FAIRNESS IN INSURANCE PRICING

Mary Hardy, Carlos Araiza Iturria, and Paul Marriott

Outline

- What do we mean by discrimination?
- What do we mean by “fair”?
- A causal inference perspective
- A Case Study from Wisconsin, USA
- A microsimulation
- Summary and References

Discrimination

- Risk classification vs discrimination
  - discriminatory = protected/sensitive/unfair
- Potential rating factors
  - ***X*** denotes non-discriminatory covariates
  - ***D*** denotes discriminatory covariates
- Y denotes the loss random variable
- $\hat{Y}$ denotes a pure premium estimate

Direct & indirect discrimination

- Direct discrimination
  - when ***D*** is used as a rating factor
- Indirect discrimination
  - ***X*** variables are correlated with ***D***
  - or pricing has a ***disparate impact*** on minority/equity-seeking groups

If *D* and *Y* are correlated, why not discriminate?

1. *D* is not the real issue
   - e.g. race & immigrant mortality
2. Social solidarity
   - e.g. EU gender neutral pricing laws
3. Mitigate systemic inequality
   - e.g. anti-redlining initiatives
4. Differential inaccuracy
   - e.g. lack of credible data

FAIRNESS

What is fair?

- T and J have identical careers – employers, salaries etc
- T and J both retire at 65 with a DC pension fund of $1 million.
- T and J both use the fund to purchase a pension
  - Tom gets 6,500 per month for life
  - Jane gets 5,000 per month for life
- Is that fair?

Group and individual fairness

- Are some groups excluded from coverage?
- Are some groups subsidized by others?
- How granular is the premium rating process?
- Insurance as economic commodity vs social good
- See Frees and Huang (2021) and Xin and Huang (2023)

Fair Premium Criteria

(Xin & Huang, 2023)

- Fairness through unawareness
- Counterfactual fairness
  - The premium would be the same if *D* were different
- Discrimination-free premium (Lindholm et al, 2022)

$$P(\boldsymbol{X}) = \sum_{d \in D} E\left[Y \mid \boldsymbol{X}, d\right] \Pr{}^*[d]$$

Fair Premium Criteria

- Demographic Parity

$$F_P\left(\hat{Y}\middle|D=a\right)=F_P\left(\hat{Y}\middle|D=b\right)$$

- Conditional Demographic Parity

$$F_P\left(\hat{Y}\middle|X_a,D=a\right)=F_P\left(\hat{Y}\middle|X_a,D=b\right)$$

where $X_a \subseteq \boldsymbol{X}$ are designated allowable covariates

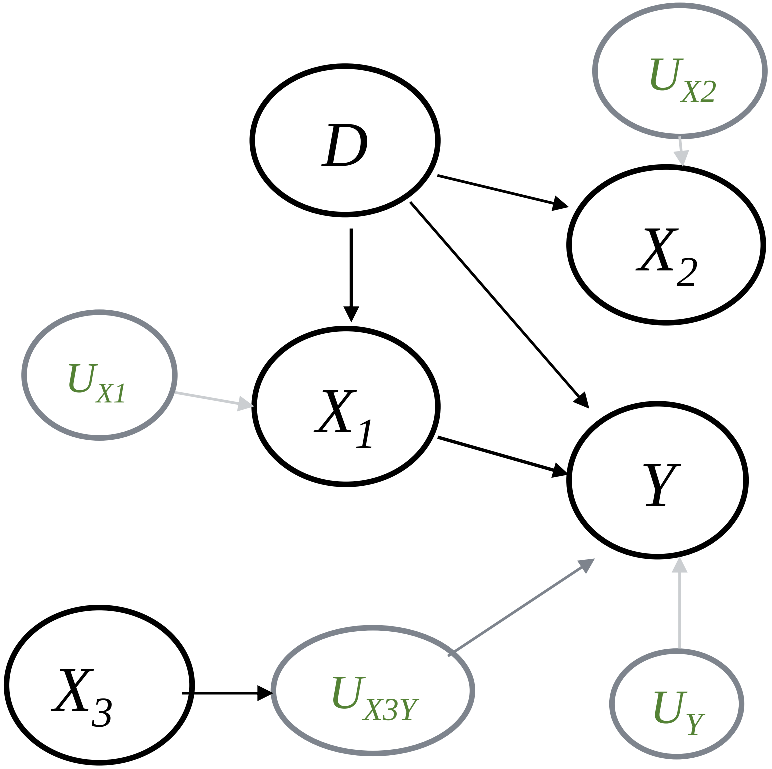
- Weak Demographic Parity

$$\mathrm{E}\left[\hat{Y}\middle|D=a\right]=\mathrm{E}\left[\hat{Y}\middle|D=b\right]$$

CAUSAL INFERENCE PERSPECTIVE

Causal Framework

- Directed Acyclic Graphs (DAGs) model *causality*, eg



D, X_1, X_2, X_3 : Rating factors

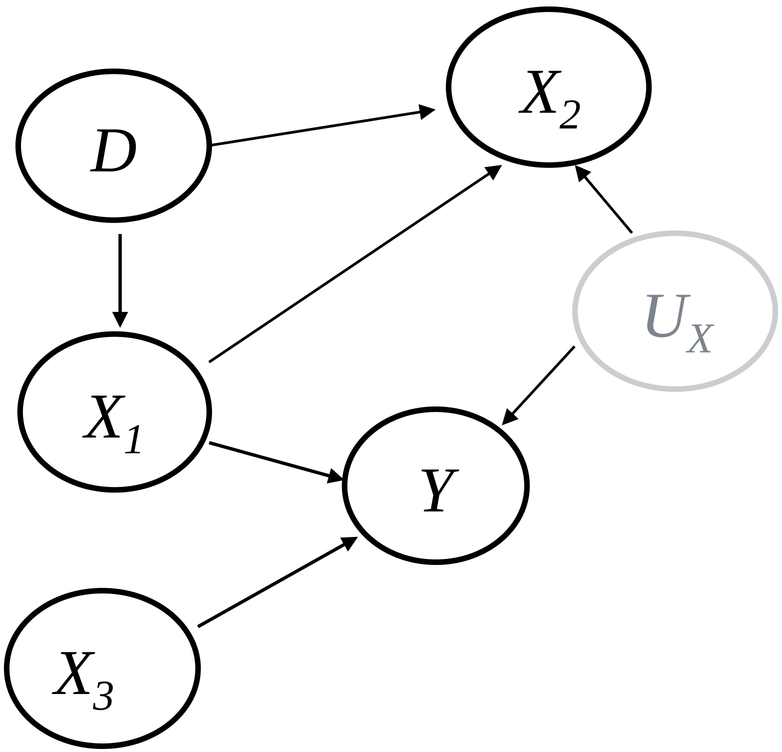
Y : Insurance Loss

U_X : Latent/Unobserved factors

Insurance DAG example

- D = Race $\in$ {a, b}
- X_1 = Zip Code
- X_2 = Credit Rating
- X_3 = age group
- U_X = risk aversion

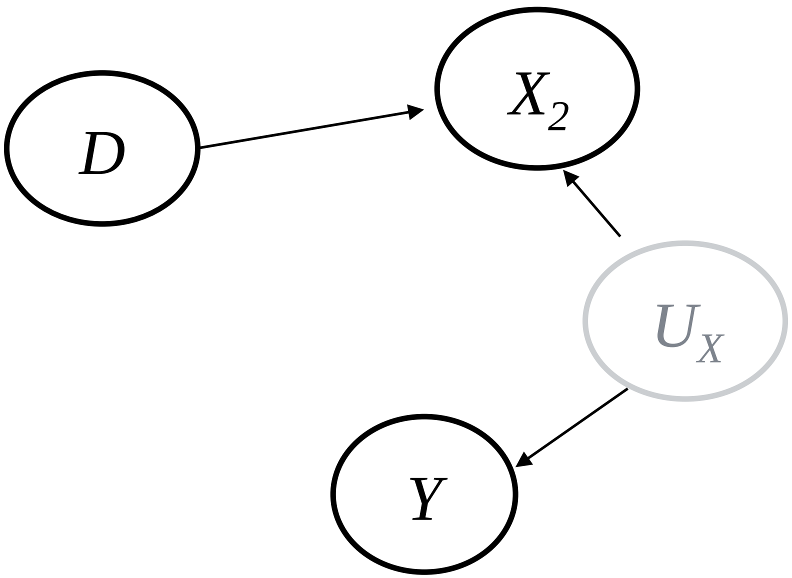
Suppose the impact of X_2 on Y depends on D….



Insurance DAG example

- $D = a$ and $X_2 < 500$
  $\Rightarrow$ low risk aversion
  $\Rightarrow$ higher premium

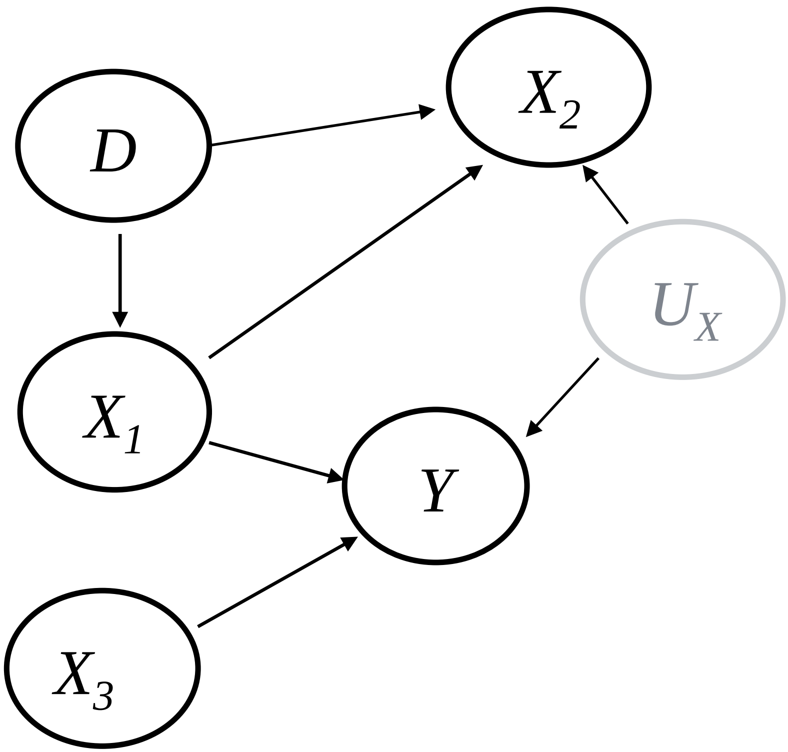
- $D = b$ and $X_2 < 500$
  $\Rightarrow$ no information on risk aversion
  $\Rightarrow$ no impact on premium



Fair premiums & causal model

Unawareness Premium (UP) – ignore *D*

- But X_1 is a proxy for *D*
  - → disparate impact?
- *D* provides information about how X_2 impacts *Y*
- ***Unawareness ≠ Fairness***



Fair premiums & causal model

- Counterfactual Fairness
  - Requires

$$\hat{Y} \,\Big|\, \{D = a, \boldsymbol{X} = (x_1, x_2, x_3)\}$$

$$= \hat{Y} \,\Big|\, \{D = b, \boldsymbol{X} = (x_1, x_2, x_3)\}$$

  - If X_2 is a rating factor, we do not have counterfactual fairness

Fair premiums & causal model

- Discrimination-Free Premium (DFP)

$$P(x_1, x_2, x_3) = \hat{Y}\Big|\{D=a, X=(x_1,x_2,x_3)\}\,\mathrm{Pr}^*(D=a) + \hat{Y}\Big|\{D=b, X=(x_1,x_2,x_3)\}\,\mathrm{Pr}^*(D=b)$$

- If $X_2 < 500$
  - DFP overcharges for $D = b$
  - and undercharges for $D = a$
- If $\boldsymbol{X}=(x_1, x_3)$ then D is irrelevant and DFP $\equiv$ UP
- DFP is useful when D is a confounder

Fair premiums & causal model

- Strict Demographic Parity

$$\hat{Y} \big| \{D = a\} = \hat{Y} \big| \{D = b\}$$

- Conditional DP
  - Suppose age (X_3) is the only permitted rating factor. Then

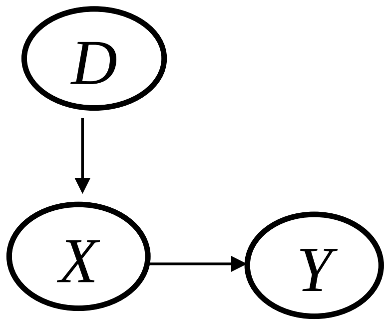
$$\hat{Y} \big| \{D = a, X_3 = x_3\} = \hat{Y} \big| \{D = b, X_3 = x_3\}$$

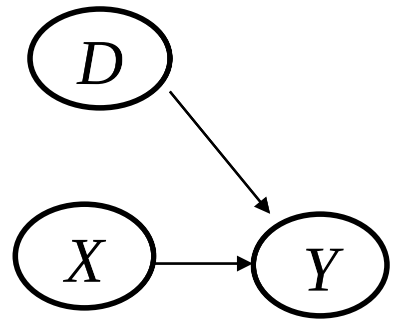
- Weak DP

$$\sum_{\boldsymbol{x}} \hat{Y}(D = a, \boldsymbol{X} = \boldsymbol{x}) p_X(\boldsymbol{x}) = \sum_{\boldsymbol{x}} \hat{Y}(D = b, \boldsymbol{X} = \boldsymbol{x}) p_X(\boldsymbol{x})$$

Notes

- If D is not a confounder? eg





- In principle, the Unaware premium (UP) = DF premium (DFP)
- But (eg) measurement error can create spurious correlations → UP $\neq$ DFP

Notes from the causal perspective

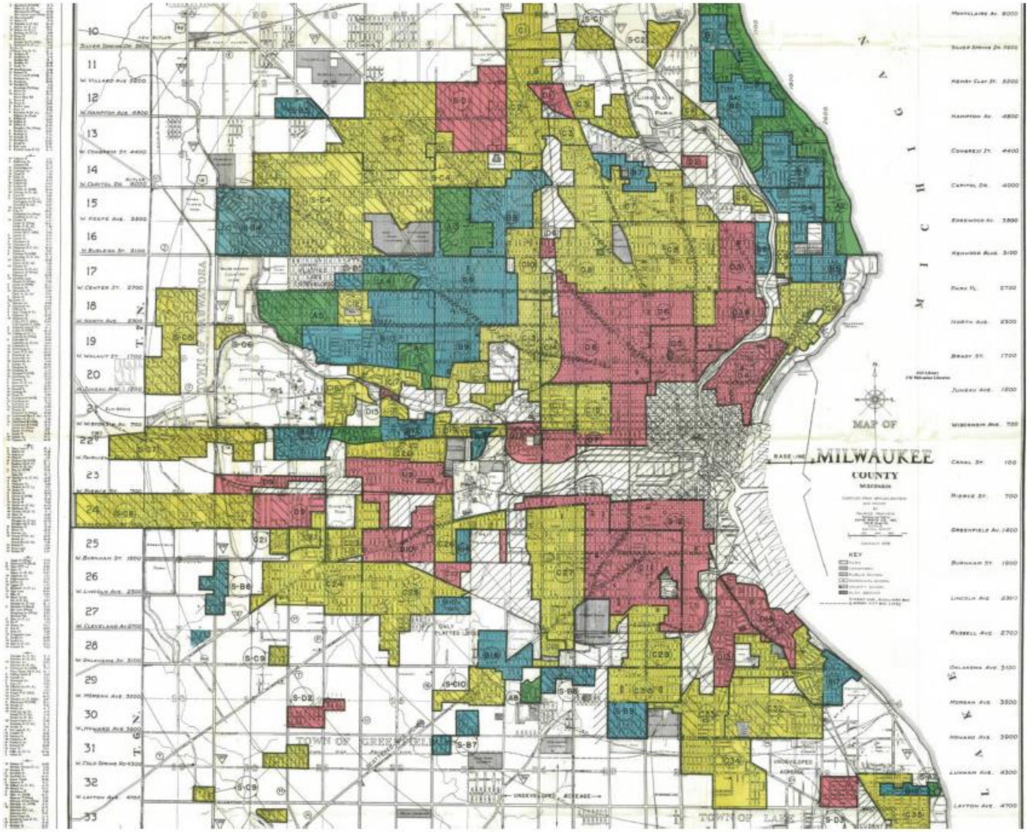
1. The DAG helps identify risk factors (causal) from other rating factors.
2. The DAG helps achieve counterfactual fairness.
3. DAGs are subjective.
4. Variable aggregation, discretization, use of proxies, will impact the reliability of causal inference.
5. Problems with high dimensions
6. Availability of Discriminatory covariate
7. Impact of telemetrics; interpretation; algorithmic fairness

DFP vs UP

- If *D* is a confounder, then the DFP is generally preferred to the UP.
- If the DFP mitigates historic social disadvantage, then the DFP is generally preferred to the UP.
- If *D* is not a risk factor (no direct impact on Y)
  - In principle UP = DFP
  - Proxies, measurement error may lead to spurious correlations
  - The UP is generally preferred to the DFP.

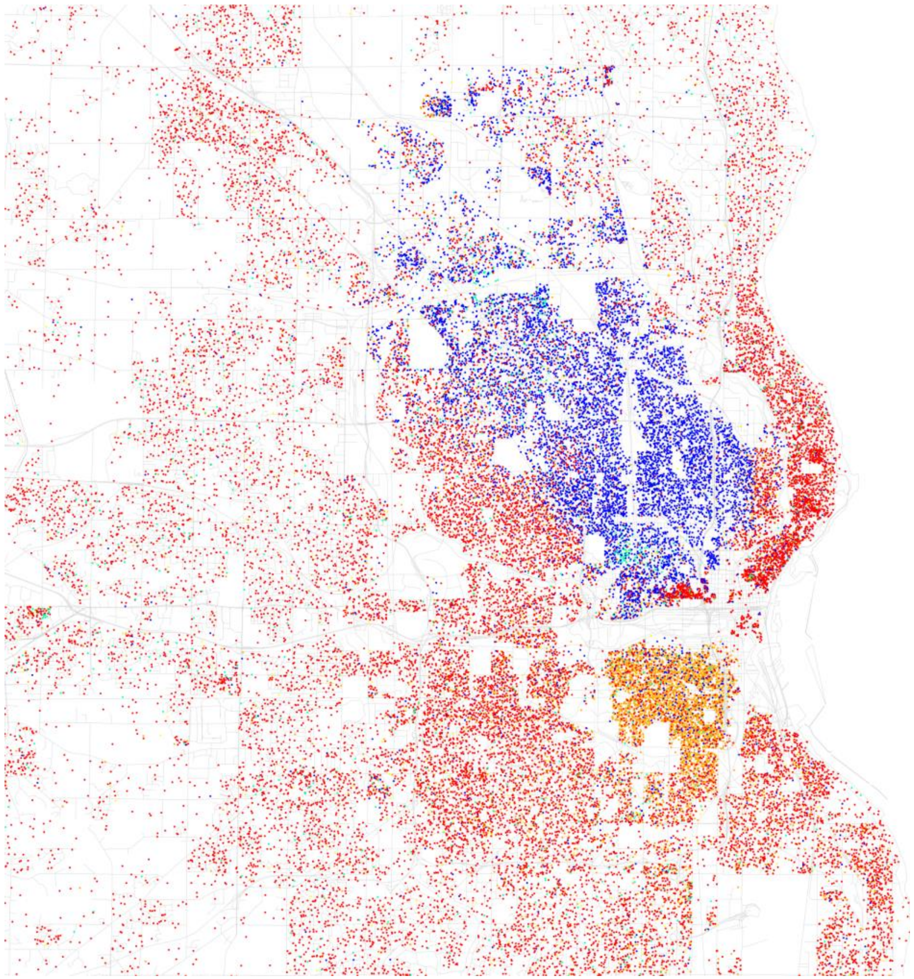
CASE STUDY: WISCONSIN AUTO INSURANCE

Redlining in Milwaukee, 1964



link to [Auto insurance premiums Milwaukee Redlining in Milwaukee,...]()

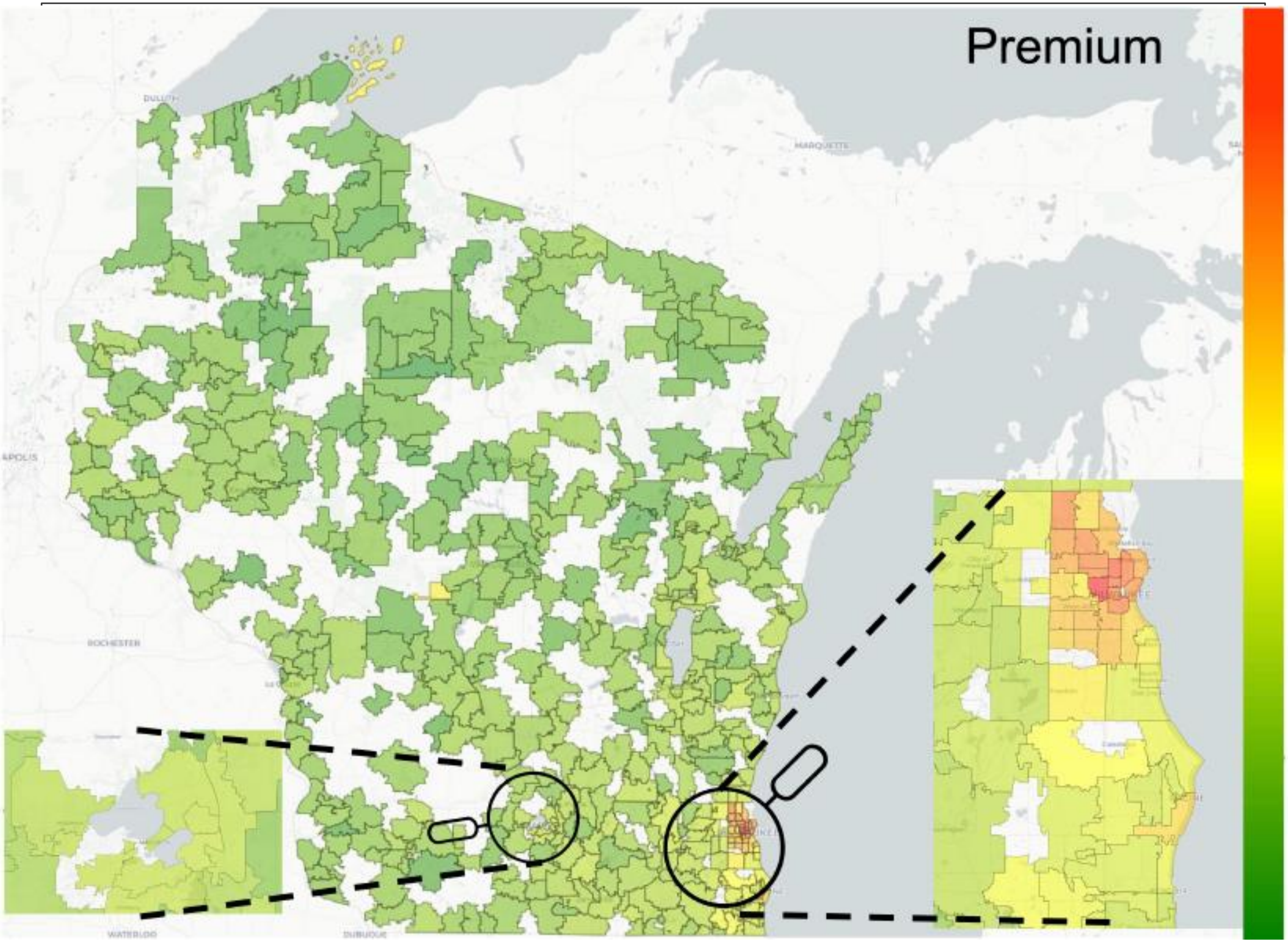
Racial distribution Milwaukee, 2010



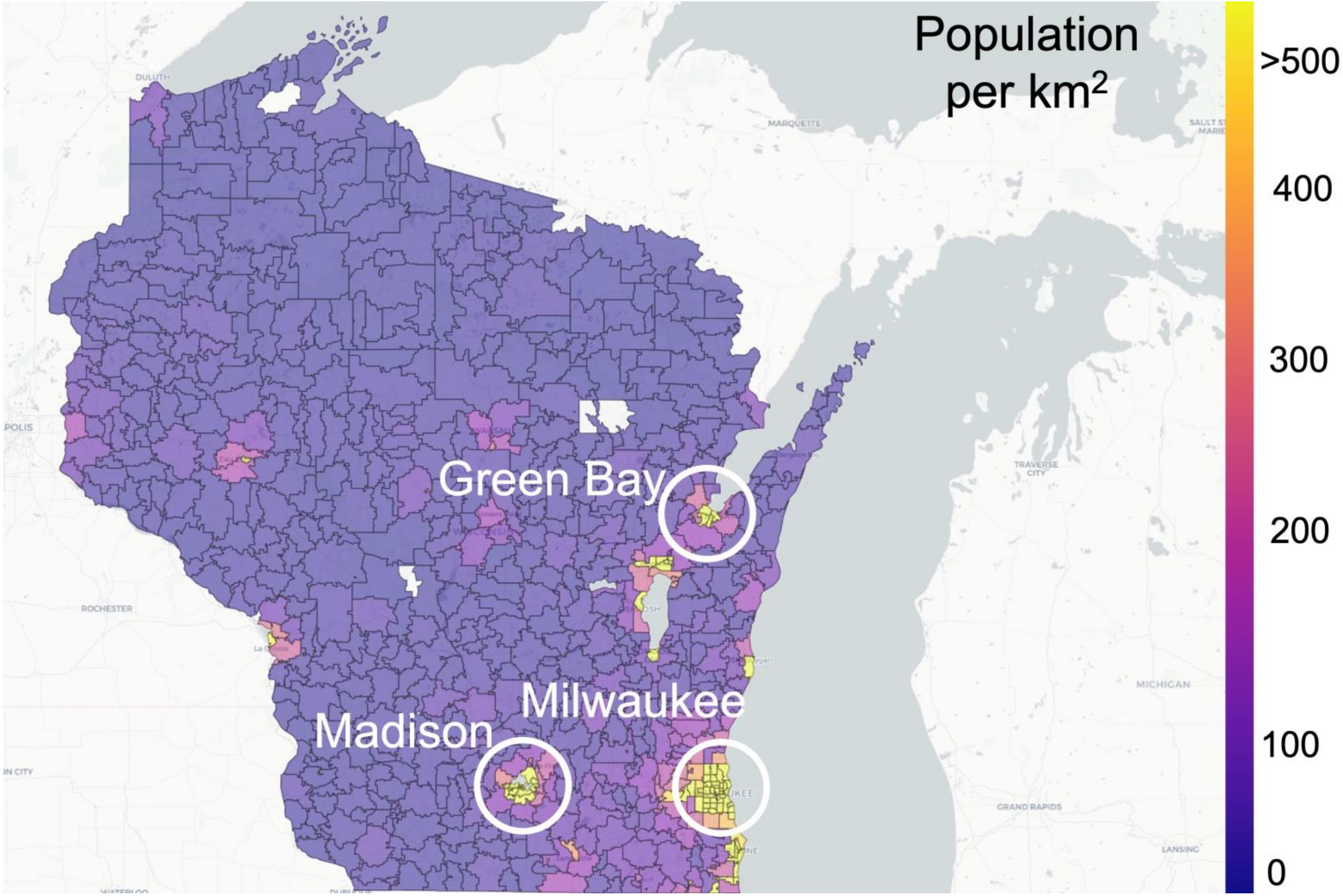
Each dot is 25 people:
White Black
Asian Hispanic
Other

Questions

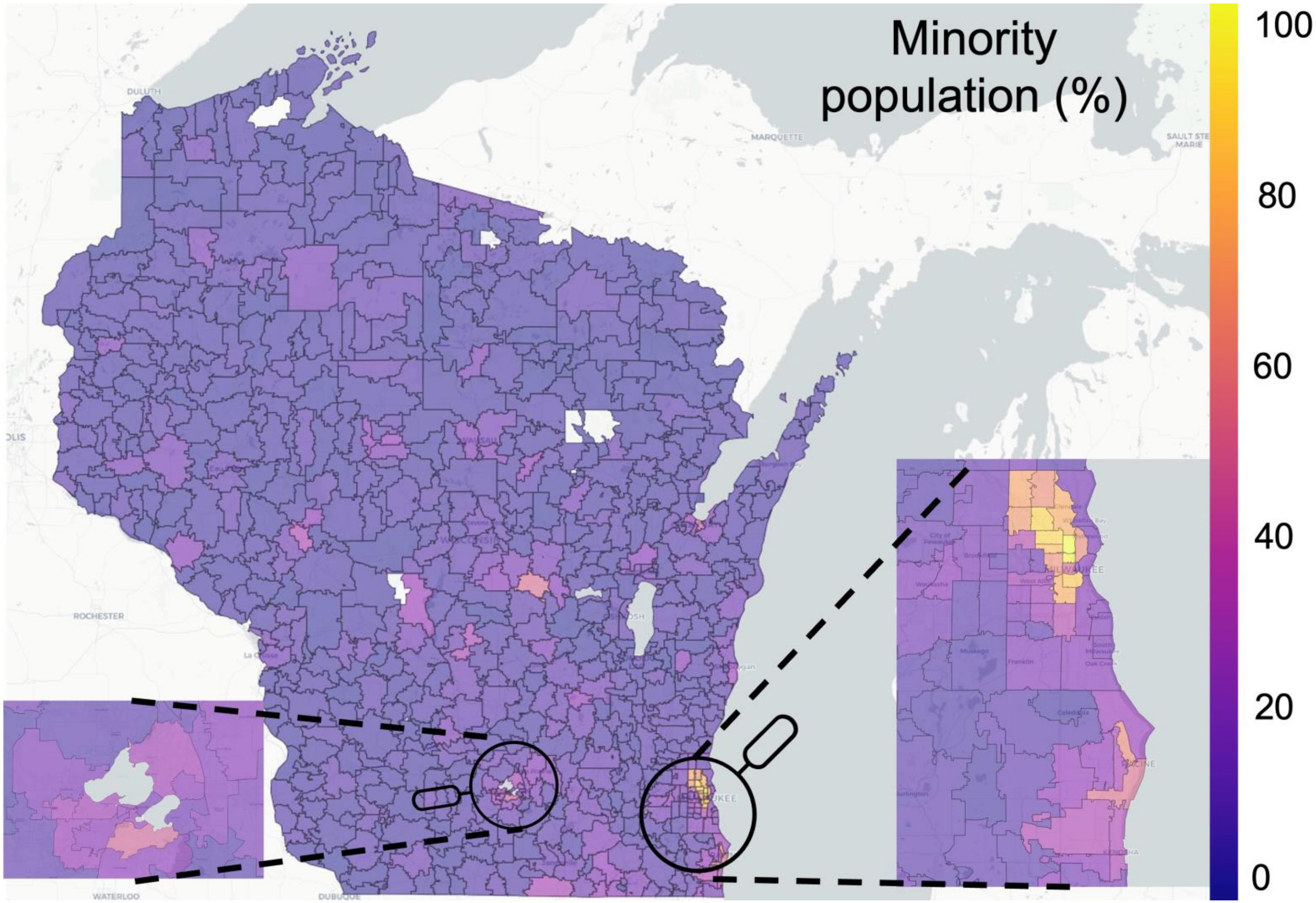
- Are there ongoing effects of the redlining era on auto insurance premiums?
- Could actuarial models, assumptions and policy design be perpetuating systemic discrimination?
- Should insurers and regulators take historic inequality into consideration?


Premium
DULUTH
MARQUETTE
ROCHESTER
WATERLOO
DUBUQUE

Population density Wisconsin

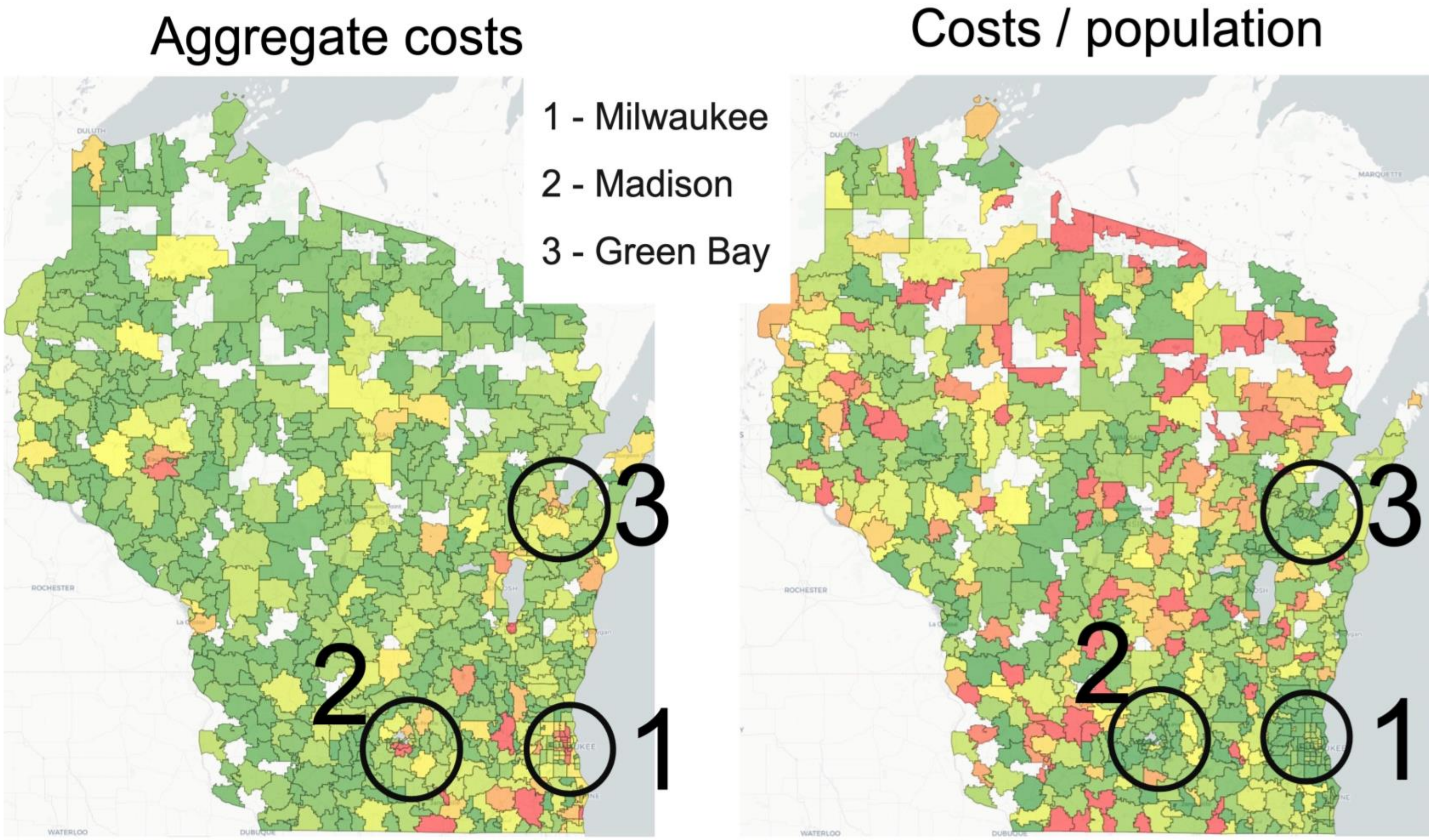


Minority population Wisconsin



Maybe Milwaukee is an outlier?

- We examined an extensive database on accidents attended by police, 2001-2020.
  - Age, gender, car, location, other vehicles, speed, time of day, severity of injuries.
- Transform accident severity into approximate cost.


Aggregate costs
Costs / population
1 - Milwaukee
2 - Madison
3 - Green Bay
1
2
3

[Link](#)

Notes

- Premiums are ~ 50%-100% higher in the predominantly Black/Hispanic neighbourhoods.
- Population density does not appear to be the reason.
- Auto insurance is vital to participate in workforce and access amenities.
- Third party liability insurance is mandatory in Wisconsin.

A MICROSIMULATION STUDY

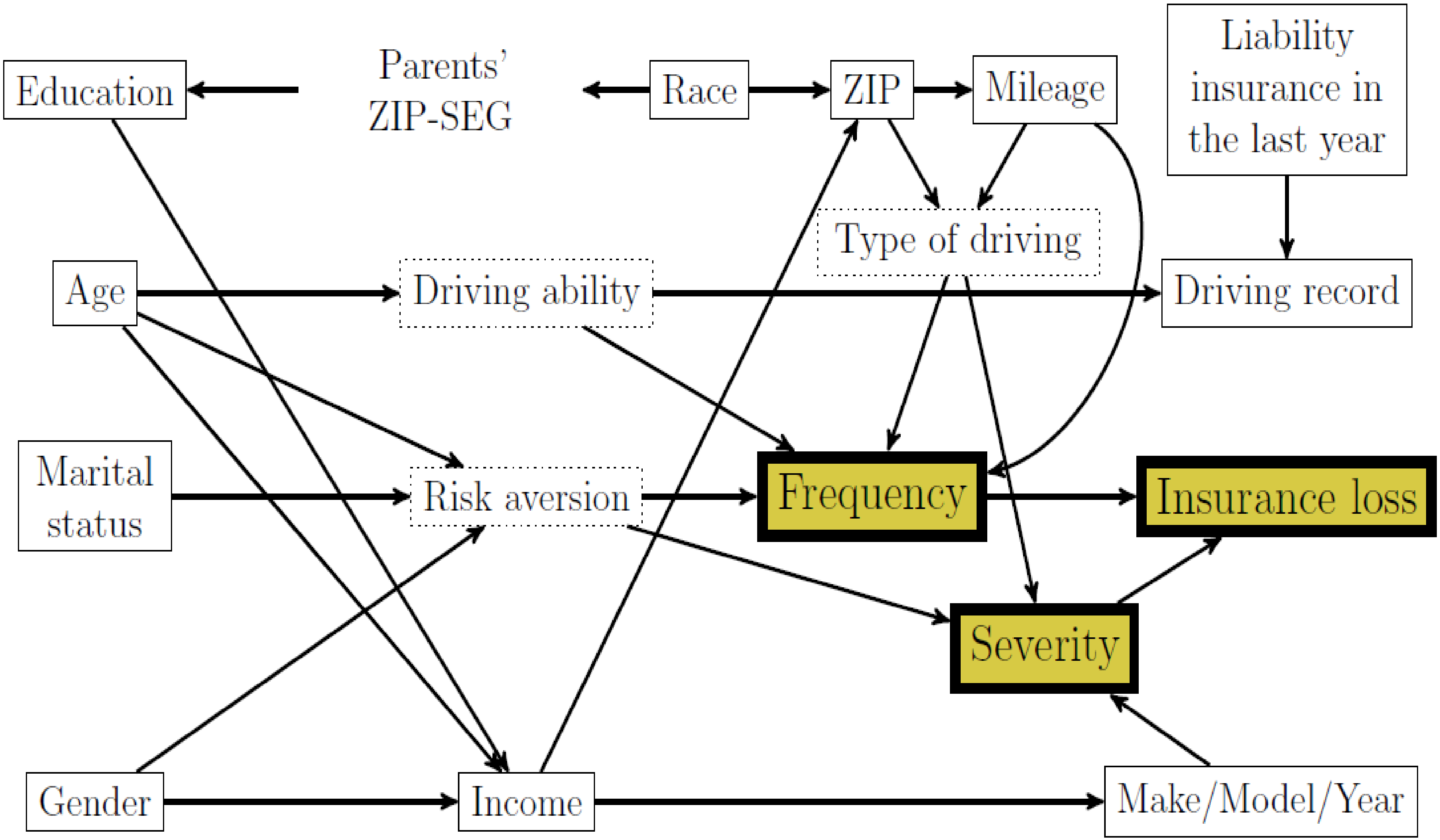
Microsimulation Experiment

- Simulate a portfolio of policyholders
- Calibrated to Wisconsin 2020 population data.
- Calibrated to Wisconsin premiums
- Calibrated to Dept of Transport data on accident frequency and severity

Simulated Variables

- Age, gender, marital status, education, driving record, car, mileage, zip code — **Rating Factors**
- Income ← **Proxied Rating Factor**
- Risk aversion, driving ability, driving type ← **Latent variables**
- Race ← **Discriminatory covariate**
- Insurance frequency / severity ← **Response Variables**

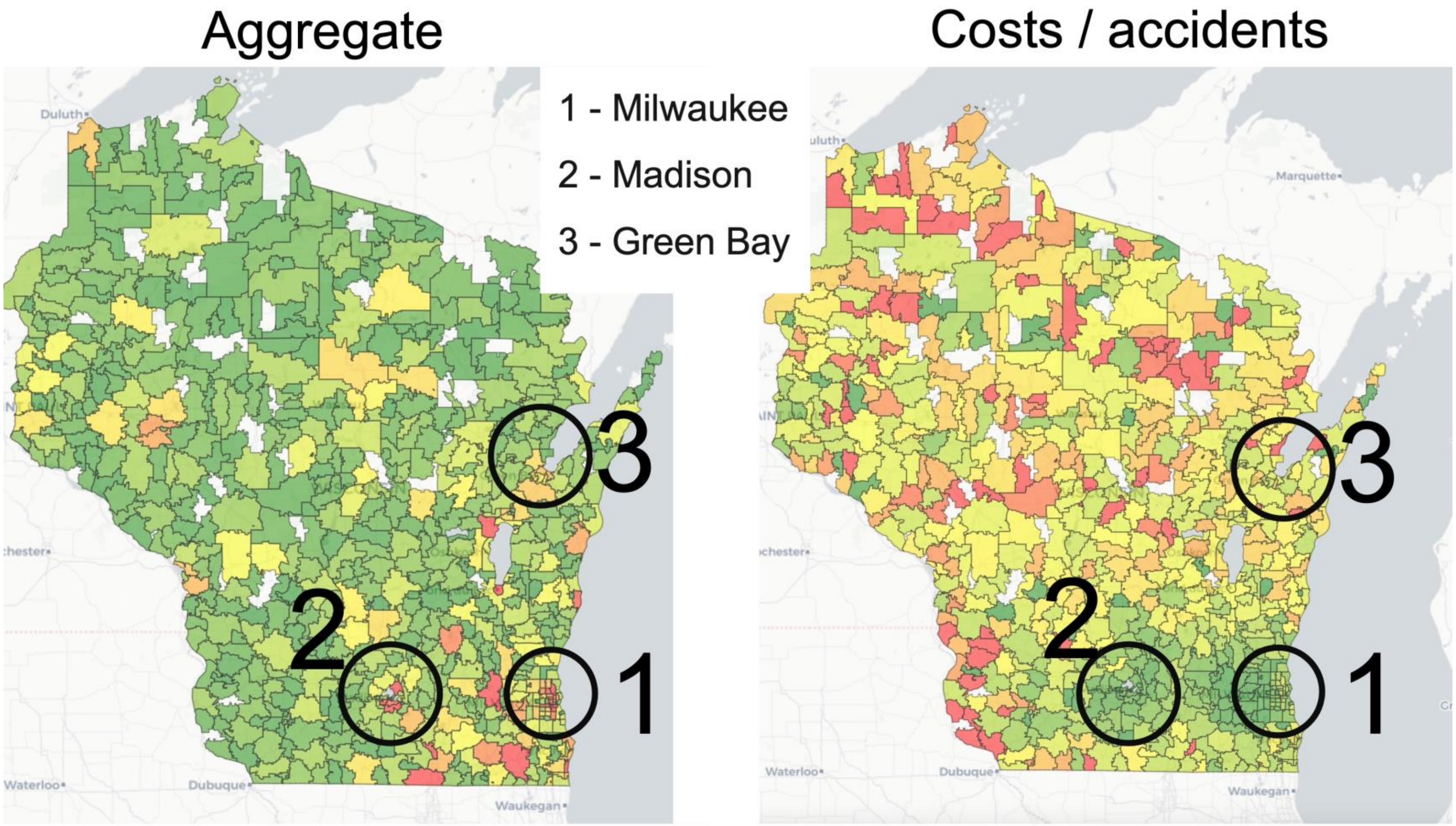
Microsimulation DAG



$\mathcal{G}_m$

Simulation results - summary

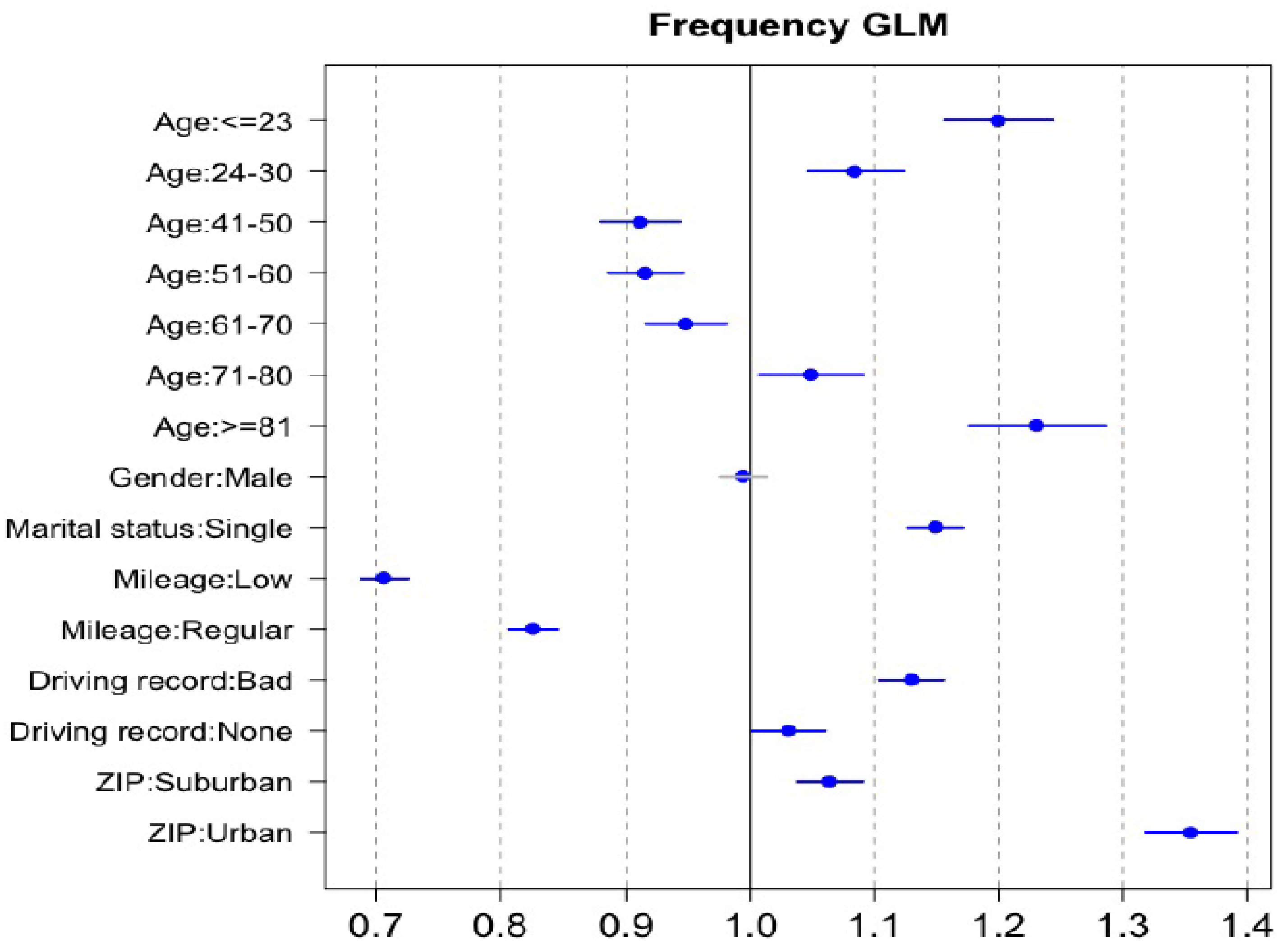
	White	Black	Asian	Hispanic	Total
Policyholders	83%	6%	3%	7%	100%
- urban	24%	86%	53%	63%	31%
- rural	48%	5%	12%	19%	42%
Claim freq.	0.187	0.200	0.189	0.190	0.188
Ave severity	6,536	5,282	6,207	5,803	6,388
Ave pure premium	1,227	1,058	1,158	1,115	1,206


Aggregate
Costs / accidents
1 - Milwaukee
2 - Madison
3 - Green Bay
1
2
3
Duluth
Marquette
Waterloo
Dubuque
Waukegan

Link

What does the insurer see?

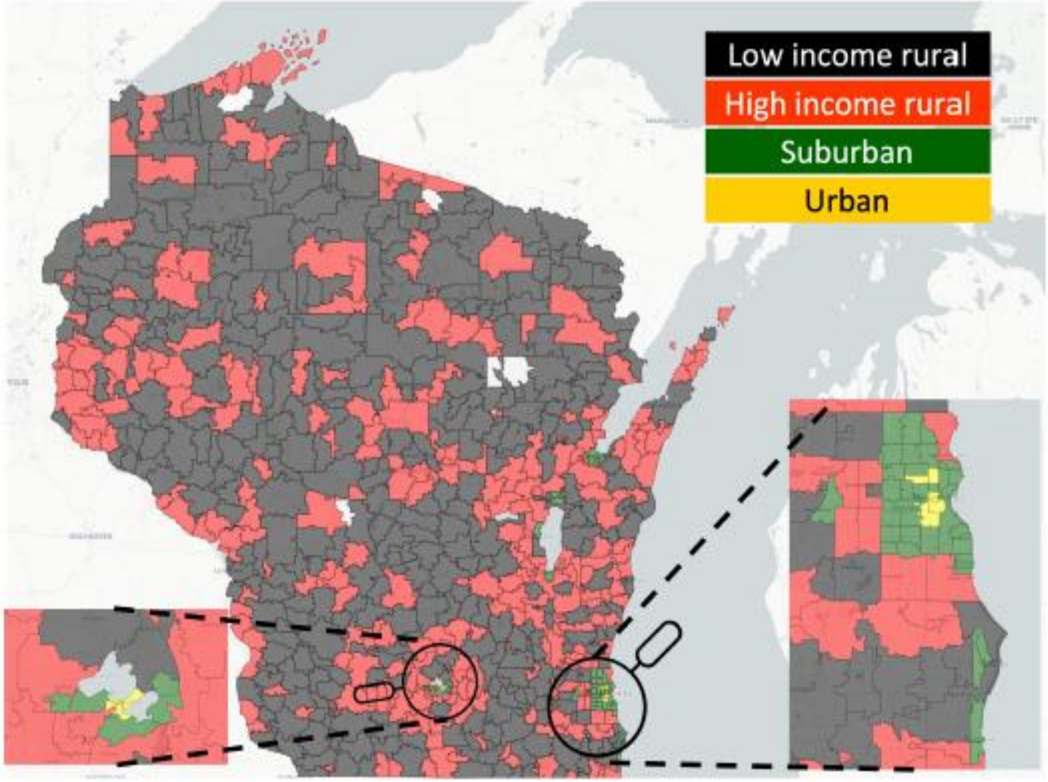
Insurer’s analysis of frequency data



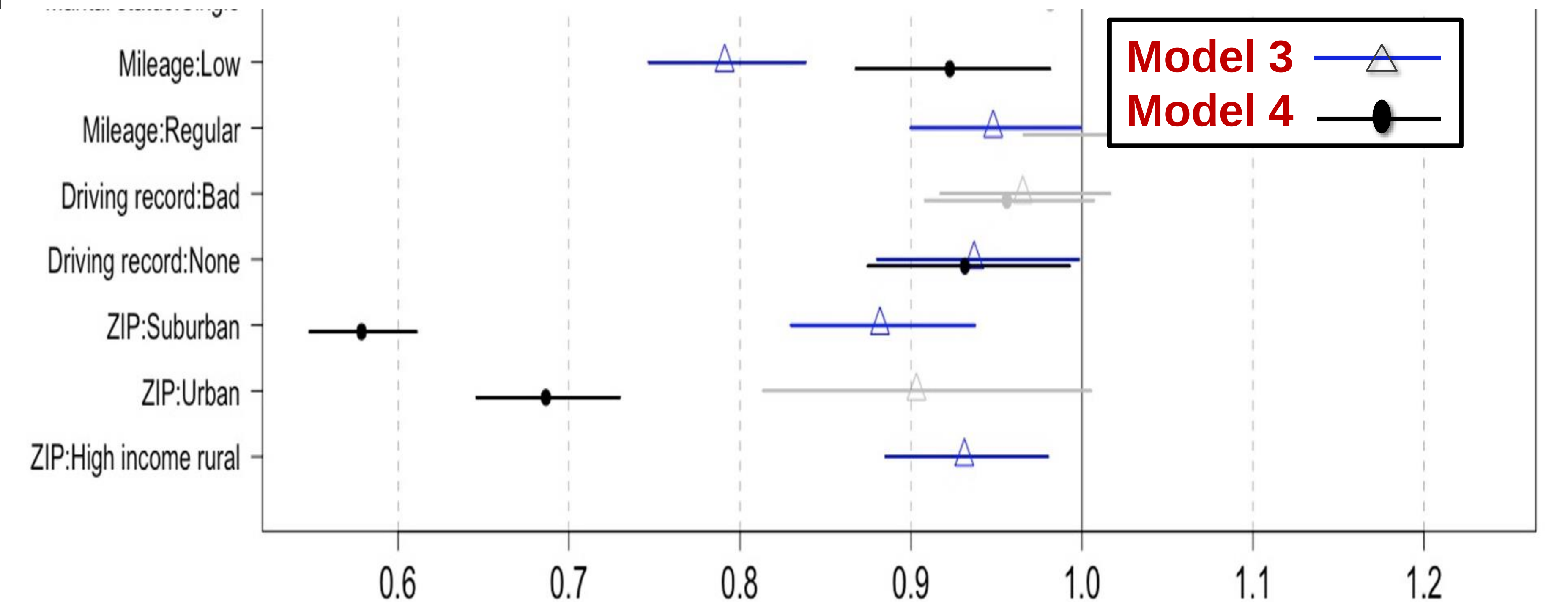
Insurer's analysis of severity data

- Model 1: Independent of all covariates
- Model 2: Dependent on car make/model/year
- Model 3: GLM, area income proxy for zip code, all other rating factors
- Model 4: GLM, all rating factors

Model 3: income ←→ zip code proxy



Insurer's analysis of simulated severity data



Discrimination metrics

Relative bias tests:

- Relative bias for Model *j*: $$\phi_j(d) = \frac{\hat{\mu}_j(d)}{\mu(d)} - 1$$
- where $\hat{\mu}_j(d)$ is the Model *j* estimated average pure premium for group $d \in D$
- and $\mu(d)$ is the true average pure premium for group $d \in D$

Why use relative bias?

- We know the exact premium.
- We are testing for disparate impact of model assumptions
  - specifically, leading to larger values of ϕ

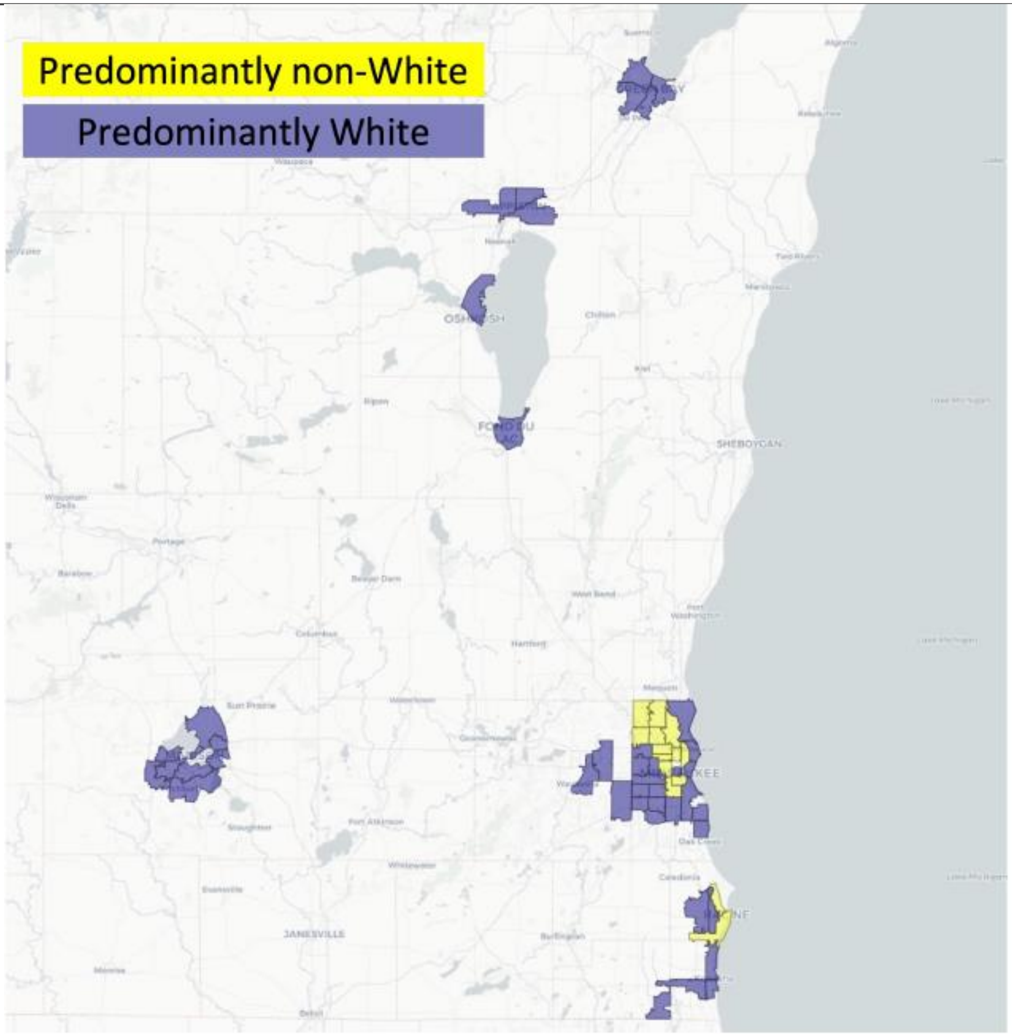
Discrimination metrics

1. Frequency or severity parameters $\perp$ race?
2. Relative bias ratio:

$$\frac{\phi(d = \text{White})}{\phi(d = \text{ Black/Asian/Hispanic})} <> 0.8$$

3. Pairwise comparison of relative bias
   - 14 mostly minority ZC vs 48 mostly White

Proxy discrimination sample



Discrimination Results

	Direct	Indirect (1)	Indirect (2)	Indirect (3)
Frequency	**No**	**No**	**No**	**Yes**
Model 1	**No**	**Yes**	**Yes**	**Yes**
Model 2	**No**	**Yes**	**Yes**	**Yes**
Model 3	**No**	**No**	**No**	**Yes**
Model 4	**No**	**No**	**No**	**No**

Other issues

- Claim rate given loss
  - wealthier p/h may claim less often
- Impact of No Claims Bonus systems
  - frequency dependent, not severity dependent
- Bias in policing
- Bias in claims management (Huskey vs State Farm)
- Potential impact of telemetrics
  - Algorithmic bias

Other issues

- Would demographic parity be appropriate here?
- Would the Discrimination-Free Premium be appropriate here?
- How might counterfactual fairness be taken into consideration?
- Causal inference requires subjective judgement – might this introduce more bias?

Summary

- The causal framework can illuminate sources of unfairness.
- Discrimination can arise from a common actuarial model of frequency/severity.
- Our results are consistent with empirical evidence in terms of premiums and insurance losses.
- Using causal rating factors can mitigate discrimination.
- Rating systems (underwriting, NCB, rating factors) can exacerbate unfairness

References

- Araiza Iturria, C A (2023) Discrimination in Insurance Pricing. PhD Thesis. University of Waterloo.
- Araiza Iturria CA, Hardy MR, Marriott P (2022) A discrimination-free premium under a causal framework. SSRN.
- Frees EW, Huang F (2021) The discriminatory pricing actuary. NAAJ.
- Kusner, M.J., Loftus, J., Russell, C. and Silva, R. (2017) Counterfactual fairness. Advances in neural information processing systems
- Lindholm M, Richman R, Tsanakas A, Wuthrich M (2022). Discrimination-free insurance pricing. ASTIN 52(1).
- Lindholm, M., Richman, R., Tsanakas, A. and Wuthrich, M.V., (2023). What is Fair? Proxy Discrimination vs. Demographic Disparities in Insurance Pricing. SSRN
- Xin, X. and Huang, F., 2022. Anti-discrimination insurance pricing: regulations, fairness criteria, and models. Fairness Criteria, and Models