Efficiency and Equity in Government

Service allocation & Congestion Pricing

Nikhil Garg, Cornell Tech (ngarg@cornell.edu)

Special thanks to Urban Tech Hub at Cornell Tech & NYC Department of Parks and Recreation

High level vision

Opportunity

Large, granular, open data

Modern computational, machine learning, algo fairness literature

Willing, equity-focused practitioners, important decisions to make

Challenge

Reality + data is complicated: censoring, distribution shifts, spatio-temporal heterogeneity, strategic behavior

Status quo ⇔ Estimation ⇔Decision-making

Status quo disparities, both spatially and demographically

Estimation

- What data do you need to make decisions?
- Is this data missing at random? Does it vary spatially/demographically?

Decision-making

- What sorts of interventions are feasible?
- Can you “price discriminate” or make “online” decisions?

Need to center institutional context, work with domain experts, and develop new methods informed by domain characteristics

Today

Today:

- Understanding and improving government service allocation
- Near the end: Designing equitable congestion pricing

Government service allocation

Local government manages many services

- ~8k miles of streets in NYC
- ~700k trees lining streets in NYC
- Housing, sanitation, transportation, etc.

Operational tasks

- [Learning] What problems are there?
- [Allocation] Which ones to address?
- [Auditing] Did we do a good job?

Desiderata: Efficiency & Equity

Street trees on Upper East Side in NYC

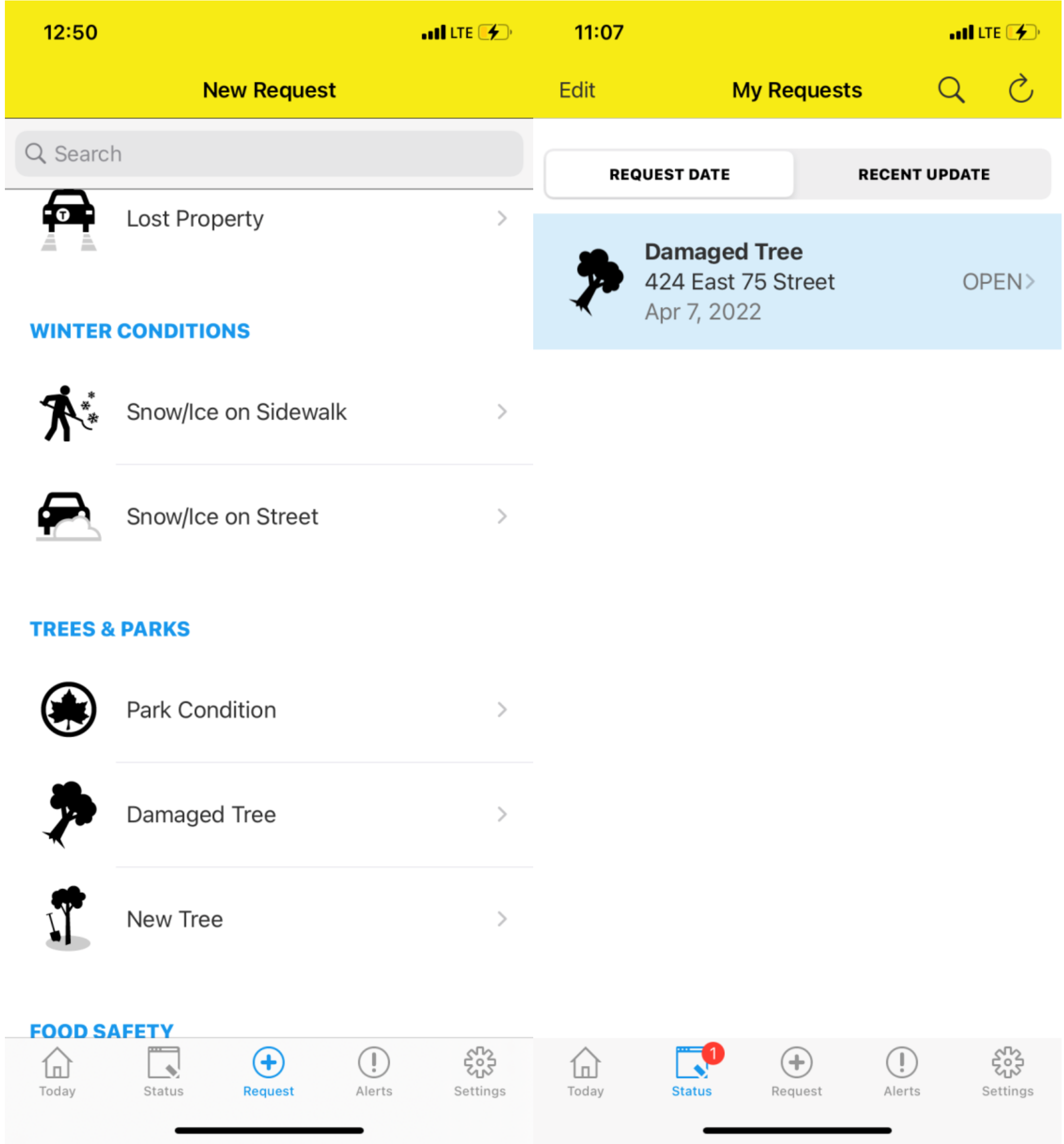
311 (crowdsourcing) systems

Cities have a phone number & app to complain to the local government

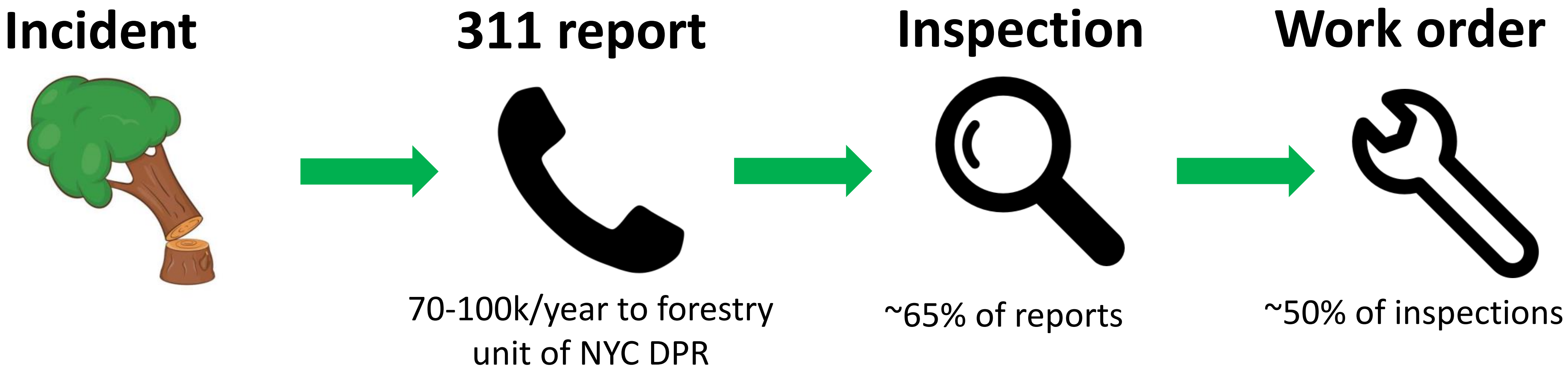
NYC's 311 system receives about 3 million service requests per year

These are the primary way the government learns about problems

- Urgent problems: street floods
- "Routine" problems: falling trees, potholes



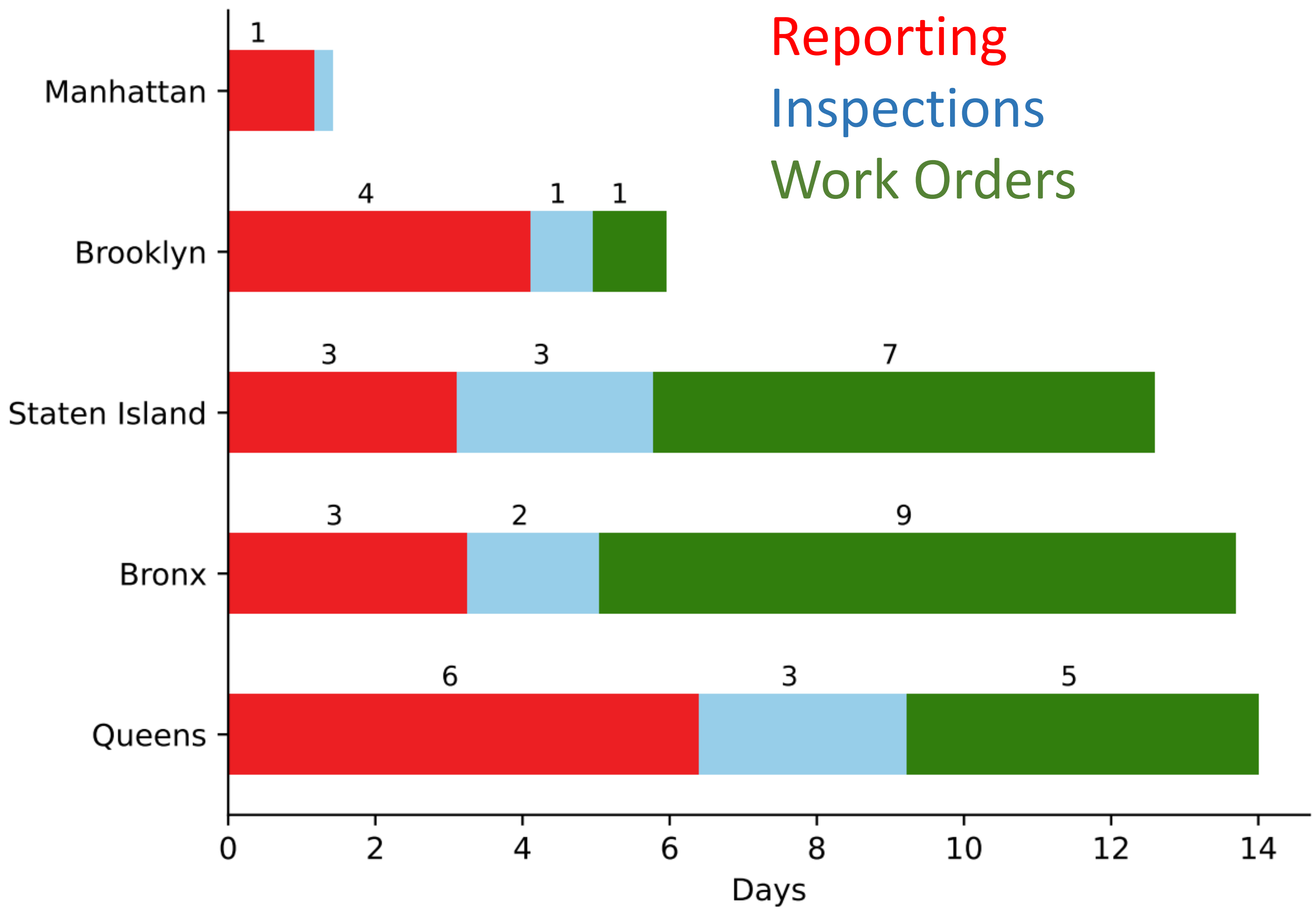
Pipeline: from incident to work orders



Why is this hard? Uncertainty, heterogeneous + strategic behavior, distribution shifts over time, capacity constraints, pipelined decisions

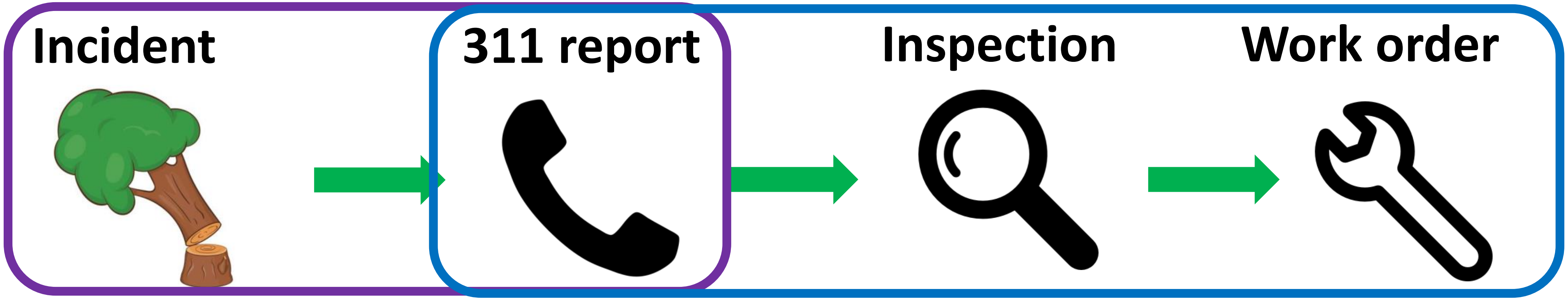
Research agenda: Understand and improve efficiency and equity

Disparities in every stage of the pipeline



End-to-end delays for the highest priority incidents

Research agenda + talk overview



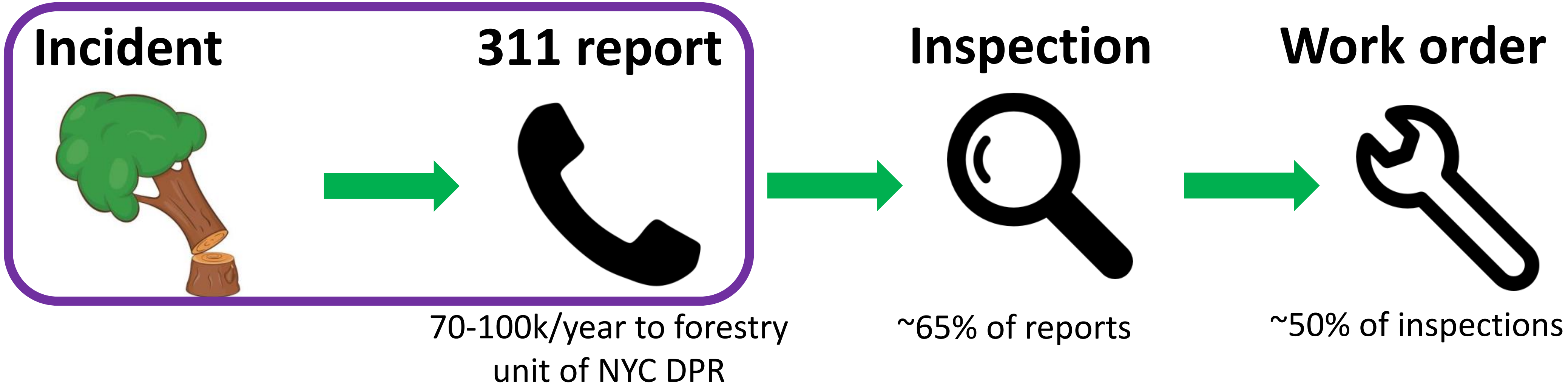
Estimation: Understanding (Heterogeneous) Reporting Behavior

Overcoming missing data challenges using cross-report data

Decision-making: Understanding (and making) optimal decisions

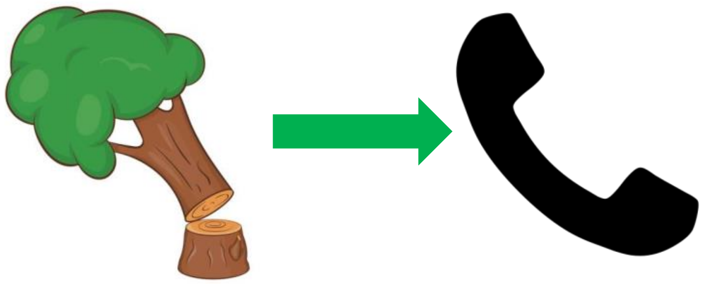
with heterogeneously missing data & needs

Understanding reporting behavior



Why? If there are disparities in who reports problems, there will be disparities in what work gets done. Need to understand them to mitigate (or at least not reinforce) them with decision-making.

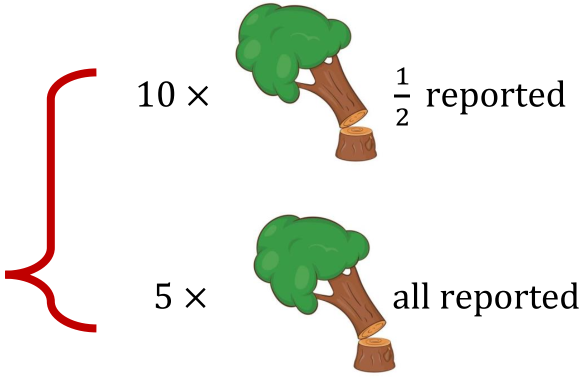
Statistical challenge



How do we distinguish between under-reporting, and some neighborhoods truly having fewer problems?

By definition, we don't observe data on missing reports

If a tree falls in a forest, and no one reports it...
(how) does the city know about it?



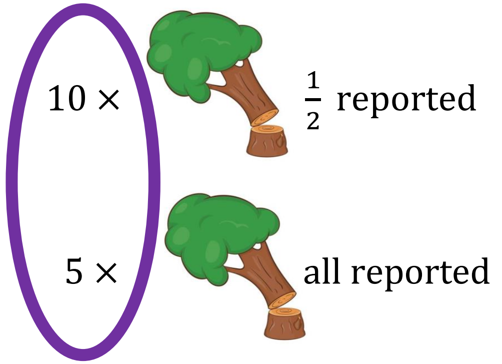
This "Benchmark" problem is a fundamental challenge across contexts

Policing: crimes committed vs inequitable policing

Healthcare: under-testing vs better health

Ecology: recording effort vs species population

How to measuring under-reporting?



“Standard” approach: Use ground-truth data on incident rate: “how many incidents of each type (hazards, root issues, tree pruning requests…) do we expect to see in each neighborhood?”

- Go out and walk the streets and get a snapshot, uncensored view
- Construct proxy measures (number of trees, their size, species, etc)

Our questions:

- Can we measure under-reporting, *without ground truth?*

  (or limited access to ground truth)
- Can we recover the ground truth events?

Key idea 1: “Missing Species”

Leverage the rate of duplicate reports about the same incident to identify the reporting rate, *given* that an incident has occurred

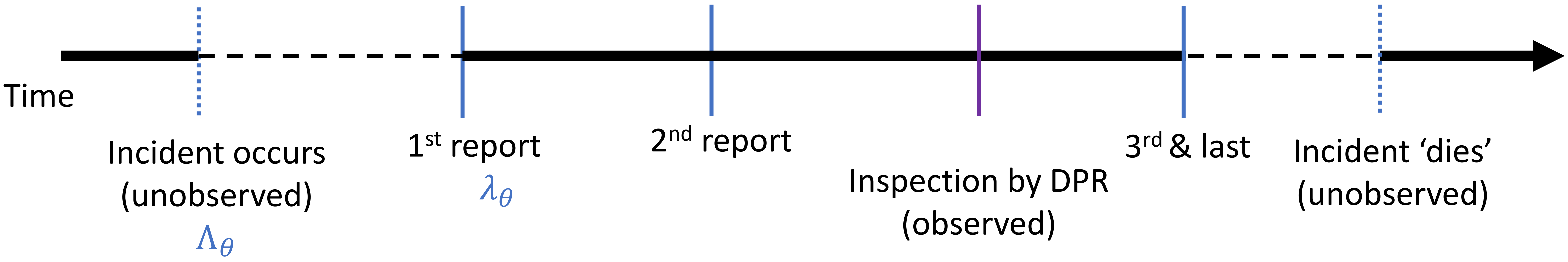
Complication: time. Incidents happen and are fixed;

& we care about reporting *delays*

This work: We develop the statistical method, and then apply it to audit reporting behavior of street tree incidents over 3 years

“Quantifying Spatial Under-reporting Disparities in Resident Crowdsourcing”
w/ **Zhi Liu** and **Uma Bhandaram**
(*Nature Computational Science* 2023; preliminary version in *ACM EC* 2022)

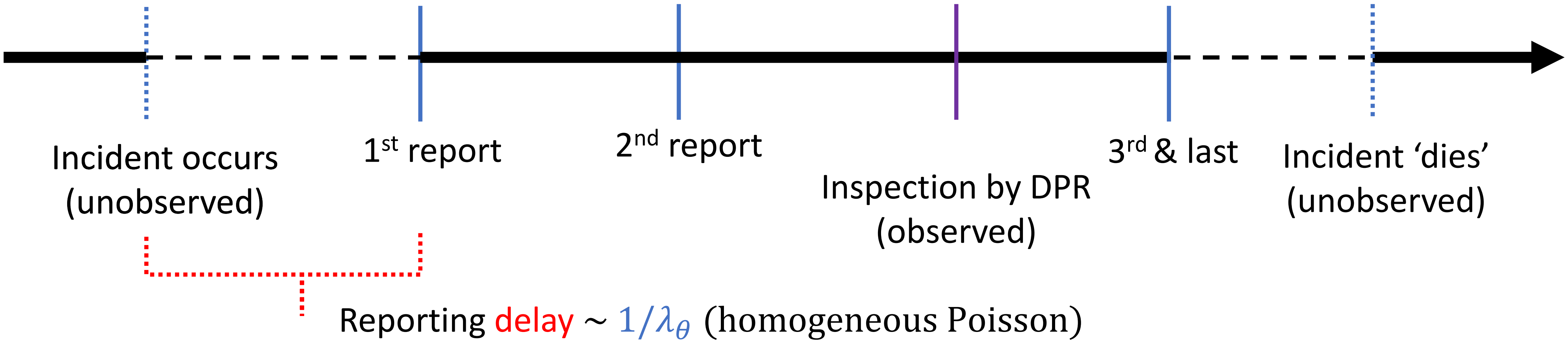
Data generating process



- Incident type θ encodes geography, incident characteristics
- Incident i ***occurs*** according to process Λ_θ at time t_i
- People ***report*** the incident according to Poisson process with rate λ_θ
- Might be *inspected* by DPR (agency)
- Incident "dies" (fixed by agency or someone else) at time $t_i + T_i$

Research question

How does reporting Poisson process λ_θ depend on type θ?



Why does it matter?

- Reporting delay → delay in inspections & fix
- Some incidents may never be reported if low enough reporting rates

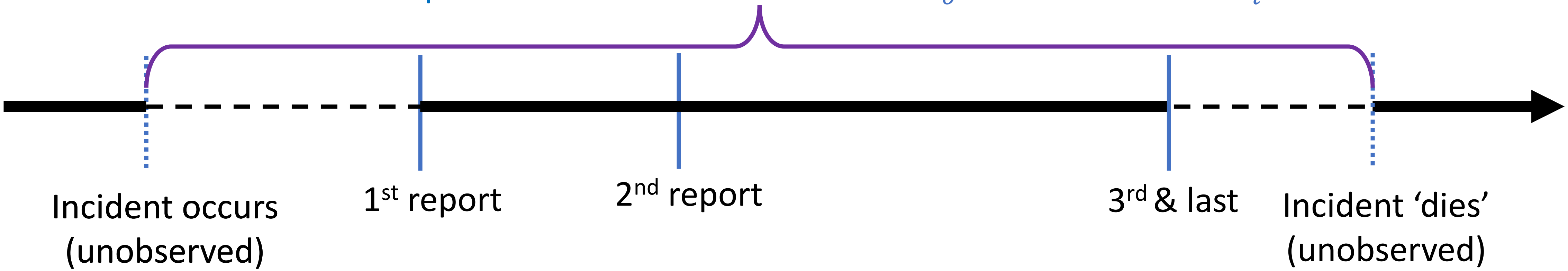
Ideally: more dangerous incidents are reported more quickly

Concern: some neighborhoods use 311 system less → lower reporting rate

Insight: "Missing species"

Duplicate reports about observed incidents give us the reporting rate

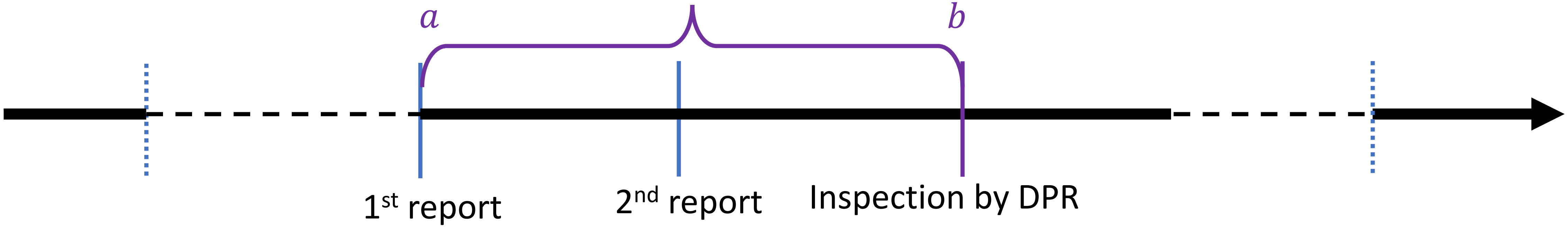
Number of Reports distributed Poisson at rate λ_θ for time duration T_i



If know reporting duration T_i, then reduced to Poisson (potentially non-homogeneous) rate estimation problem: $\lambda_\theta \approx \dfrac{\#\text{reports}}{T_i}$

But we don't know T_i

Method summary



- For incident i of type θ, construct (a_i, b_i) and count $\#\ \text{reports}$ *inside*
- Then, $\#\ \text{reports(i)}$ is Poisson inside the interval.
- For example, if assume **Homogeneous Poisson process**:

$$\#\ \text{reports(i)} \sim \text{Poisson}\big(\lambda_\theta \times (b_i - a_i)\big)$$

- But same method works for time inhomogeneous estimation
- **Zero inflated Poisson**:

$$\#\ \text{reports(i)} \sim \text{Bernoulli}(\alpha) * \text{Poisson}\big(\lambda_\theta \times (b_i - a_i)\big)$$

High dimensional: Bayesian Regression in Stan

Homogeneous process:

$$\#\ \text{reports(i)} \sim \text{Poisson}\big(\lambda_\theta \times (b_i - a_i)\big)$$

Zero inflated Poisson process:

$$\#\ \text{reports(i)} \sim \text{Bernoulli}(\alpha) * \text{Poisson}\big(\lambda_\theta \times (b_i - a_i)\big)$$

Where $$\log \lambda_\theta = \alpha_0 + \sum_k \alpha_k\, \theta_k$$

Spatial smoothing: ICAR Model [Morris et al. 2019]

- Type θ contains an indicator for census tract (2000+ in NYC)
- Then, α_k for each tract is drawn with mean of α_j of neighboring tracts

Applying the method to understand NYC 311 reporting behavior

Data: NYC street trees

- Mostly public* data, augmented with internal data on inspections, work orders, and anonymized caller information
- ~220k service requests over a 3-year period (June 2017 – June 2020)
  - Of these, ~140k correspond to service requests that were inspected
  - ~100k unique incidents inspected

  (after data cleaning + exploration data exclusion, we analyze ~80k incidents)
- Incident covariates:
  - Location [latitude longitude → census tract]
  - Reported characteristics [category (e.g., hazard vs sidewalk damage), …]
  - Inspector report [Risk rating, tree condition, …]

[*Forestry Service Requests | NYC Open Data (cityofnewyork.us)]

$$\log \lambda_\theta = \alpha_0 + \sum_k \alpha_k \theta_k$$

Results: Efficiency

Reporting rates higher for more urgent incidents

Covariate	Coefficient Mean	Standard Deviation
Category[Hazard]	1.500	0.0170
Category[Prune]	-0.076	0.0280
Category[Root/Sewer/Sidewalk	-1.600	0.0380
INSPCondtion[T.Excellent Good]	-0.300	0.0270
INSP RiskAssessment	0.240	0.0120

Higher reporting rate:

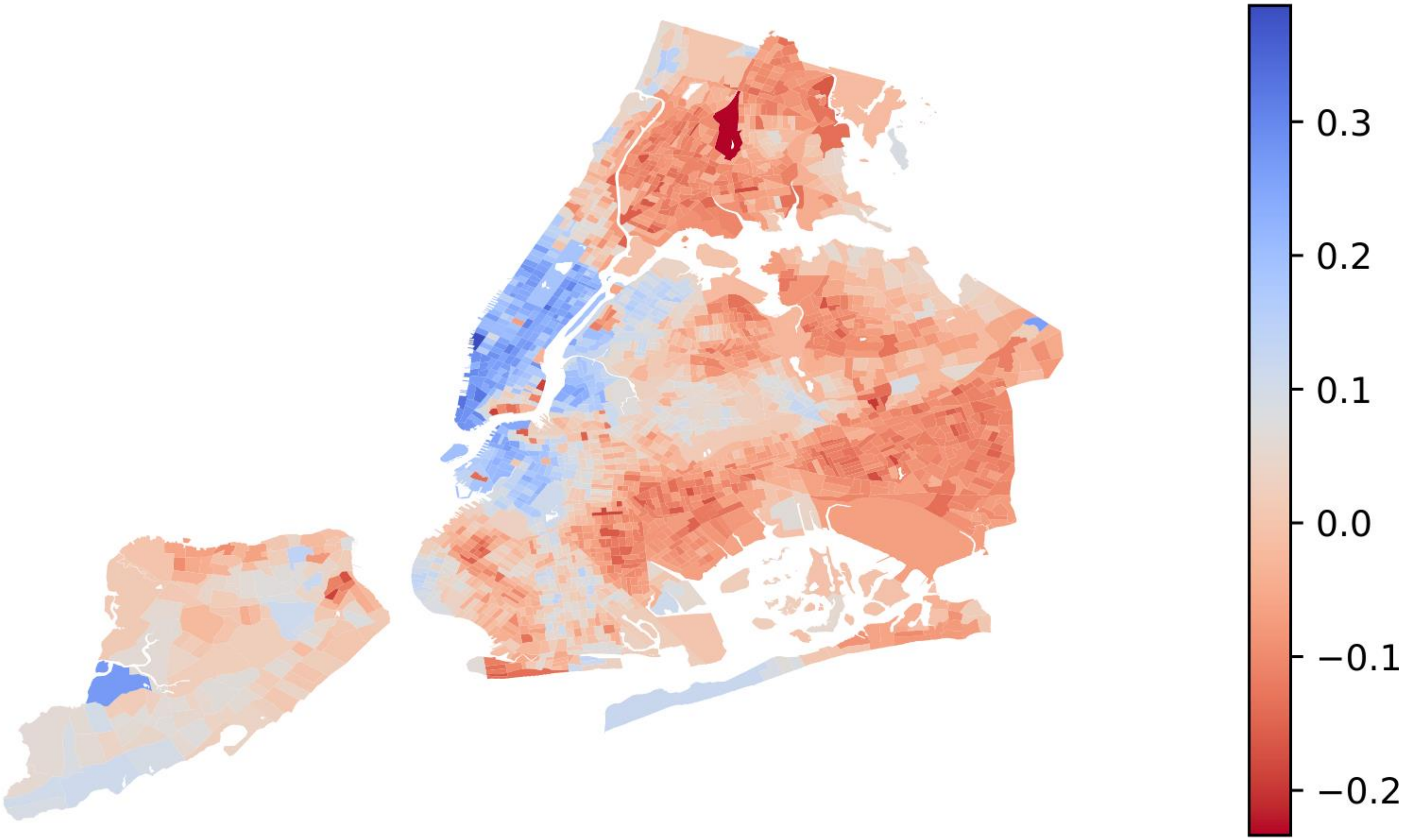
- Hazards (3-7x higher)
- Higher risk incidents
- Trees in Poor/Critical condition

Results: Equity

Reporting rates also vary substantially by neighborhood, even *conditional on incident characteristics*

Difference in reporting rate can be more than 3x between census tracts

Vary by socio-economic characteristics of neighborhoods



Contextualizing the reporting rates

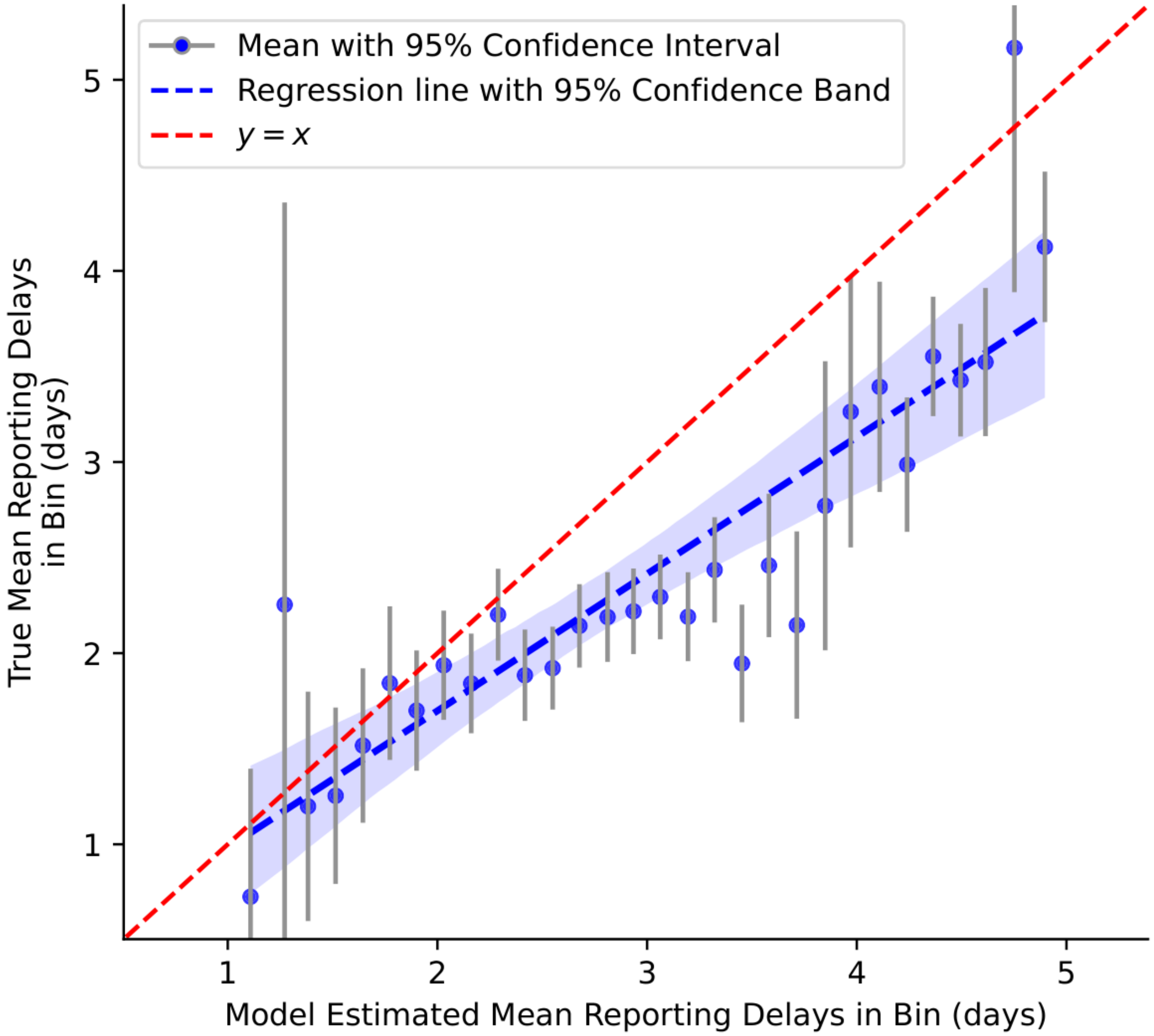
Implied average reporting delay

	Manhattan	Queens
Hazard, tree in Poor condition, High risk	2.5 days	4.7 days

Is the method actually correct?

Idea: when a storm hits, we *know* the timing of the incident (the storm causes many incidents!)

Measure: How long did it take to first report the incident after the storm, versus how long did our model say it would take?



"Quantifying Spatial Under-reporting Disparities in Resident Crowdsourcing"
w/ **Zhi Liu** and **Uma Bhandaram**

Key idea 2: spatial correlation

Question: Can we recover information about *unseen events?*

Idea: Use *spatial correlation* to identify under-reporting!

Model:

- Whether a flood *has occurred* is correlated with neighbors
- Whether a flood *is reported* depends on socio-economic factors

=> Recover ground truth (probabilistically) using reporting data alone

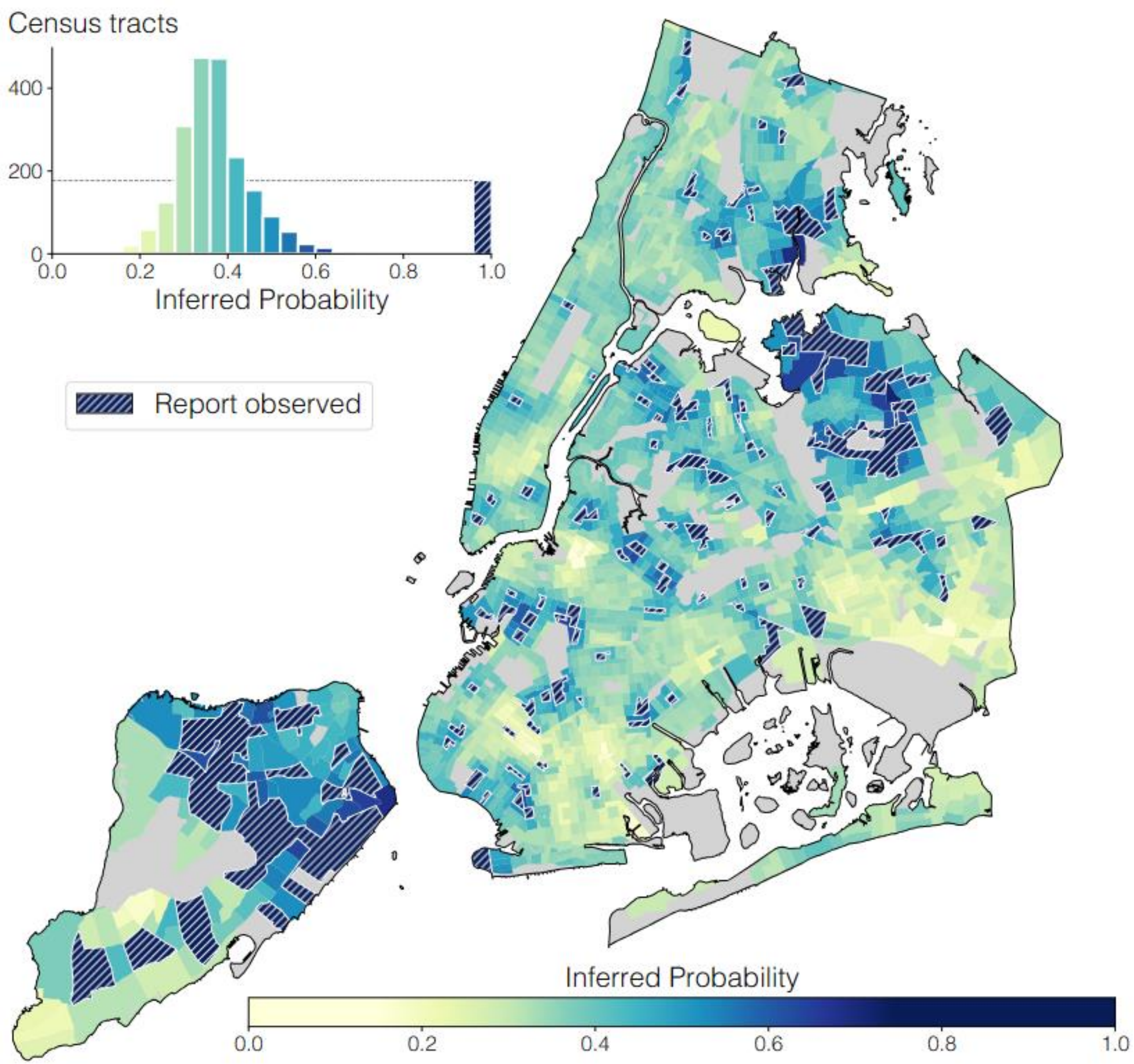
"A Bayesian Spatial Model to Correct Under-Reporting in Urban Crowdsourcing"
w/ **Gabriel Agostini** and **Emma Pierson**
AAAI 2024 (Oral presentation)

Results

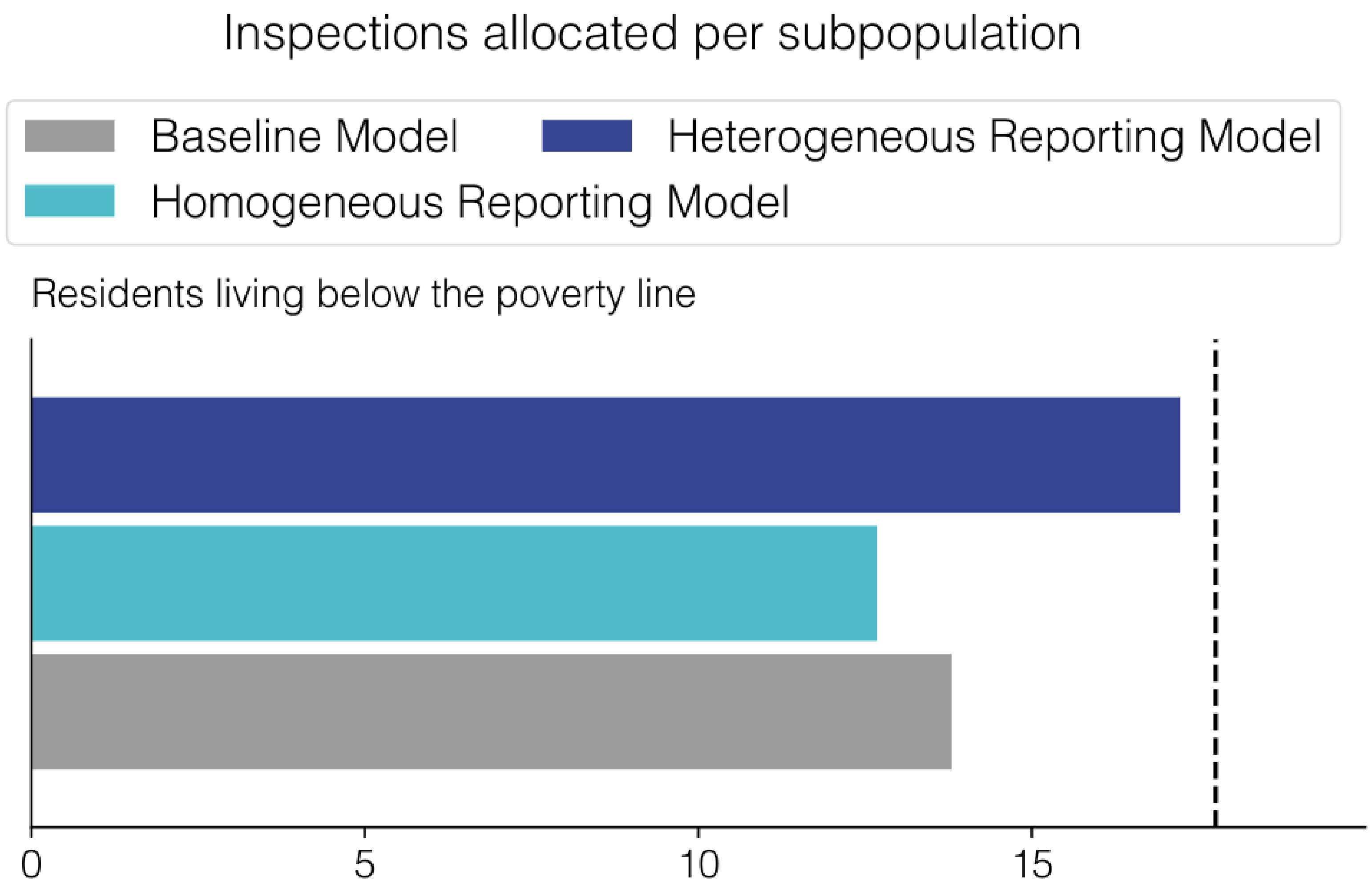
Better predict *future reports* than other methods

Uncovering socio-economic disparities in reporting rates:

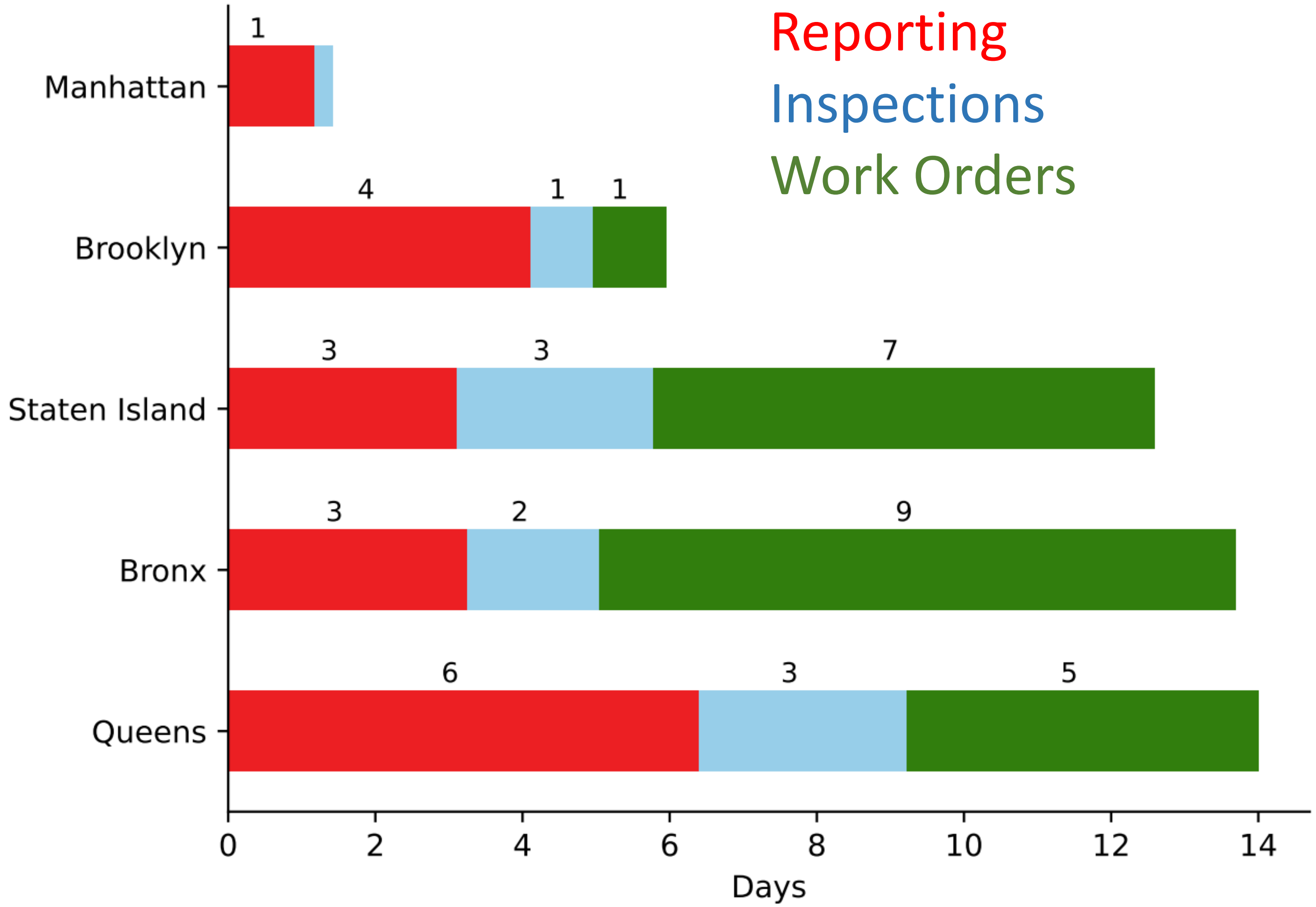
- higher populations, more white residents, owner-occupied households



Results: equitable allocation of Inspections



Is heterogeneous reporting important?
If so, what to do about them?



End-to-end delays for the highest priority incidents

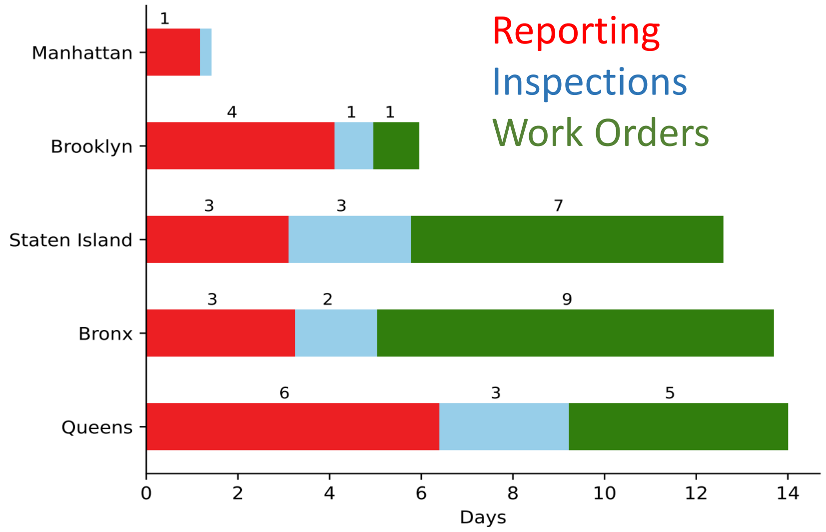
Ongoing work: understanding + making decisions

- Response workers (inspectors, maintenance) are spatially distributed
- Policy + individual worker decisions:
  - # of workers in each location (Borough, neighborhood, County, etc)
  - Which incidents do they prioritize (highest by risk, risk + age…)

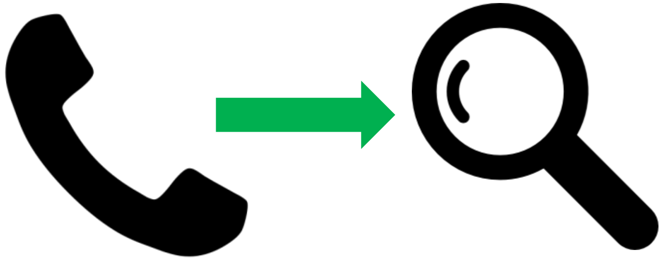
=> response delays for each incident and location

How do these decisions induce further delays?

How do we make “equitable” policies?



Understanding decision-making



Each day, agency officials are solving a subset selection problem

We observe

- Decision inputs (available open incidents, report characteristics)
- Decisions (which incidents were inspected, when)
- Outcomes (inspection results)

Question: did the agency make efficient/equitable decisions?

- Prioritize the riskiest incidents?
- Distribute the risk “fairly”?

Key idea: Cast it as a choice modeling estimation problem

“Detecting Disparities in Capacity-Constrained Service Allocations”
w/ **Benjamin Laufer** and **Emma Pierson**

Making decisions: designing SLAs

We optimize departmental policy:

- # of workers in each location (Borough, neighborhood, County, etc)
- Which incidents do they prioritize (highest by risk, risk + age…)

Desiderata: Efficiency and equity

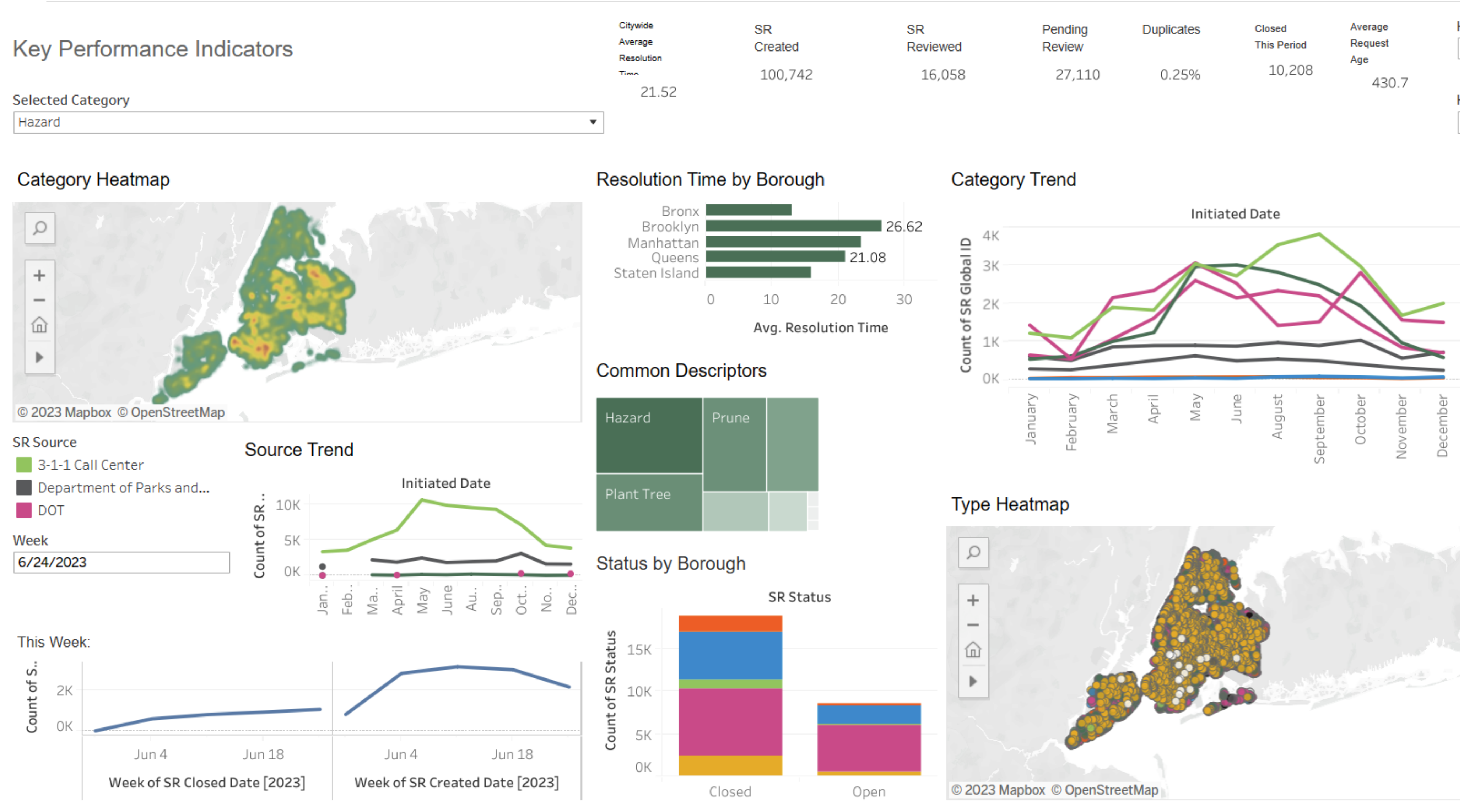
[Potential] worry: Efficiency-equity trade-off

We find: in practice, small trade-off

especially compared to status quo suboptimality

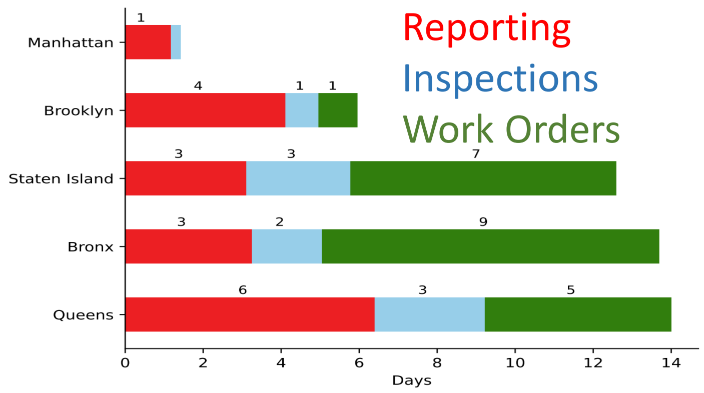
“Redesigning Service Level Agreements: Equity and Efficiency in City Government Operations”
w/ **Zhi Liu**

Translating insights to practice



[parks_dashboard | Tableau Public]

w/ **Anum Ahmad** and **Zhi Liu**



Bringing things together

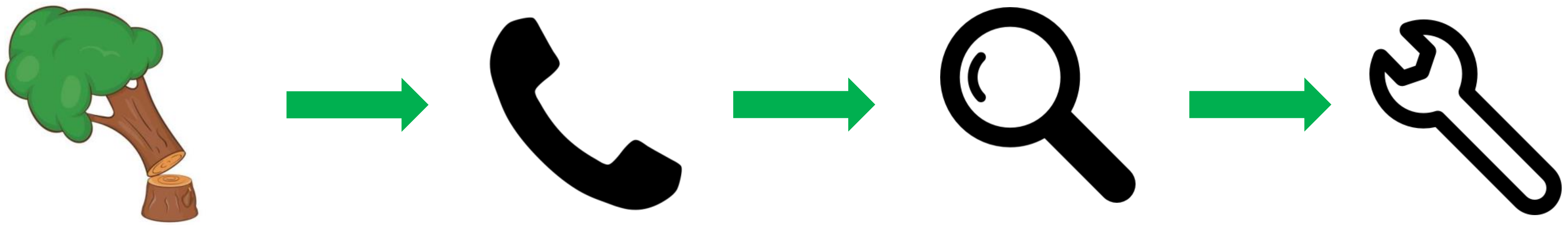
- Inefficiency and inequity can arise from many parts of the pipeline
- Some of it (here: heterogeneous reporting) is hard to fix

Within the policies within our control, how do we:

Mitigate inequity in other parts of the pipeline?

…Or at least not reinforce them?

Incident **311 report** **Inspection** **Work order**



Other applications

Improving public libraries (with NYPL)

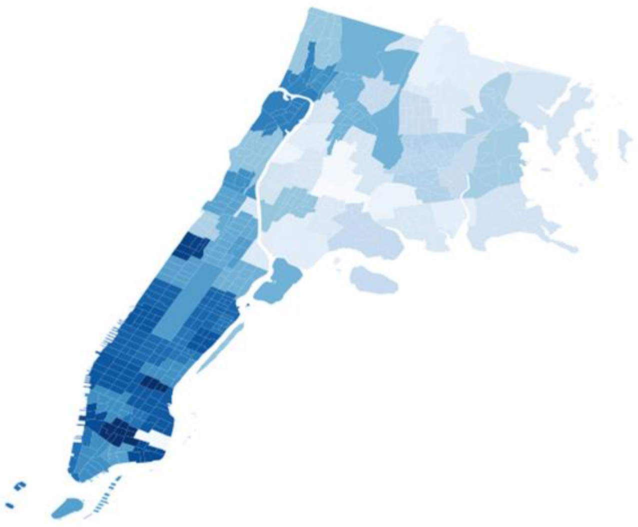
Books are distributed across branches

…but, people can request books (through online holds) from other branches

…who does so is a function of socioeconomics

=> "good" books flow to branches in

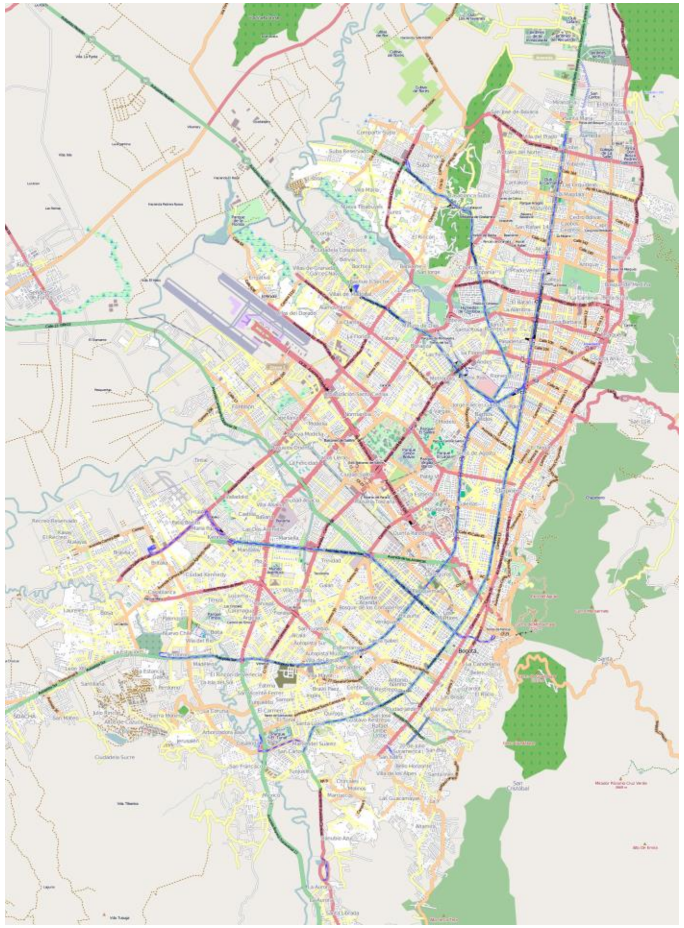
richer areas



Computational intervention: changing where we pull books *from* in order to fulfill a hold request

Identifying and Addressing Disparities in Public Libraries
with Bayesian Latent Variable Modeling
w/ **Zhi Liu** & **Sarah Rankin, AAAI 2024**

Congestion pricing



Congestion pricing heterogeneously affects people by:

- Ability to pay
- Where they live

Pricing has complex equilibrium effects

How do we set (spatial or personalized) congestion pricing to be “equitable” for different groups

w/ Alfredo Torrico, Natthawut Boonsiriphatthanajaroen, Hugo Mainguy, & Andrea Lodi

Discussion

Model and explicitly address data challenges and institutional details

- Missing or censored data
- Capacity constraints
- Institutional knowledge or external data

Decision-making to mitigate (or not reinforce) existing disparities

Similar participation challenges when you study details of any system

New York Public Library

NYC Parks

Questions?

ngarg@cornell.edu