Combining Data and Crime

Sharon Lohr

March 7, 2018

sharon.lohr@asu.edu

Federal Statistics

- Increased nonresponse to surveys
- Decreased funding
- Demand for more granular data
  - Faster, more frequent
  - More geographic detail
- Demand for more privacy

Crime Statistics

- Brief history
- Quality of current data sources
- Using multiple sources of data
- Measuring accuracy
- Unintended consequences?

Why Crime?

Book for general public:

Measuring Crime: Behind the Statistics.

ASA/CRC Press Series on Statistical Reasoning in Science and Society

In print, late 2018

Edith Abbott

- PhD Economics 1905
- Hull House resident
- First female dean of a US graduate school (U Chicago, 1924)
- Statistician, Chicago City Council Committee on Crime, 1914-15

www.ssa.uchicago.edu/edith-abbott

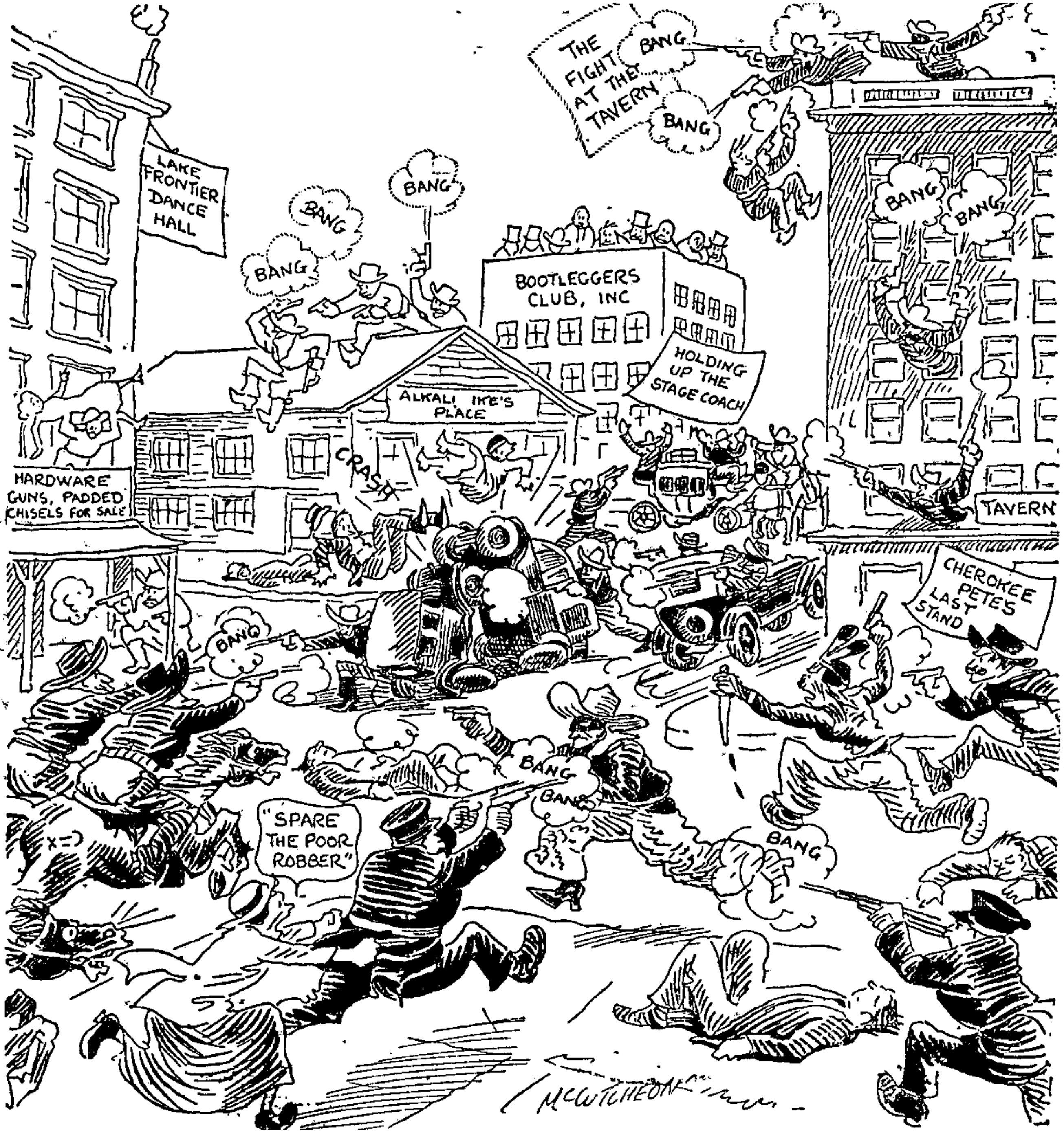
Statistics Relating to Crime in Chicago

by

MISS EDITH ABBOTT, Ph. D., CHICAGO SCHOOL OF CIVICS AND PHILANTHROPY

Abbott (1915)

- Need accurate statistics:
  - Use police resources efficiently
  - Dispel myths about crime
- Vital stats: model for crime stats
- Comparison groups: identify outliers
- Bring together statistics about crime from various sources

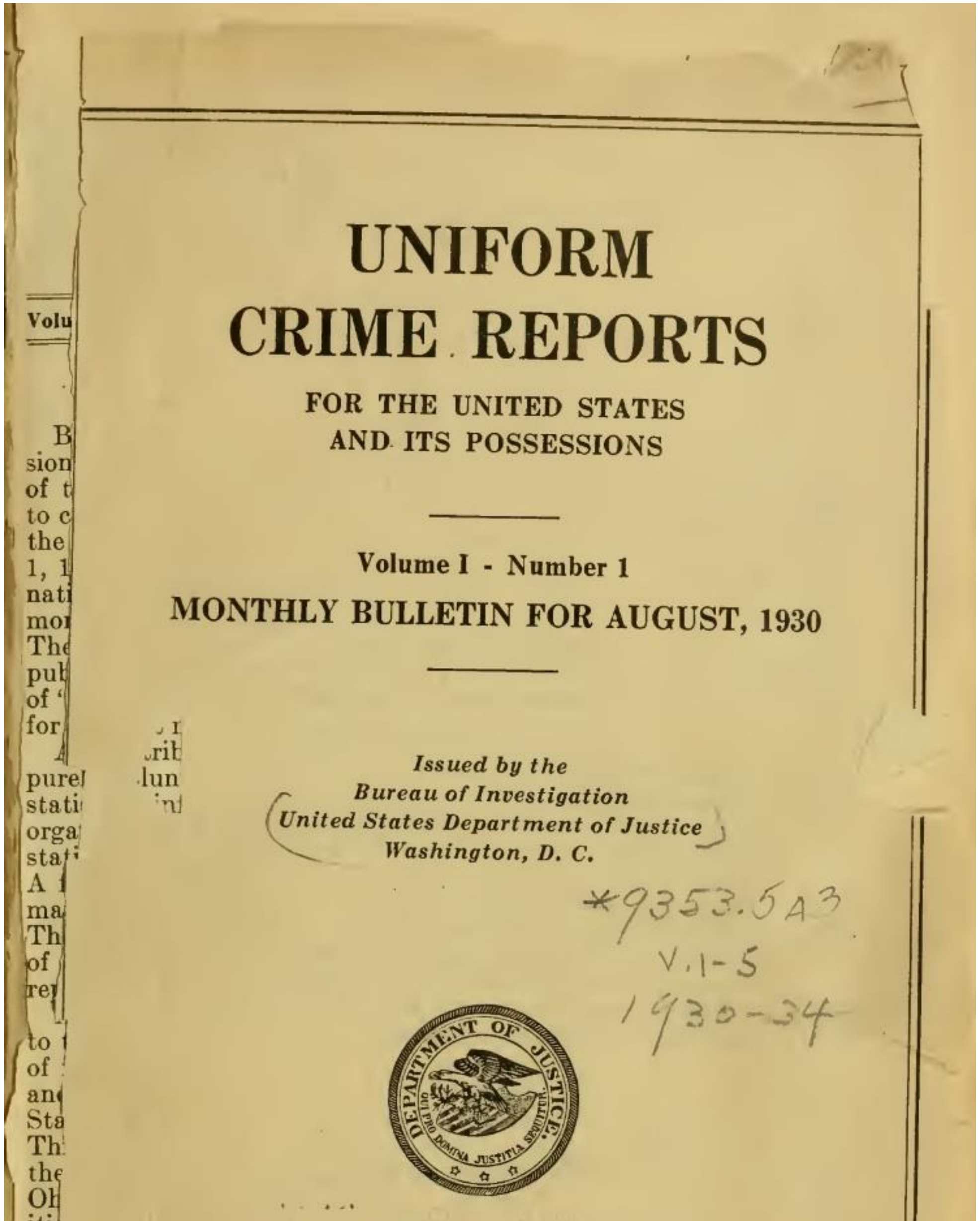

LAKE FRONTIER DANCE HALL
BANG
BANG
BANG
THE FIGHT AT THE TAVERN
BANG
BANG
BANG
BANG
BOOTLEGGERS CLUB, INC
ALKALI IKE'S PLACE
HOLDING UP THE STAGE COACH
CRASH
HARDWARE GUNS, PADDED CHISELS FOR SALE
TAVERN
CHEROKEE PETE'S LAST STAND
BANG
BANG
BANG
"SPARE THE POOR ROBBER"
BANG
McCUTCHEON

1920s

CHICAGO –
AS THE REST
OF THE WORLD
MUST IMAGINE IT

Chicago Tribune,
1925

1930 Uniform Crime Reports



UNIFORM
CRIME REPORTS

FOR THE UNITED STATES
AND ITS POSSESSIONS

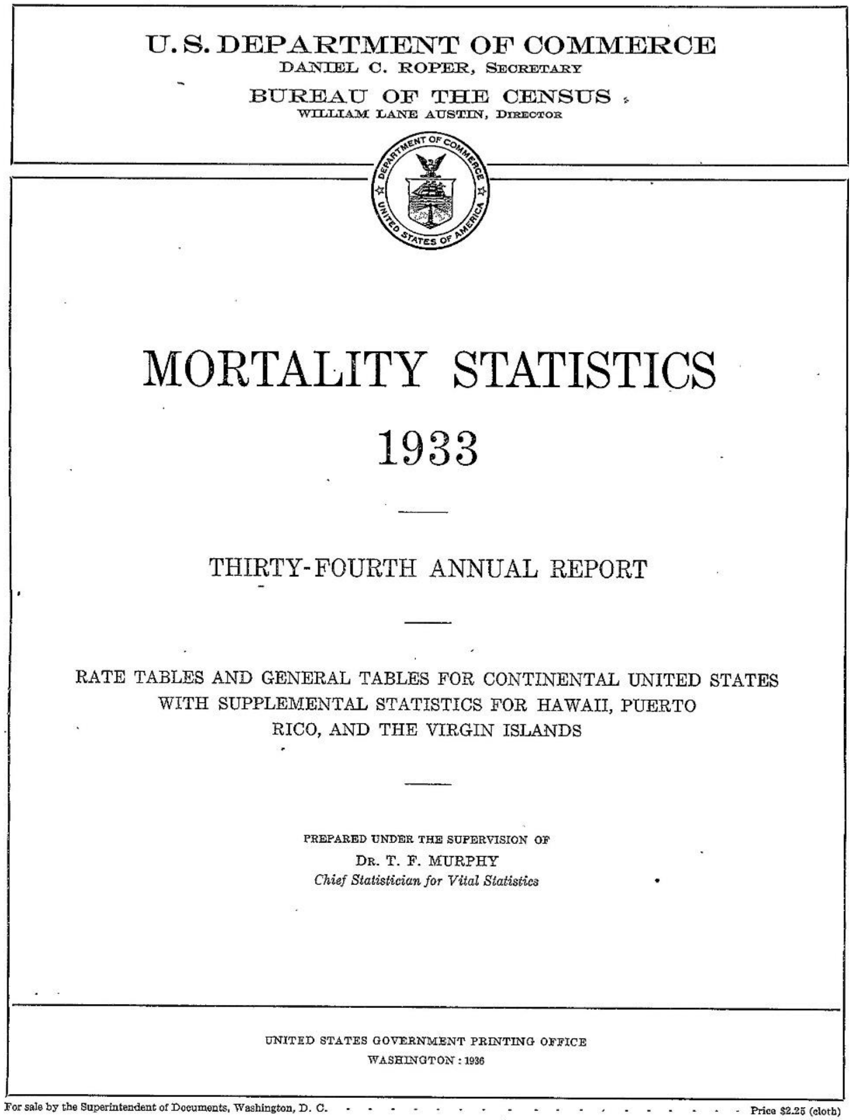
Volume I - Number 1
MONTHLY BULLETIN FOR AUGUST, 1930

Issued by the
Bureau of Investigation
United States Department of Justice
Washington, D. C.

*9353.5A3
V.1-5
1930-34

- Offenses known to police
- Uniform crime definitions
- Originally only counts
- Crime details 1961, 1991

Vital Statistics



U. S. DEPARTMENT OF COMMERCE
DANIEL C. ROPER, SECRETARY
BUREAU OF THE CENSUS
WILLIAM LANE AUSTIN, DIRECTOR

MORTALITY STATISTICS
1933

THIRTY-FOURTH ANNUAL REPORT

RATE TABLES AND GENERAL TABLES FOR CONTINENTAL UNITED STATES
WITH SUPPLEMENTAL STATISTICS FOR HAWAII, PUERTO
RICO, AND THE VIRGIN ISLANDS

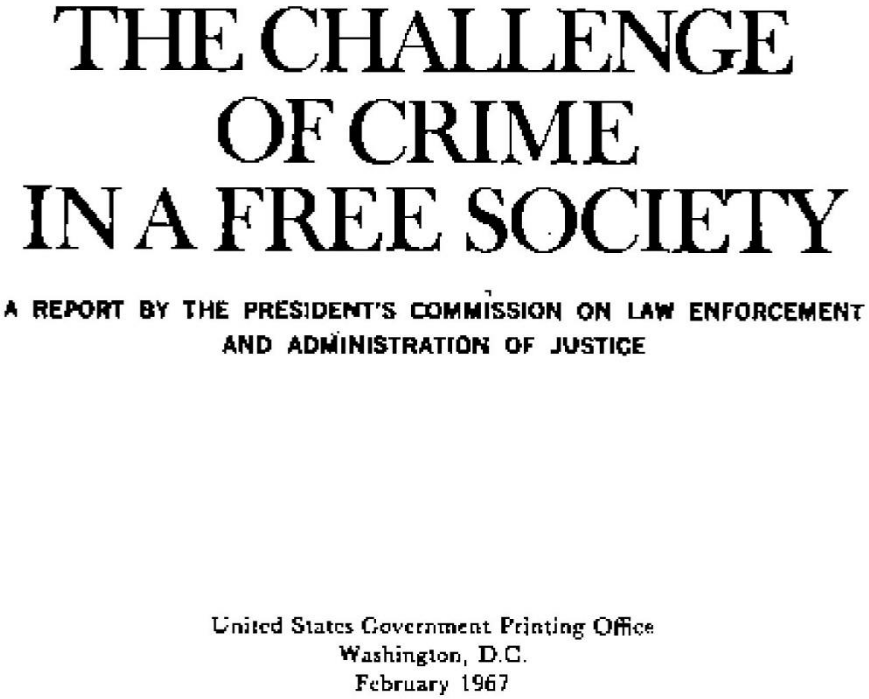
PREPARED UNDER THE SUPERVISION OF
DR. T. F. MURPHY
Chief Statistician for Vital Statistics

UNITED STATES GOVERNMENT PRINTING OFFICE
WASHINGTON : 1936

For sale by the Superintendent of Documents, Washington, D. C. - - - - - - - - - - - - - - - - - - - Price $2.25 (cloth)

- 1933: All states
- Homicide only

1967



THE CHALLENGE
OF CRIME
IN A FREE SOCIETY

A REPORT BY THE PRESIDENT'S COMMISSION ON LAW ENFORCEMENT
AND ADMINISTRATION OF JUSTICE

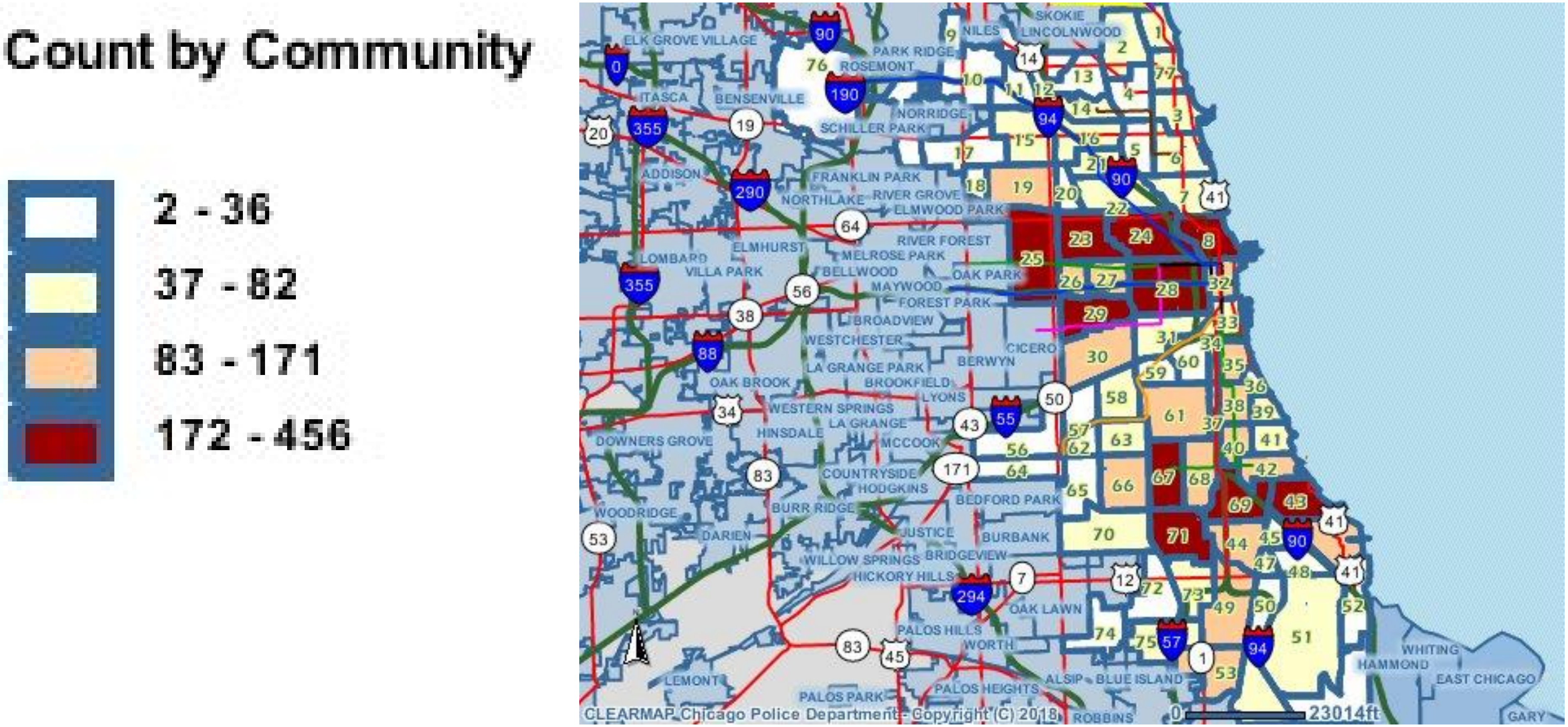
United States Government Printing Office
Washington, D.C.
February 1967

- “No one way of describing crime describes it well enough.”
- Fraud
- Survey: capture crimes not in UCR
- National
  Crime
  Victimization
  Survey

1990s: CompStat

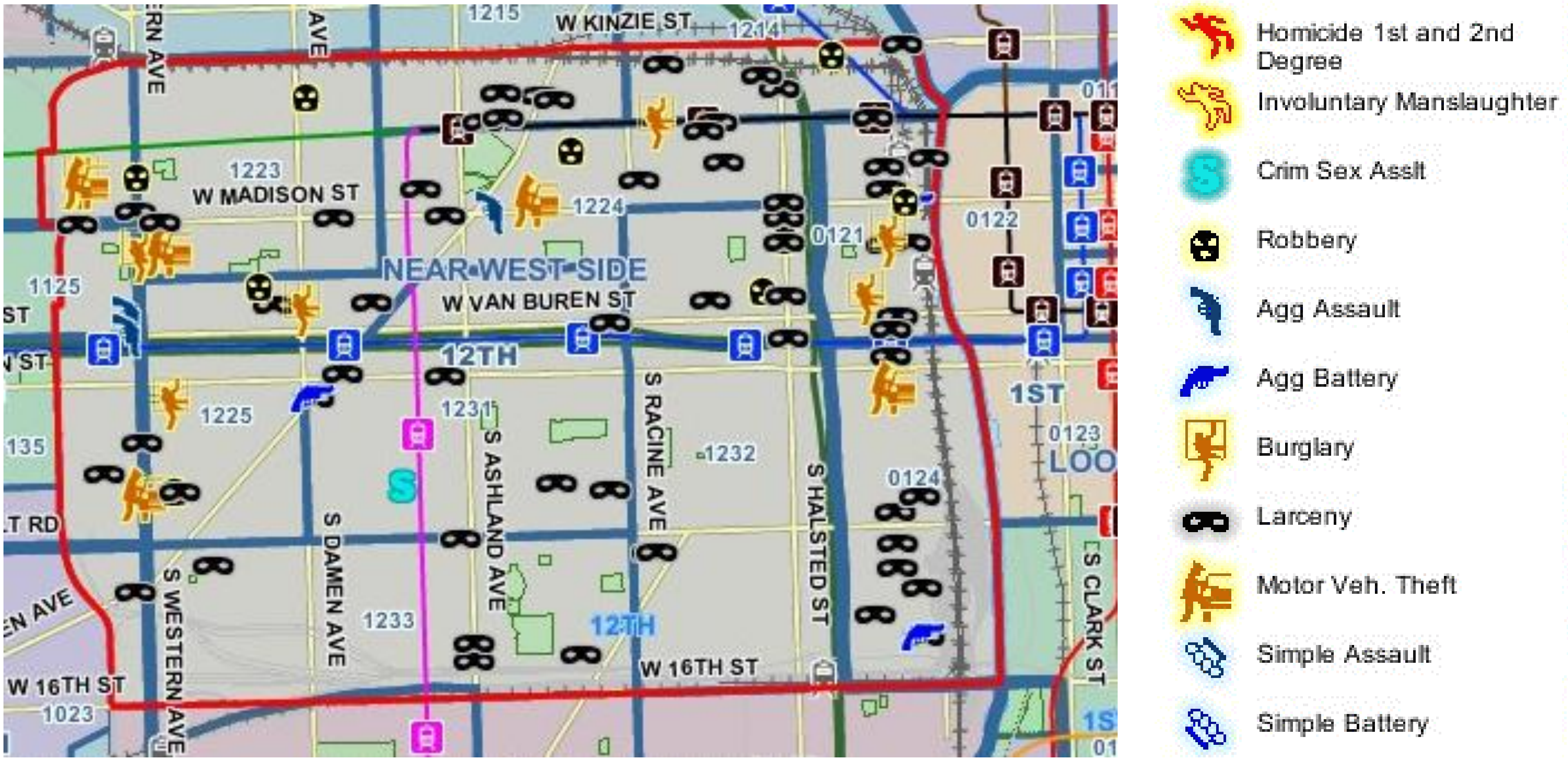
- Computer-intensive predictive policing
- Data available
  - Quickly
  - Geographic detail

Chicago Violent Crime


Count by Community
2 - 36
37 - 82
83 - 171
172 - 456
ELK GROVE VILLAGE
ITASCA
BENSENVILLE
ADDISON
LOMBARD
VILLA PARK
ELMHURST
NORTHLAKE
FRANKLIN PARK
RIVER GROVE
ELMWOOD PARK
RIVER FOREST
MELROSE PARK
BELLWOOD
MAYWOOD
FOREST PARK
OAK PARK
BROADVIEW
WESTCHESTER
CICERO
BERWYN
LA GRANGE PARK
BROOKFIELD
LYONS
OAK BROOK
WESTERN SPRINGS
LA GRANGE
HINSDALE
MCCOOK
DOWNERS GROVE
COUNTRYSIDE
HODGKINS
BEDFORD PARK
WOODRIDGE
BURR RIDGE
DARIEN
JUSTICE
BURBANK
WILLOW SPRINGS
BRIDGEVIEW
HICKORY HILLS
OAK LAWN
PALOS HILLS
WORTH
LEMONT
PALOS PARK
PALOS HEIGHTS
ALSIP
BLUE ISLAND
ROBBINS
ROSEMONT
PARK RIDGE
NORRIDGE
SCHILLER PARK
NILES
SKOKIE
LINCOLNWOOD
WHITING
HAMMOND
EAST CHICAGO
GARY
CLEARMAP Chicago Police Department - Copyright (C) 2018
0
23014ft

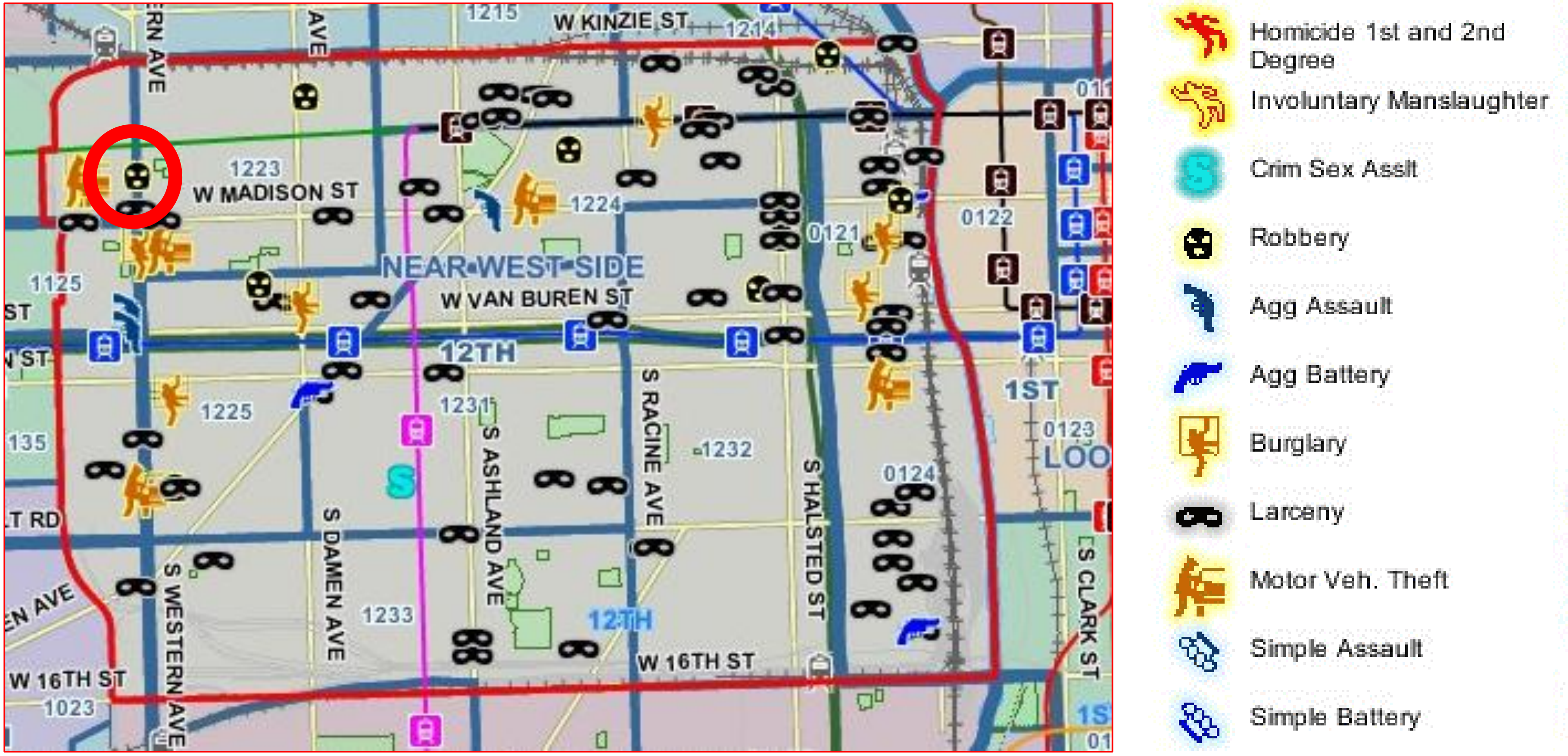
Violent crime past 90 days, Feb 18, 2018

Chicago Near West Side, Feb 5 to 18


W KINZIE ST
1215
1214
1223
W MADISON ST
1224
0121
0122
NEAR WEST SIDE
W VAN BUREN ST
1125
12TH
1ST
1231
1225
1232
0123
0124
S ASHLAND AVE
S RACINE AVE
S HALSTED ST
S DAMEN AVE
S WESTERN AVE
S CLARK ST
1233
12TH
W 16TH ST
W 16TH ST
1023
Homicide 1st and 2nd Degree
Involuntary Manslaughter
Crim Sex Asslt
Robbery
Agg Assault
Agg Battery
Burglary
Larceny
Motor Veh. Theft
Simple Assault
Simple Battery

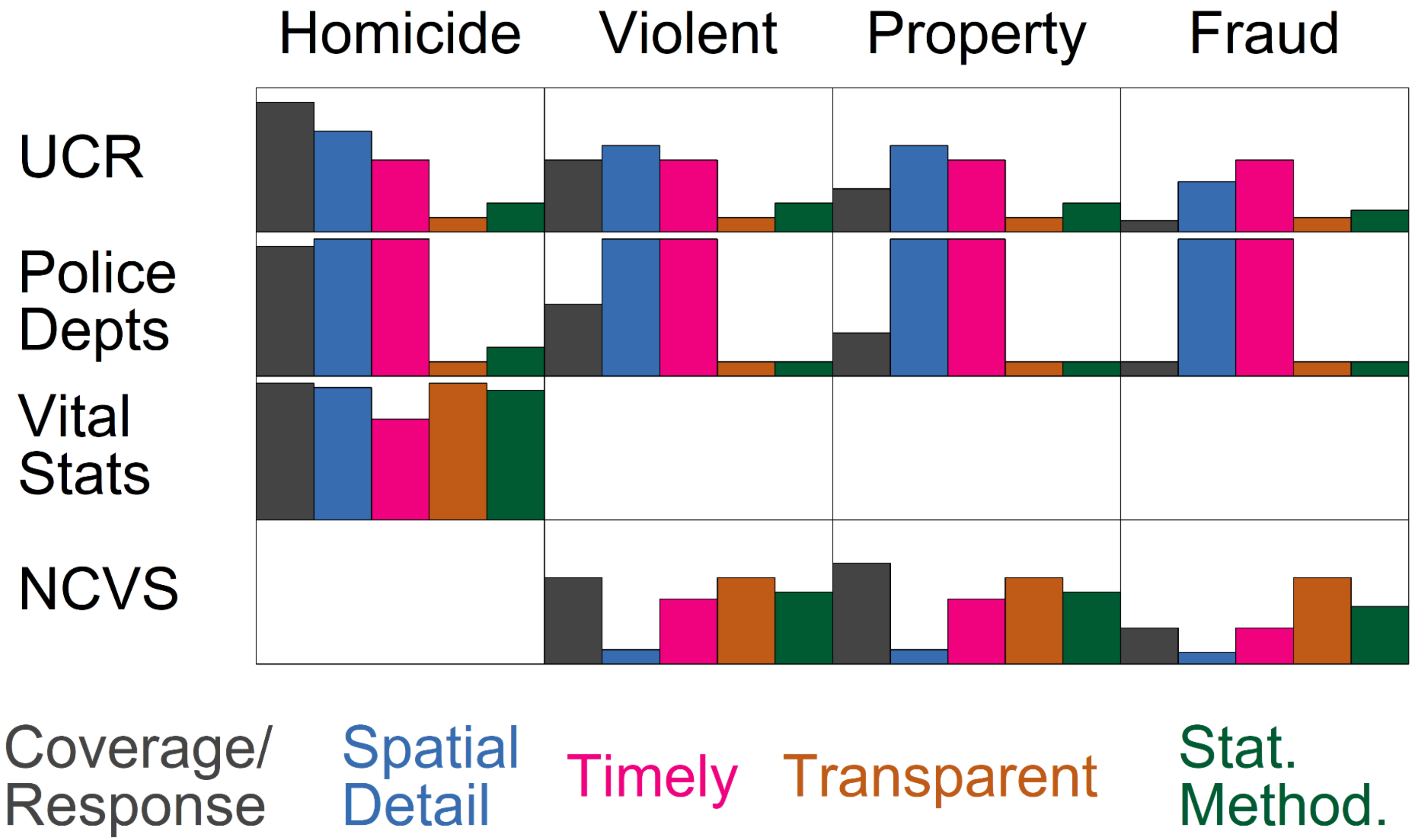
gis.chicagopolice.org

Chicago Near West Side, Feb 5 to 18


1215
W KINZIE ST
1214
ERN AVE
AVE
011
1223
W MADISON ST
1224
0121
0122
NEAR WEST SIDE
W VAN BUREN ST
1125
ST
N ST
12TH
1ST
1225
1231
135
S ASHLAND AVE
S RACINE AVE
1232
S HALSTED ST
0123
LOO
0124
T RD
S DAMEN AVE
S WESTERN AVE
1233
12TH
S CLARK ST
EN AVE
W 16TH ST
W 16TH ST
1023
1S
01
Homicide 1st and 2nd Degree
Involuntary Manslaughter
Crim Sex Assit
Robbery
Agg Assault
Agg Battery
Burglary
Larceny
Motor Veh. Theft
Simple Assault
Simple Battery

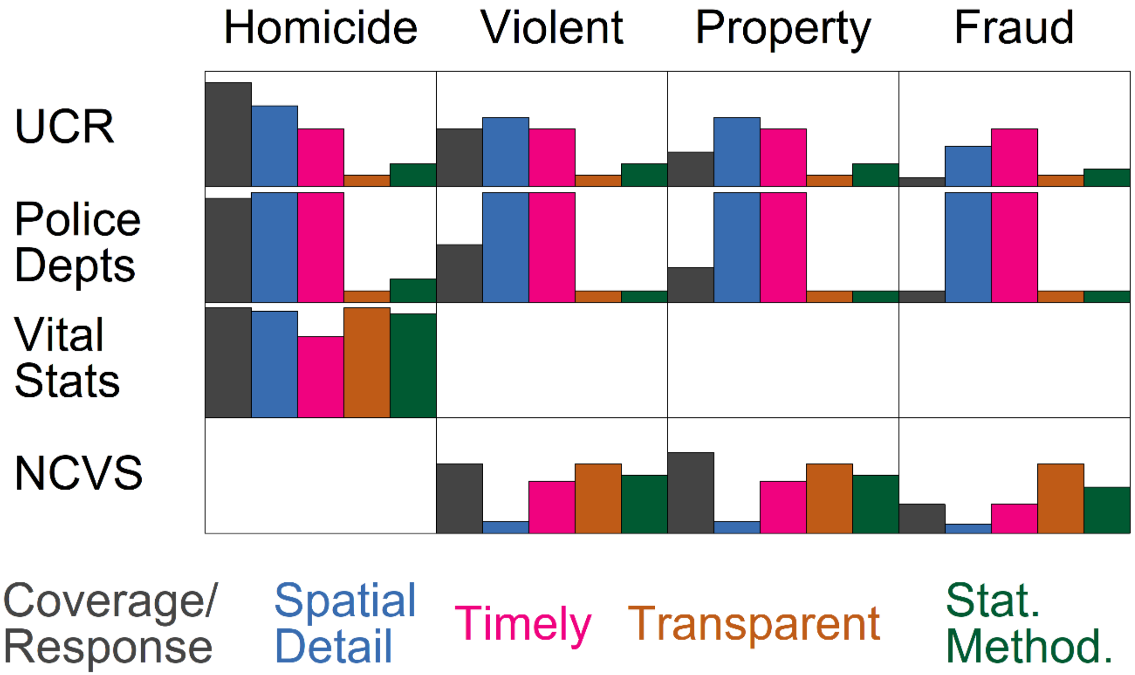
DATE	BLOCK	CODE	CRIME_TYPE	SECONDARY	ARREST	LOCATION	DOMESTIC	BEAT
Fri, 09 Feb 2018 11:29:00	100 N WESTERN AVE	031A	ROBBERY	ARMED: HANDGUN	N	GAS STATION	N	1222

Data Sources: Quality



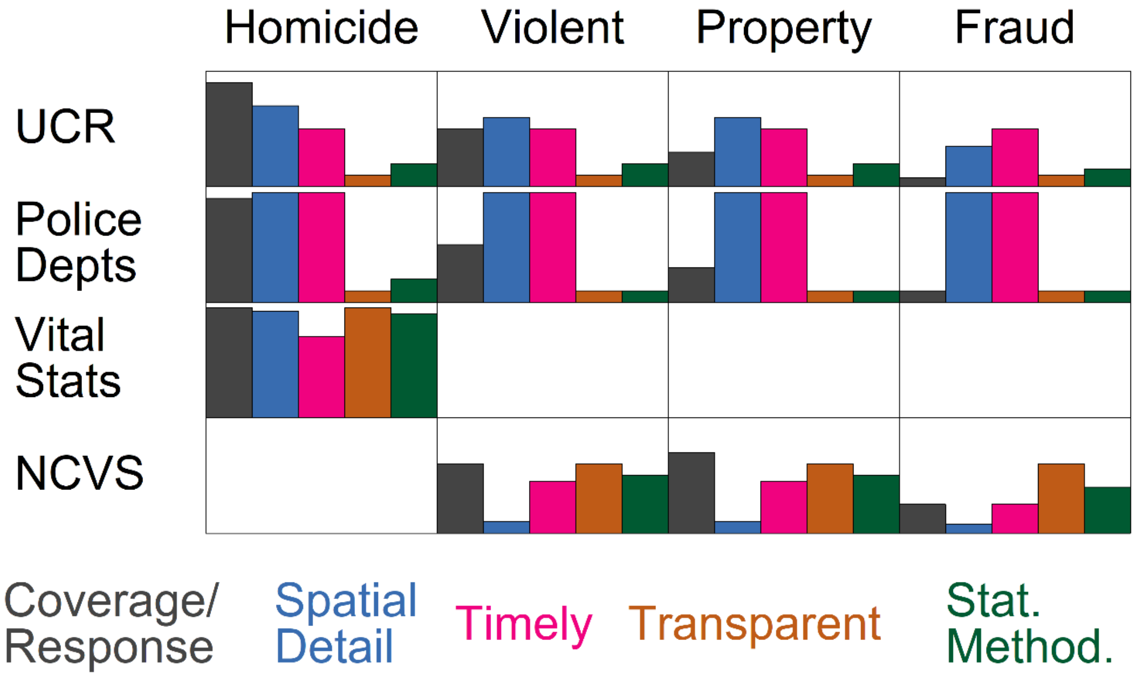
UCR

- Voluntary
- Measurement error
- Only crimes recorded by police
- No measures of uncertainty



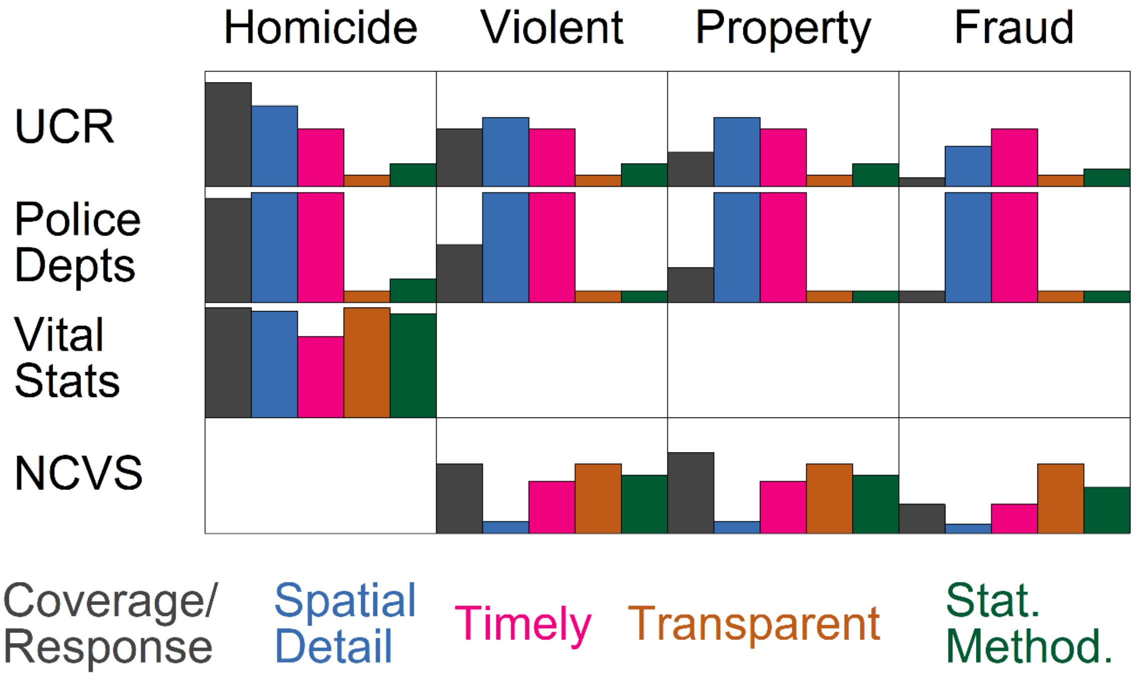
Police Depts

- Only large depts.
- Similar to UCR
- More detail
- Faster
- Algorithms secret



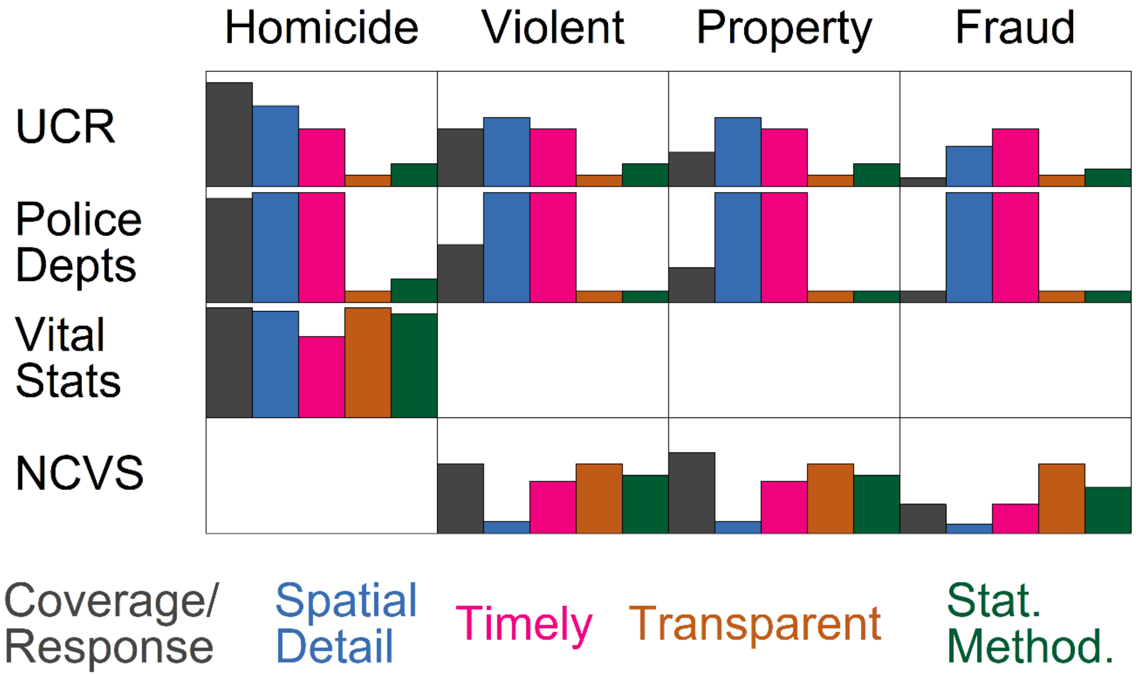
Vital Statistics

- Best data we have
- QI program
- National Violent Death Reporting System
- Homicide

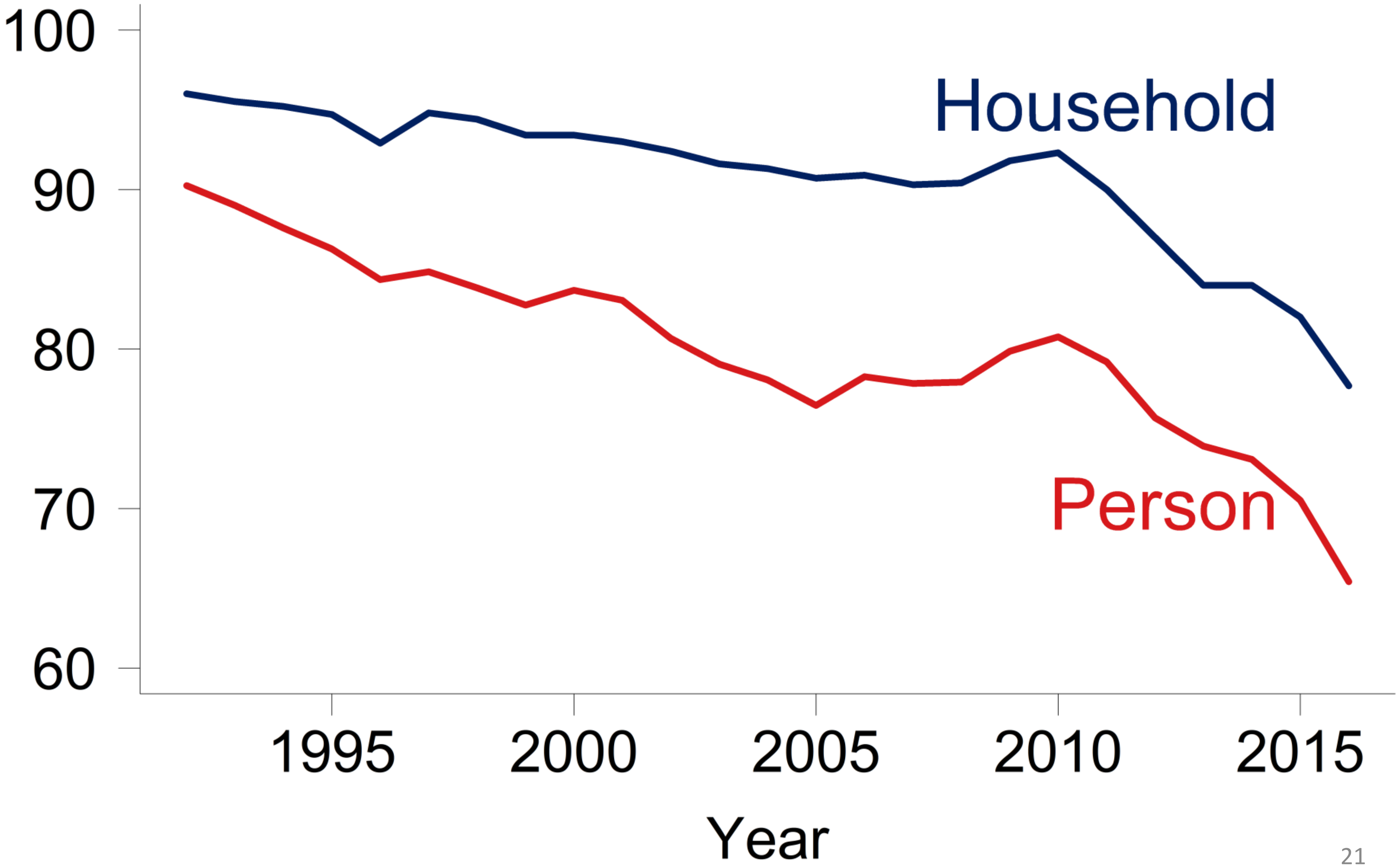


Natl Crime Victimization Survey

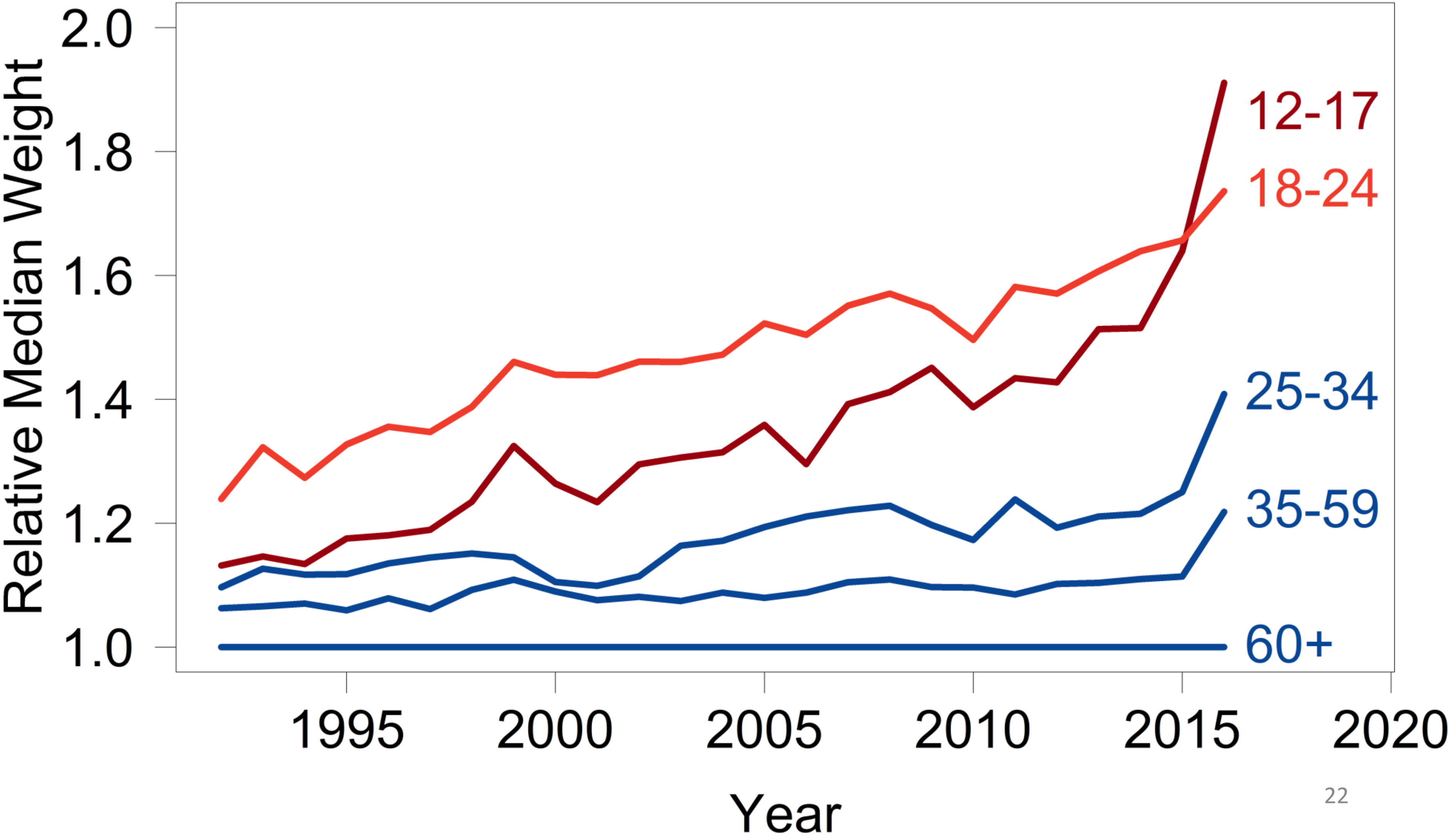
- Crimes not in other sources
- Small n for victims



NCVS Response Rates



$$\text{NCVS: } \frac{\text{Median weight of age group}}{\text{Median weight of age 60+}}$$



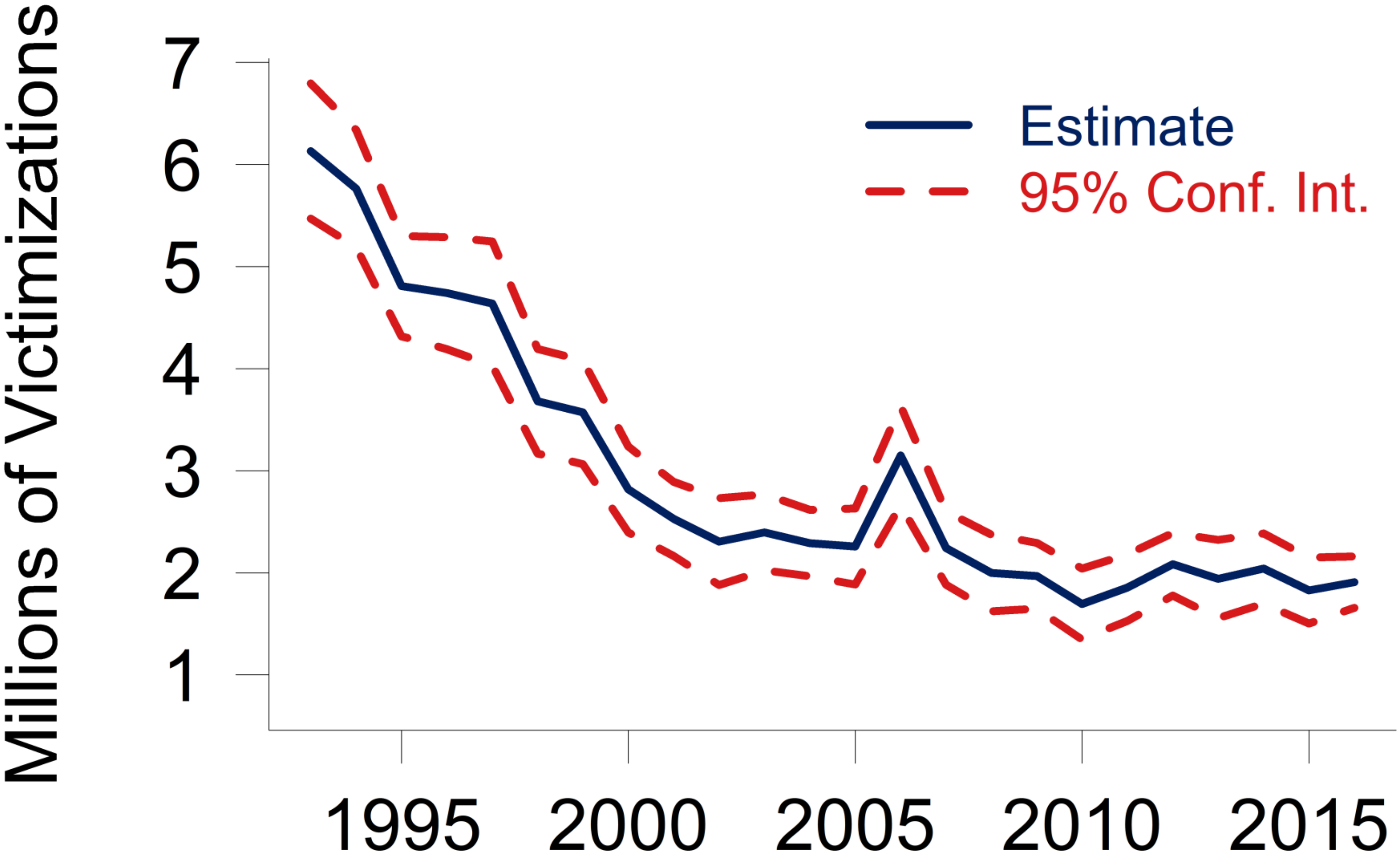
At what response rate is design-based SE invalid?

W. Edwards Deming

“A sample that is **95** or **98** percent a probability-sample ... **may** still be an excellent sample, although it is important to investigate the remaining 5 or even 2 percent as soon as possible.”

www.deming.org

NCVS Rape, robbery, agg. assault



NCVS Standard Errors

- Too small
- Only sampling error
- Poststratification
  - lowers SE further
  - assumes bias is removed

- UCR, police stats don’t attempt SE

Common and special causes

- 2006, 2016 anomalies treated as special causes
  - New areas
  - New interviewers
- These are systems features
- Standard errors need to include causes of variation from the system

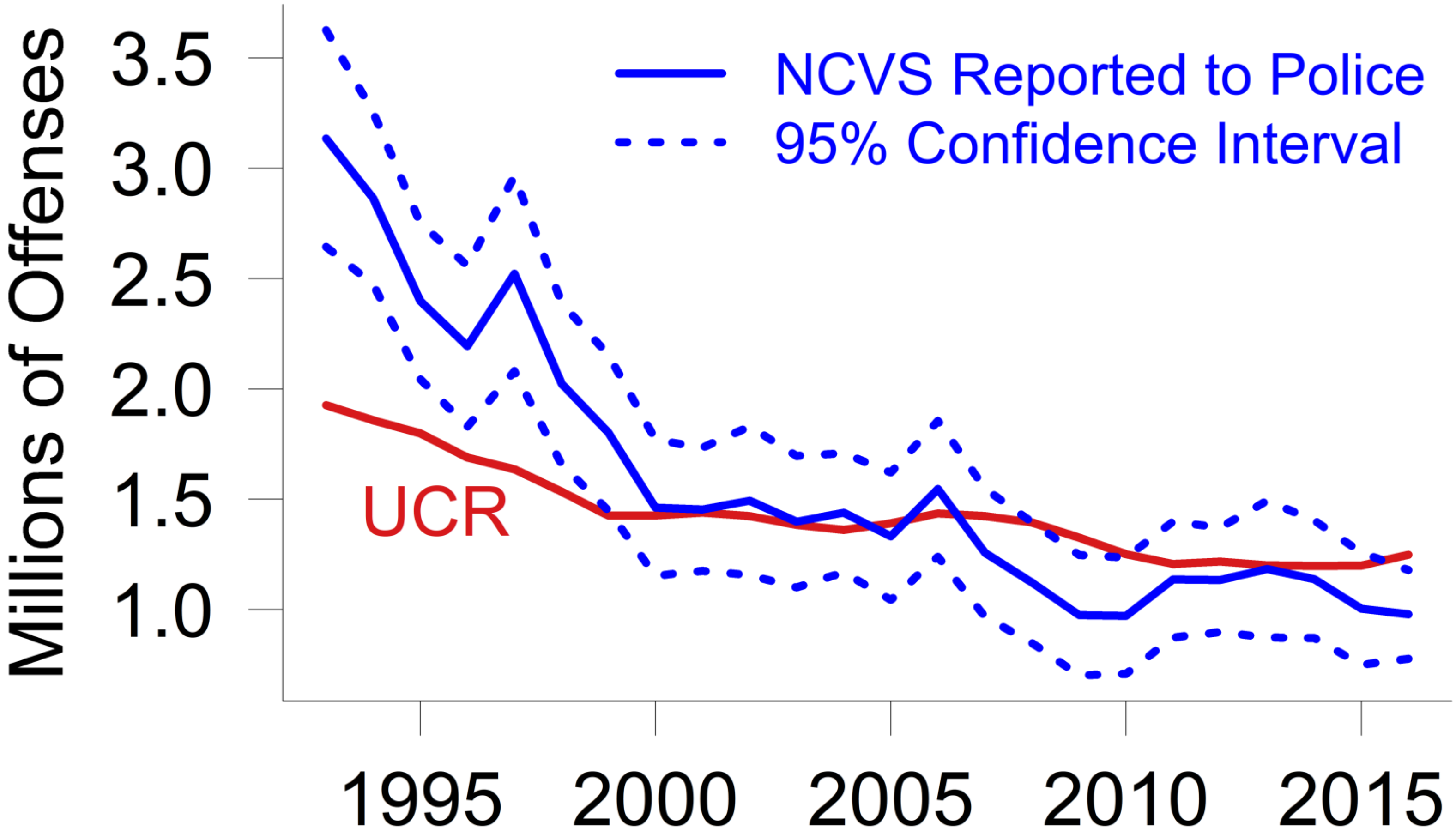
How?

- Nonresponse adjustments use models
- Include model uncertainty in SE
    - Spiegelhalter & Riesch (2011), risk analysis
    - Lohr & Brick (2017), Bayesian model averaging (1936 *Literary Digest* poll)
    - Always larger than design-based SE

How?

- Measurement error models
- Hierarchical models
- Use multiple sources
  - Thurs 10:15: NAS report on methods for combining data
  - Lohr & Raghunathan (2017), *Statistical Science*

Rape, robbery, agg. assault

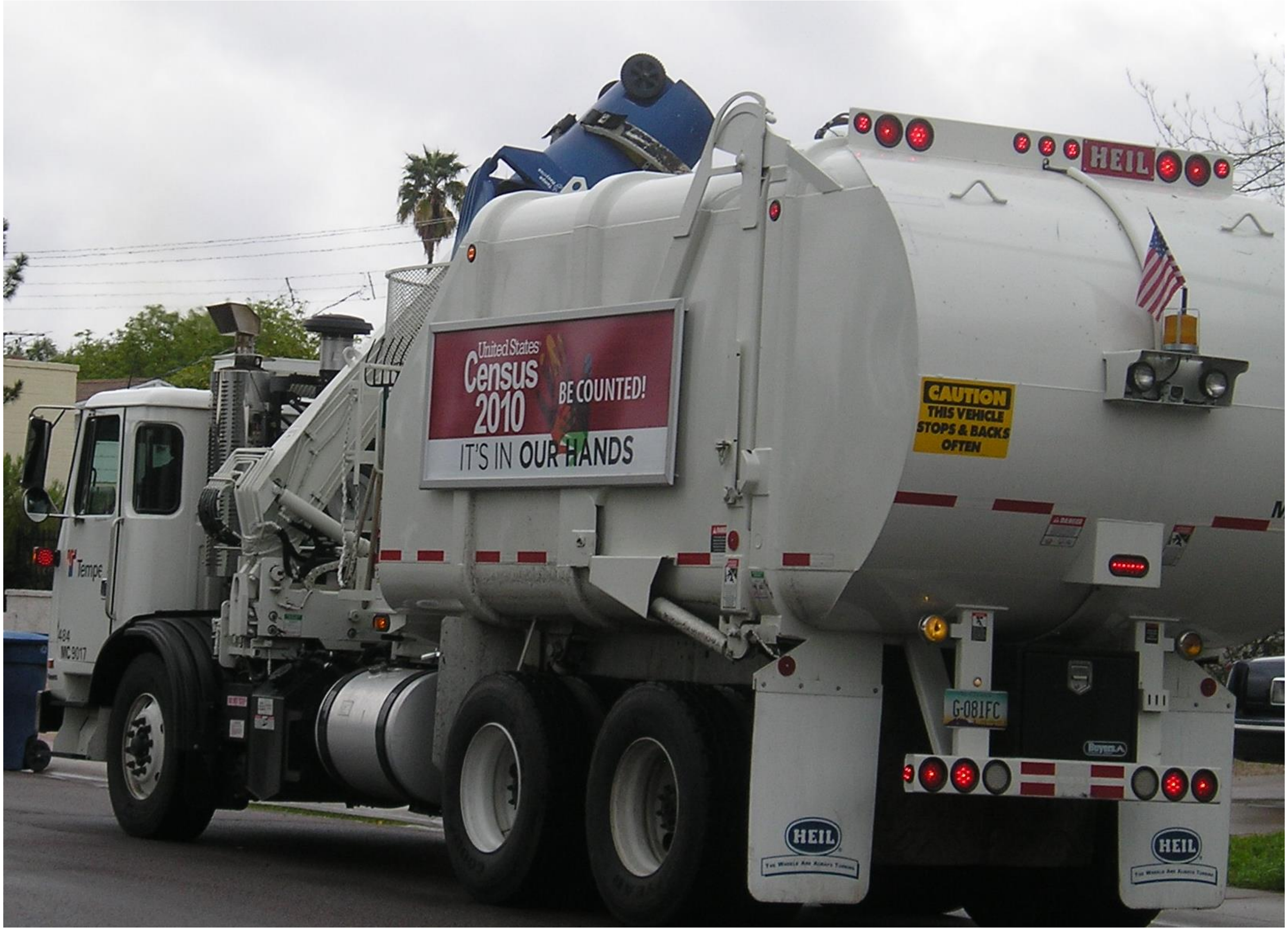


	In NCVS	Not in NCVS
In UCR	Reported to police Except →	Homicide Reported to police against: Children, institutionalized, commercial establishments
Not In UCR	Not reported to police Except →	Not reported to police: Children, institutionalized, commercial establishments Not represented in NCVS: Nonresponse Measurement error

Multiple sources

- Statistics from merged data
- Explore error properties
- Present alternative views
- Diversity is a strength
- Alternative sources for lower right quadrant, fraud
- 1960s-70s: expts, checks

Unintended Effects?


United States Census 2010
BE COUNTED!
IT'S IN OUR HANDS
CAUTION THIS VEHICLE STOPS & BACKS OFTEN
HEIL
Tempe
G-081FC

Unintended Effects?

- Linking records
  - Need affirmative consent to link survey
  - Public trust, response rates?
- Replacing data with models?
  - What if relationships change?
  - Convenient data sources manipulable
  - Need high-quality surveys or other independent sources

Unintended Effects?

- More honest standard errors?
  - Standards for sensitivity analyses
  - It’s easy to find model giving no nonresponse bias
  - Incentives penalize thorough analysis
- Errors in data sources correlated
  - Victimizations missing from all
- Privacy?

Abbott (1915)

- Need accurate statistics:
  - Use police resources efficiently
  - Dispel myths about crime
- Vital stats: model for crime stats
- Comparison groups: identify outliers
- Bring together statistics about crime from various sources

Future Research

- Use multiple sources to study quality
- Standard error:
  - Need systems change
  - Include measurement, nonresponse
- “We of the West are not afraid of crossing the frontiers”