Motivation
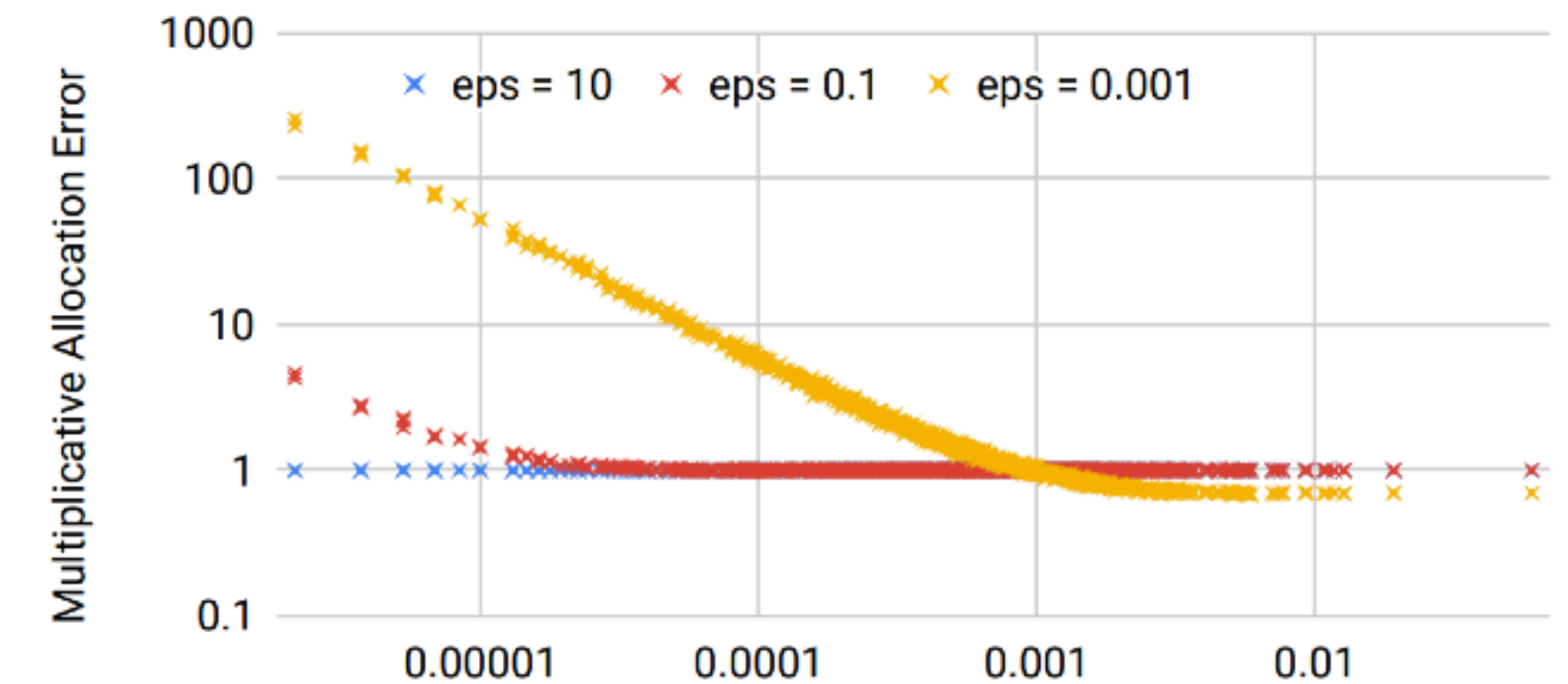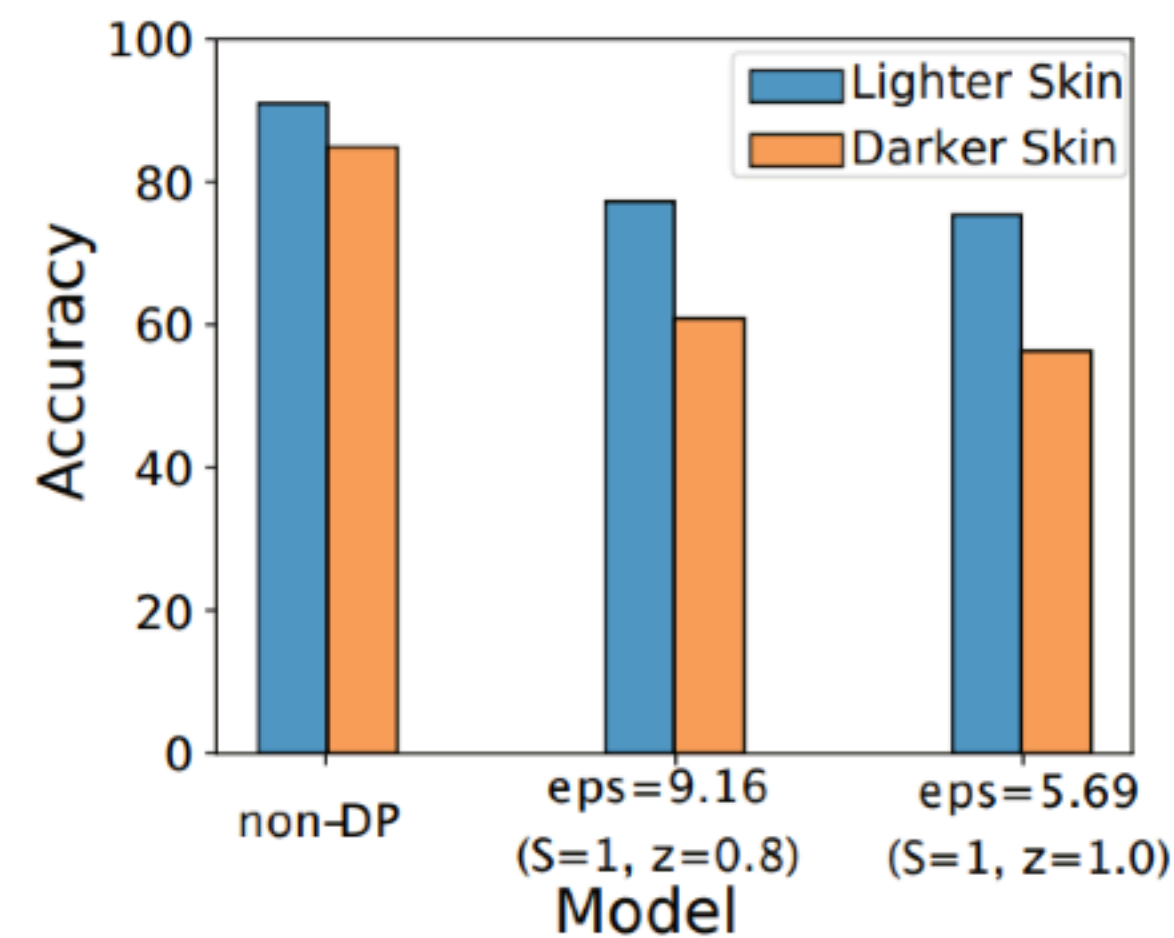
DP has a disparate effect[1] on:

- Statistics[2]
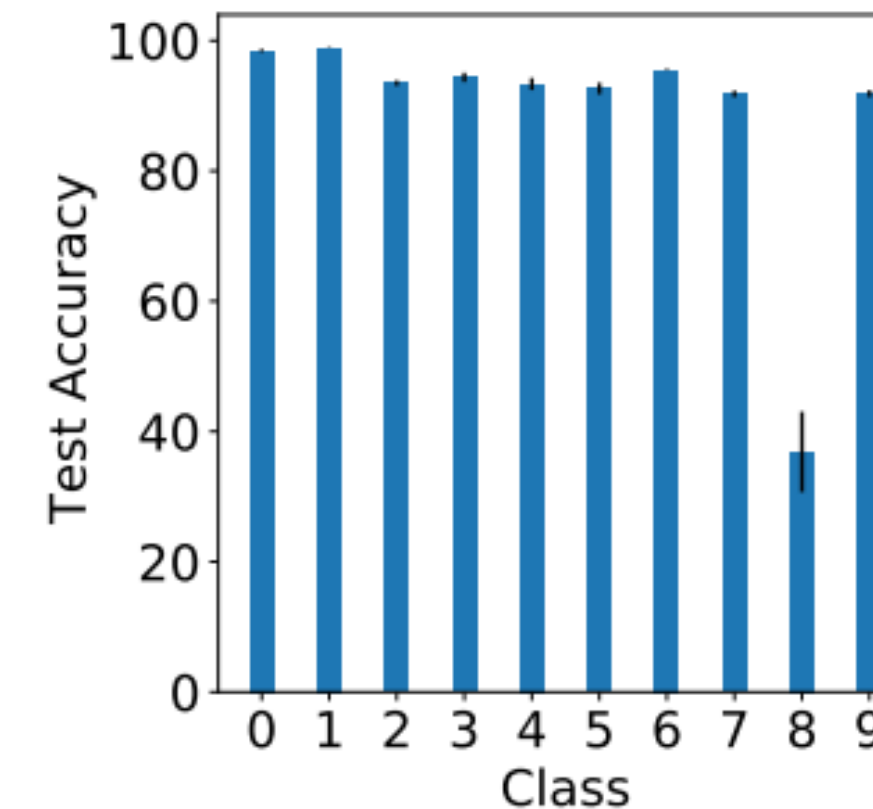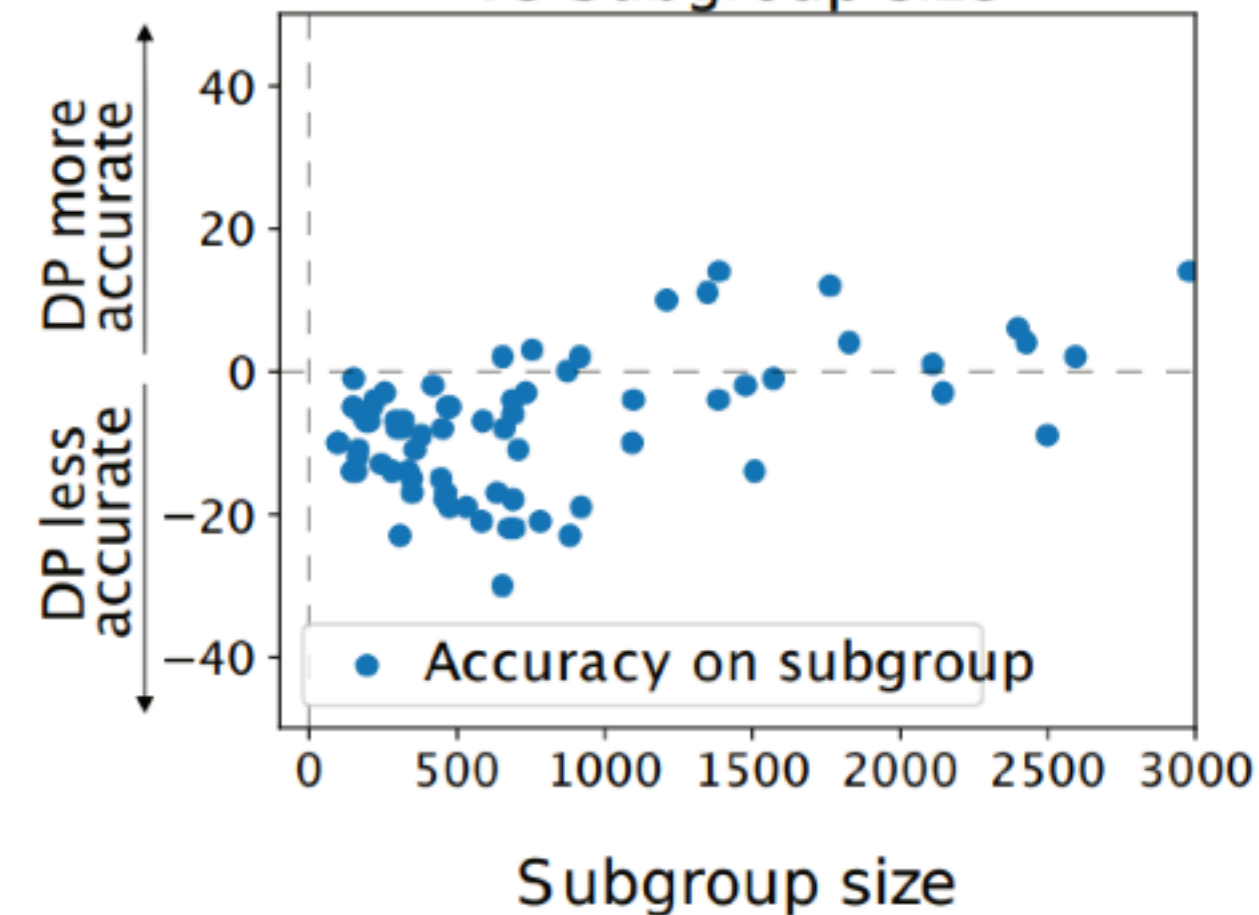
- Deep learning classifiers[3,4]

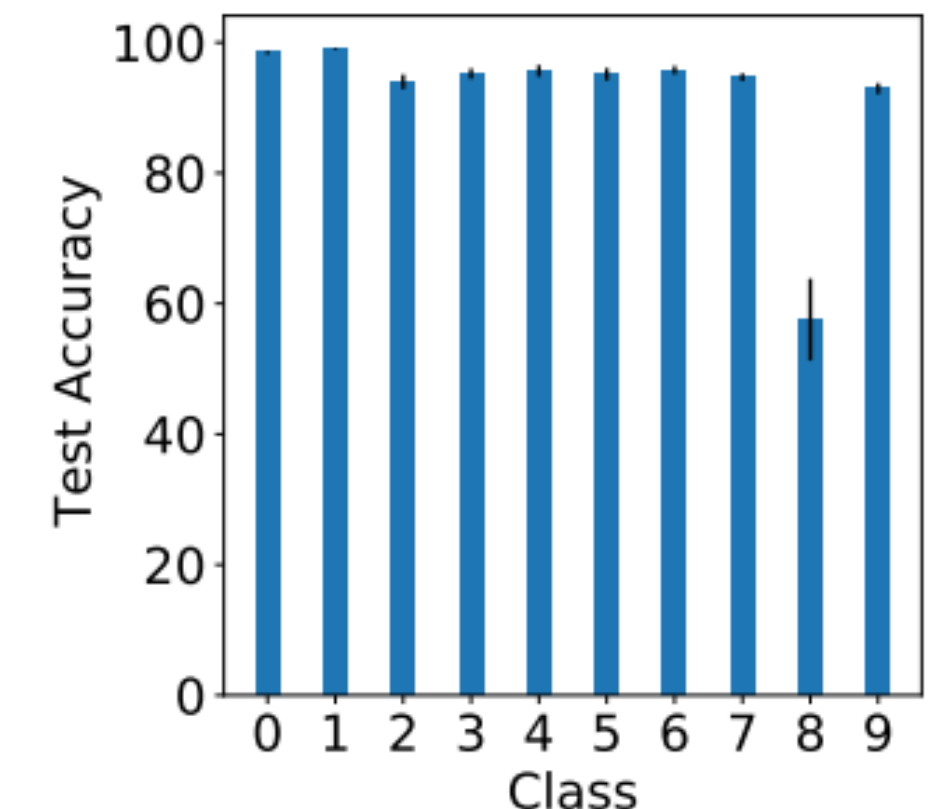Multiplicative Allocation Error in Michigan with Laplace

(a) Accuracy vs Model type

(b) DP model accuracy relative to non-DP vs Subgroup size

(b) DP-SGD for $\varepsilon = 5$

(e) PATE for $\varepsilon = 5$

[1]Fioretto et al., Differential Privacy and Fairness in Decisions and Learning Tasks: A Survey
[2]Kuppam et al., Fair Decision Making using Privacy-Protected Data
[3]Bagdasaryan et al., Differential privacy has disparate impact on model accuracy
[4]Uniyal et al., DP-SGD vs PATE: Which Has Less Disparate Impact on Model Accuracy?

Goal

Empirically evaluate the disparate effect DP causes on generative models:

- class/subgroup size

- class/subgroup accuracy