Quantify and Punish: Data, Race, and Policing
From the Burgess Method to Big Data

Module 11
Logistics

- For next Thursday: come prepared to give a brief (3/5 minute) presentation on an empowering work of recent social science
- Individually or in pairs (if in pairs, both partners should talk, presentation should be proportionally longer)
- Slides/visuals optional…
- Post links on Canvas board
Today’s Plan

I. Statistics, policing, and racial injustice: an overview

II. Crime statistics and the “condemnation of blackness,” c. 1890-1940

III. Risk assessment and the “actuarial” approach to policing and punishment, from the Burgess Method (1928) to AI

IV. “Data-driven” policing from CompStat (1990s) to Big Data
Quantify and Punish: Overview

- Module 11 considers: what role have quantitative data, computational methods, and social science played in the construction of modern systems of criminal justice?

- How has quantification contributed to the injustices of modern policing and punishment – to the creation and maintenance of a system that disproportionately and unjustly targets, punishes, incarcerates, and kills people of color, especially Black citizens?

- What can history tell us about the role that data and computation should – or should not – play in efforts to create a more just system of justice in the future?
Quantification-Race-Policing: Big Themes

- Turn to data and quantification often driven by desire to create fairer, more accountable, less biased systems – often in response to crisis, calls for reform
- While quantitative turn has brought some positive effects and reforms, it has also brought harms
- Efforts to quantify policing and punishment rely on historical data, which reflect biases of previous system – histories of data shape their futures
- Data is always already mediated through previous criminal justice apparatus – the crime data” is in fact always “policing data”
In fact, data has played a central role in the historical construction of ideas about criminality that undergird structural racism of modern criminal justice system.

Computational tools can obscure, launder, and exacerbate existing patterns of discrimination – behind veil of “objectivity”.

The establishment of quantitative systems for guiding and evaluating policing can reshape policing behaviors in detrimental ways:

- “Ratchet effects”
- “Juking the stats”
“Cartes figurative: Crimes contre les propriétés / Crimes contre les personnes.” Two lithograph maps within one border, 21.5 × 33 cm. From Quetelet’s *Sur l’homme et le développement de ses facultes; ou, Essai de physique sociale*, 2 vols. in 1 (Brussels: Louis Hauman, 1836) [Historic Maps Collection]. (Princeton Historical Maps Collection)
Three Episodes

- Crime statistics and the “condemnation of blackness,” c. 1890-1940
- Risk assessment and the “actuarial” approach to policing and punishment, from the Burgess Method (1928) to AI
- “Data-driven” policing from CompStat (1990s) to Big Data
Crime statistics and the “condemnation of blackness,” c. 1890-1940
In 2011 book, Khalil Gibran Muhammad argues that statistics played a crucial role in the establishing a link between Blackness and criminality.

Notion that criminality was a feature of Black Americans as a group, while crime by white Americans was masked as individual failure.

“Black criminality [became] the most significant and durable signifier of black inferiority in white people’s minds” (p. 3).

Gibran argues this link was forged in the Jim Crow era (1890-1940).
The 1890 Census

- 1890 Census was a watershed moment
- Context: 25 years after emancipation; more than a decade after end of Reconstruction – question of status and future of African Americans in American society
- 1890 census publicized data about prison populations by race: Black Americans were 12% of population and 30% of prisoners
- Statistical racists like Frederick T. Hoffman (Module 7) drew heavily on 1890
Comparing Black and Foreign-Born

- Crucial to the creation of what Muhammad calls the “statistical ghetto” was comparison between Black and foreign-born white Americans.

- Progressive era social scientists interpreted (or dismissed) statistics on crime by European immigrants as evidence that Irish, Italians, Poles, etc. could be assimilated into US culture.

- Charles R. Henderson, U. Chicago sociologist, 1901: “the [evil of immigrant crime] is not so great as statistics carelessly interpreted might prove...”
Comparing Black and Foreign-Born

- Henderson (1901), cont’d: But where the “Negro factor” is concerned, “racial inheritance, physical and mental inferiority, barbarian and slave ancestry and culture…” were among the “most serious factors in crime statistics.”
Risk assessment and the “actuarial” approach to policing and punishment, from the Burgess Method (1928) to AI
“Positive Criminology” & “Individualization”

- Key finding of 19th c. quantitative soc. sci. was regularity of “laws” of crime – Quetelet’s “budget of the scaffold”
- Crime was not random, but predictably steady
- Advocates of “positive criminology” like Cesare Lombroso (1835-1909) and Charles Goring (1870-1919) held criminals were not randomly chosen from population
- Criminality seen to be correlated with home conditions, physical traits, genetic makeup, neighborhood, and race

This image is in the public domain.
"Positive Criminology" & "Individualization"

- Circa 1900: international interest in the "individualization of penal treatment"
- National Conference on Criminal Law and Criminology, 1909: U. Chicago law profs. Ernst Freund and Roscoe Pound set agenda:
  - "Modern science recognizes that penal or remedial treatment cannot possibly be indiscriminate and machine-like, but must be adapted to the causes, and to the man as affected by those causes"
Ernest Burgess and Parole Prediction

- Goal of “Individualization” encouraged move toward indeterminate sentencing for crimes: local parole boards to determine length of sentence

- 1920s: researchers like Hornell Hart (Iowa Child Research Station) argue that statistics might make it possible to create a “prognostic score for each man coming up for parole”

- Robert Burgess (1886-1966): PhD in Sociology from U. Chicago in 1913; faculty 1916

- Co-wrote *Introduction to the Science of Sociology* w. Chicago mentor Robert Park (1921)

© University of Chicago. All rights reserved. This content is excluded from our Creative Commons license. For more information, see [https://ocw.mit.edu/help/faq-fair-use/](https://ocw.mit.edu/help/faq-fair-use/)
Ernest Burgess and Parole Prediction

- Burgess emphasized quant. methods over “ecological” approach dominant at Chicago
- Burgess: “Prediction is the aim of the social sciences as it is of the physical sciences.” (1929)
- 1927-28: Burgess and four colleagues requested by chairman of Illinois parole board to study Ill. procedures
- Understaffed, overworked parole officials in Ill. had been unable to give much attention or deliberation to parole cases
- Burgess conducts study of 3,000 parole cases in Illinois ~1920-1924
Ernest Burgess and Parole Prediction

- Burgess attempted to find a statistical relationship between who did/did not violate terms of parole and 22 other independent factors

- Factors included: racial and ethnic categorization, social & personality type, mental age, circumstances of crime, and prior criminal record

- Burgess develops 21-factor test: those with high scores (16-21) had low parole-offense rates (1.5%); those with low scores (2-4) had highest offending rates (76%)

- **Reading for this module...**
The Burgess Method Takes Hold

▪ Burgess Method implemented quickly at Joliet Penitentiary, 1932-33

▪ 1932 election: Democratic wave leads to Dem. Governor elected in IL; appoints John Landesco, Burgess’s research assistant, to State Parole Board

▪ Landesco urges IL legislature to pass bill to hire sociologists and actuaries to “make analyses and predictions in the cases of all men being considered for parole”

▪ 1930s-40s: Adoption of Burgess Method in IL triggers outpouring of academic research
The Burgess Method Takes Hold

- Others seek to critique, improve Burgess Method through more data and development of new risk-assessment tools using different variables and weighting different variables more/less
- Some explored more sophisticated statistical techniques (e.g. multiple regression) to develop assessment tools
- Researchers studying Burgess Method found prominent positions at top institutions like U. Chicago, Harvard Law School, University of Southern California, and University of Minnesota
For the first few decades, application actuarial parole methods confined to Illinois

Other states begin to adopt similar tools in 1960s

Prior to 1970s, race and nationality were common factors in these tools

Early 1970s: federal gov’t adopts slim, 7-factor “Salient Factor Score”

Federal adoption stimulates widespread interest in parole prediction tools – first in California, then wave of other states in the 1980s-90s


© Bernard Harcourt. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/
The Actuarial Approach Takes Off

![Graph showing historical trend in the number of states using parole-prediction instruments.

**Figure 1.1** Historical trend in the number of states using parole-prediction instruments. Source: Survey conducted by Marylynne Hunt-Dorta in July and August of 2004. Hunt-Dorta surveyed the parole authorities of the fifty states to determine whether they used prediction instruments, which ones they used, and when they began using them. Documentation of the interviews is available upon request from the author.

From Harcourt, *Against Prediction: Policing, Punishing, and Profiling in an Actuarial Age* (Chicago, 2007), p. 9. © Bernard Harcourt. All rights reserved. This content is excluded from our Creative Commons license. For more information, see [https://ocw.mit.edu/help/faq-fair-use/](https://ocw.mit.edu/help/faq-fair-use/)
Example of a criminal sentencing worksheet from the Virginia Criminal Sentencing Commission (Source)
By 1951, Burgess argued actuarial risk assessment could be applied much more widely than parole selection.

Later part of the 20th century saw adoption of risk-assessment methods to policing potential crimes as well ("profiling").

1968: Federal Aviation Administration implements "airline-highjacker" profile – 25 characteristics.

1970s+: increasing use of profiles for specific crimes "drug-courier".
Starting in 1969, IRS began to use computerized assessment of past income tax returns to develop predictive tools (secret “Discriminant Index Function”) to flag income tax return for audit.

Image from hijacking of TWA flight 847, 1985 (CNN)
© JOEL ROBINE/AFP/Getty Images. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/
Social Costs of Prediction: the “Ratchet”

- Legal scholar Bernard Harcourt and others have identified crucial social costs to the actuarial, predictive approach to criminal justice

- Using statistical techniques to target policing/punishment will increase success rate of searches, audits, parole decisions, etc.
  - More searches (of specific group) will find more contraband…
  - More parole denials (of specific group) will prevent re-offenses…
Social Costs of Prediction: the “Ratchet”

- If the actuarial methods disproportionately target a specific group – e.g. drywall contractors for tax audits, young Black men for street frisks – then over time the extra policing attention will produce more infractions from that group.
- This may be taken as confirmatory evidence that the specific group offends at a higher rate (not that the group is under heightened attention).
- This creates a “ratchet”: increased police attention → more evidence of infractions → increased police attention …
Social Costs of Prediction: the “Ratchet”

- subject to police contacts, even with small/no punishment or for minor infractions, can have significant social costs

- Harcourt: “Disproportionate criminal supervision and incarceration reduces work opportunities, breaks down families and communities, and disrupts education.” (161)
Algorithmic Predictions and Their Biases

- Today, “risk assessment” for criminal sentencing and parole are increasingly conducted by algorithms
- E.g. the COMPAS -- Correctional Offender Management Profiling for Alternative Sanctions – algorithm from for-profit company Northpointe
- Algorithmic risk scores widely used in assigning bail, criminal sentencing, and parole decisions
- As of 2016, risk scores given directly to judges during sentencing in at least 11 states
- COMPAS algorithm based on 137 questions/data points; not race
‘The survey asks defendants such things as: “Was one of your parents ever sent to jail or prison?” “How many of your friends/acquaintances are taking drugs illegally?” and “How often did you get in fights while at school?” The questionnaire also asks people to agree or disagree with statements such as “A hungry person has a right to steal” and “If people make me angry or lose my temper, I can be dangerous.”’

Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, “Machine Bias,” ProPublica (May 23, 2016).
2016 study by ProPublica found that COMPAS algorithm was…

...very bad at predicting violent crime: 20% of those predicted to commit future violent crime did

...only moderately good at predicting all crime (including misdemeanor): 61% of those deemed likely recidivists committed future crimes

...racially biased: Black defendants much more likely to be flagged incorrectly as likely re-offenders (false positive), and white defendants more likely to be mislabeled as low risk (false negative)
Algorithmic Predictions and Their Biases

Prediction Fails Differently for Black Defendants

<table>
<thead>
<tr>
<th></th>
<th>WHITE</th>
<th>AFRICAN AMERICAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeled Higher Risk, But Didn’t Re-Offend</td>
<td>23.5%</td>
<td>44.9%</td>
</tr>
<tr>
<td>Labeled Lower Risk, Yet Did Re-Offend</td>
<td>47.7%</td>
<td>28.0%</td>
</tr>
</tbody>
</table>

Overall, Northpointe’s assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)
“Data-driven” policing from CompStat (1990s) to Big Data
Bill Bratton and Data-Driven Policing

- No single individual more responsible for expansion of “data-driven policing” than Bill Bratton
- Across two stints as police chief in NYC and one in LA, Bratton led three major expansions of use of data and computation
- Critically: in all three cases, Bratton was brought in to reform police depts. in crisis
- In each case, Bratton turned to data and computation as a remedy for corruption, malpractice, and bias
Early 1990s, NYC experiencing high crime and raft of police corruption

Mollen Commission (created 1992) revealed widespread, unchecked corruption in NYPD: “characterized by brutality, theft, abuse of authority, and active police criminality.”

1994: Mayor Rudolph Giuliani appoints Bratton new Police Commissioner; previously Boston police chief and head of NYC Transit Police

As head of Transit Police, Bratton had overseen work of Jack Maple
Jack Maple’s “Charts of the Future”

- Jack Maple started in NYPD in 1970; became a transit police officer, inc. undercover Times Square / 42nd Street Station
- In 1970s-80s, subways major sites of crime, esp. violent robberies
- As detective, Maple took to mapping patterns of subway crime with pushpins on wall maps – “Chart of the Future”
- Charts allowed Maple to observe repeated patterns in subways crimes – large %age of crimes from serial offenders tracing set routes
- Targeting policing to key locations disrupted these patterns; led to 27% reduction in subway crime

Jack Maple (Source)

© unknown. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/
When appointed Commissioner in 1994, Bratton promoted Maple from lieutenant to Deputy Chief.

Maple and Bratton institute new COMPuterized STATistics program.

CompStat required personnel from each of city’s 77 precincts and other units to submit weekly report on complaints, arrests, summons, open crimes, etc.

NYPD CompStat system produced regular reports showing weekly, monthly, annual trends.


© Harvard. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/
CompStat in Practice

- Systems of regular meetings with leading NYPD officials and precinct/unit commanders enabled sharing of data, discussion of crime control strategies, and oversight of units.

- One key objective of CompStat including improving accountability and professionalism of local force while giving more responsibility to local commanders.

- Also sought to address problems of inadequate police attention in certain areas: e.g. major crimes in poor, predominantly Black and Latinx areas outside Manhattan core went routinely unsolved.


© Harvard. All rights reserved. This content is excluded from our Creative Commons license. For more information, see [https://ocw.mit.edu/help/faq-fair-use/](https://ocw.mit.edu/help/faq-fair-use/)
The “Marvel” of CompStat

- Institution of CompStat coincided with notable decline in crime in New York City – inc. stark decline in murders, from record 2,245 in 1990 to 673 in 2000 (and 289 in 2018)
- Bratton publicly targeted 10%+ reductions in crime; achieved 12% in first CompStat years
- NYT called results under Bratton “marvel of modern law enforcement”; “simply breathtaking”
- But debate over role of CompStat in decline in crime (decline began in 1990, before CompStat, and occurred across major cities)
The “Marvel” of CompStat

- Apparent successes of CompStat in New York led to widespread uptake of data-driven methods across other US police departments over early 2000s, inc. Philadelphia, Miami, Chicago, Baltimore, DC, San Francisco, etc.

- 2002: Bratton recruited to lead Los Angeles PD; wracked by longstanding abuses, poor community relations, poor morale, and major corruption scandal in anti-gang unit (“Rampart Scandal”)

- Bratton instituted CompStat in LA
The “Marvel” of CompStat

- For opposing perspectives on the role of CompStat in crime decline:

Lauren-Brooke Eisen, Oliver Roeder, and Julia Bowling, “What Caused the Crime Decline?” The Brennan Center for Justice (Feb. 12, 2015)

“Juking the Stats”: Critiques of CompStat

- Introduction of CompStat had major effects on practice and culture of policing
- Increasing political, managerial emphasis on reducing CompStat numbers created pressure on commanders and officers to make it appear crime was falling
- Emphasis on measurable stats distracted from public safety goals, community relationships
- “Juking the Stats”: statistical manipulation designed to make CompStat figures appear better: e.g. through reclassifying crimes into lesser categories (aggravated assaults → assaults)


© James Estrin/The New York Times. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/
“Juking the Stats”: Critiques of CompStat

- Rewarding measurable indicators of police activity – e.g. arrests and tickets – incentivized officers to increase punishment for minor offenses (e.g. transit fare evasion, vandalism, loitering)

- Especially combined with “broken windows” theory (as in Bratton in the 90s): theory that rigorously policing minor “anti-social” crimes creates environment of order, lawfulness that prevents more serious crime

- Black and Latino citizens disproportionately targets of minor-crime enforcement: 2001-2013, Blacks and Latinos in NY (57% of population) were 80% of misdemeanor arrests and summonses

© James Estrin/The New York Times. All rights reserved. This content is excluded from our Creative Commons license. For more information, see https://ocw.mit.edu/help/faq-fair-use/
"Juking the Stats"

The Rise of Big Data Policing in the 2000s

- Bratton returned as NYPD chief Jan. 1, 2014 – in wake of widespread anger about “stop-and-frisk” and legal ruling declaring it unconstitutional
- Bratton oversees construction of real-time crime-command center in Manhattan; orders tens of thousands of crime mapping tablet computers
- “Intelligence-led policing” purported to go beyond “hunches,” offer alternative to stop and frisk
Data & the Who/Where/When/How of Policing

- **Whom…**
  - In 2012, Kansas City PD used social-network analysis to identify 884 people deemed likely to commit homicides “focused deterrence”
  - Chicago developed algorithmic “heat list” to identify likely perpetrators and victims of gun violence

- **Where…**
  - Jeff Brantingham (UCLA) uses earthquake-prediction techniques to develop PredPol algorithm for identifying likely sites of crime implemented by LAPD in 2011 (25% reduction in burglaries)
Data & the Who/Where/When/How of Policing

▪ **When…**
  ▪ Networks of surveillance devices (cameras, license plate readers, etc.) used for real-time tracking and alerts; e.g. NYPD-Microsoft “Domain Awareness system” monitoring southern Manhattan

▪ **How…**
  ▪ Data mining to search “cellular, digital, and biological data trails”
  ▪ Facial recognition

▪ Notably, much of this done in concert with private for-profit companies – as in Palantir and LAPD Real-Time Analysis Critical Response section
Andrew Guthrie Ferguson proposes five-point checklist for safe application of big data techniques:

1. Can you identify the risks your technology is trying to address?
2. Can you defend the inputs into the system?
3. Can you defend the outputs (how they will impact policing practice/community)?
4. Can you test the technology (transparency/accountability)?
5. Is police use of the technology respectful of the autonomy of the people it will impact?

- Or, abolish big data? (Data for Black Lives)
MIT OpenCourseWare
https://ocw.mit.edu

RES.TLL-008 Social and Ethical Responsibilities of Computing (SERC)
Fall 2021

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms