

## Idea

How do physicians respond to changes in average characteristics in a clinical trial when inferring results on a subpopulation? Here we start with a stylized optimal approach. Later we will look at how physicians respond in the real world. First we examine a simple solution in which two subpopulations comprise the population and share equal variance. Interestingly our model suggests it is optimal to under-represent minority groups. It seems because error is linear in marginal changes in representation?

## Simulate

- $N = 100,000$  people
- Real world is one draw of gammas from the prior distribution for each person
- How much we should update our guess of that person's treatment effects towards the overall ATE depends on what happens in a bunch of simulated worlds
- So now we make 1000 more columns, each of which are a simulated world, where we redraw gamma from the prior distribution for every individual
- In each simulated world, for each individual, we can compute  $\beta_i$  – their treatment effect – and  $\beta_{ATE}$ , the average treatment effect
- How much we update in the real-world depends on how  $\beta_i$  and  $\beta_{ATE}$  correlate across the 1000 simulations for each individual
- current closest paper