$$

- Joints should also follow Brown-Silverman

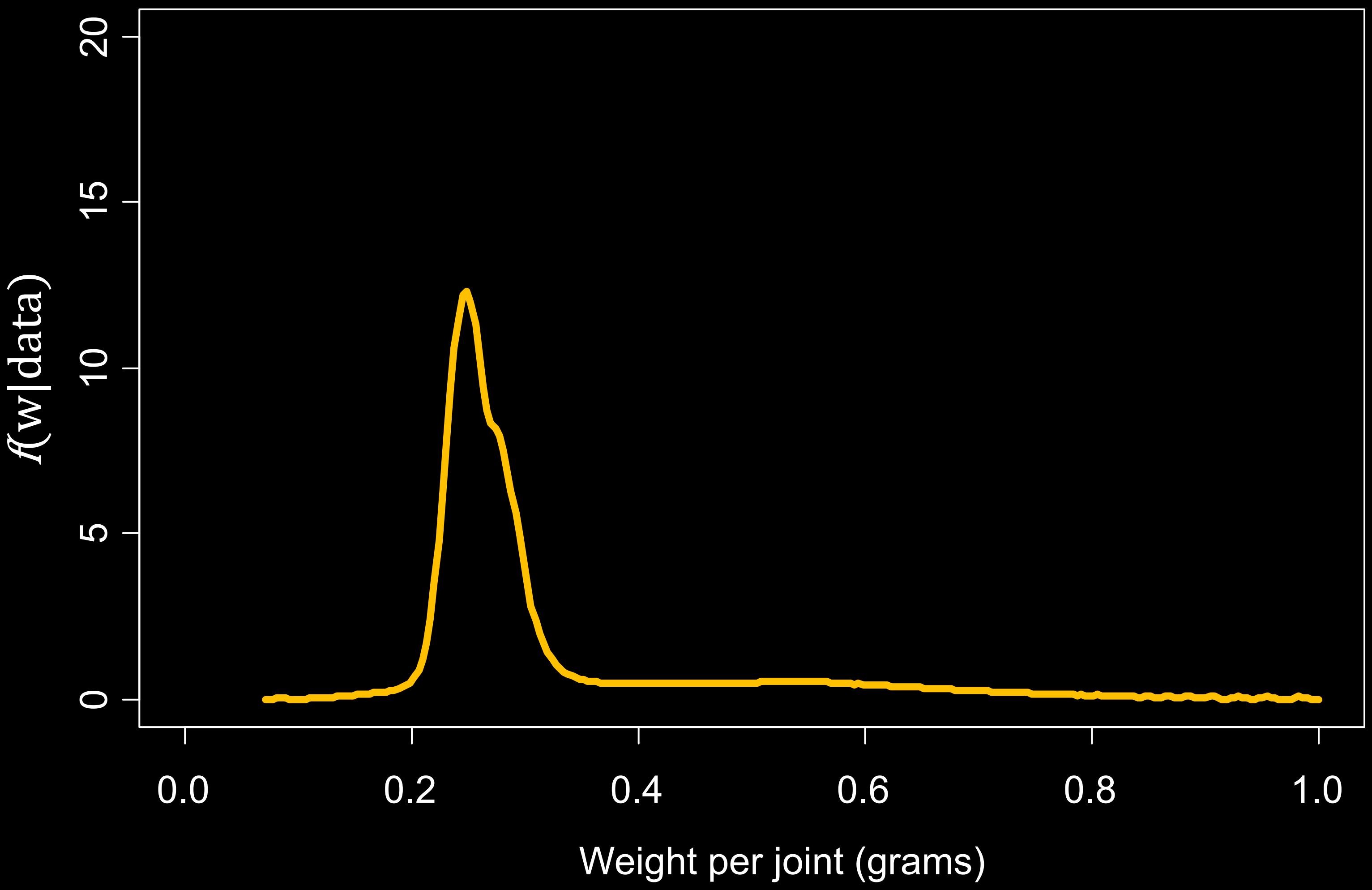
$$p_{ijk} \approx e^{\beta_j} e^{\alpha_k} \left(w_{ijk} n_{ijk}\right)^{\gamma}$$

- Estimate β_j, α_k, γ, and the distribution of w_{ijk}

Conclusion

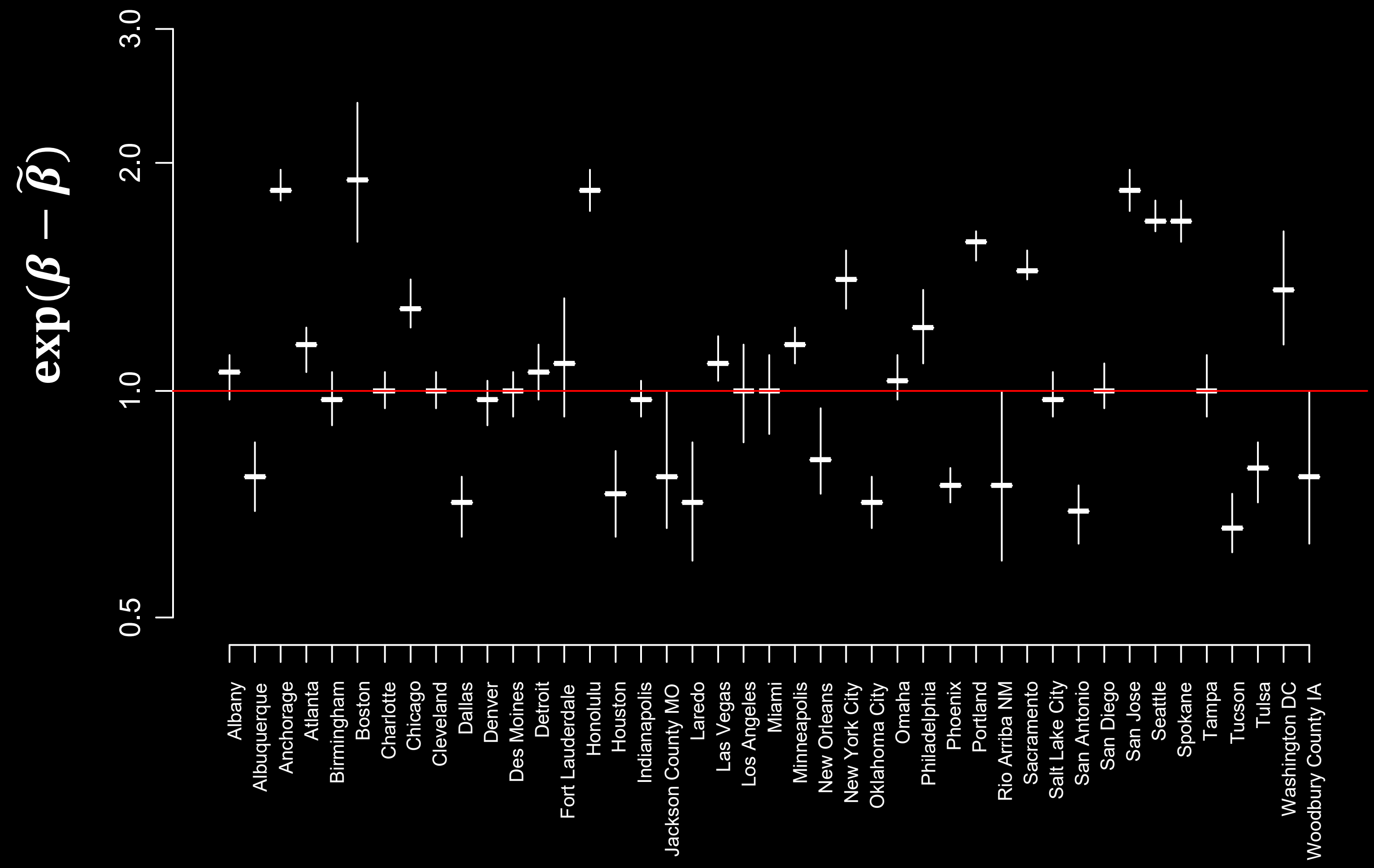
0.32 g/joint

Distribution of joint weights suggests multiple modes



Fields Sep 2018

Model allows comparison of drug market prices



Fields Sep 2018

Modern information technology now permits a massive assault on these problems at a level never before conceivable

- 1967 President's Commission on Law Enforcement and Administration of Justice

Criminological Data Science

Greg Ridgeway
Department of Criminology
Department of Statistics

G. Ridgeway (2018). "Policing in the Era of Big Data," *Annual Review of Criminology* 1:401-419.

G. Ridgeway (2019). "Experiments in Criminology," *Annual Review of Statistics and Its Applications* 6.