

Towards Data-Driven Affirmative Action Policies under Uncertainty

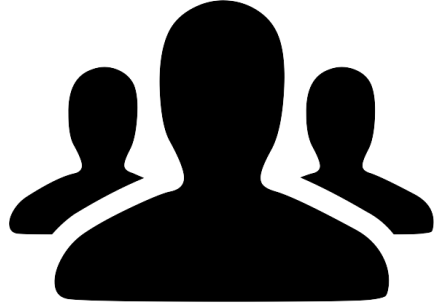
Corinna Hertweck

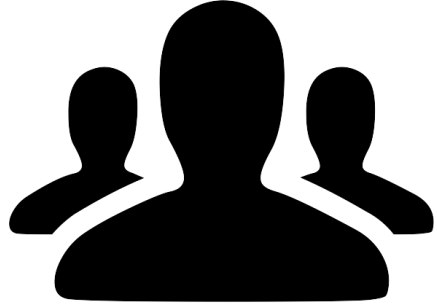
Joint work with Carlos Castillo and Michael Mathioudakis



Case study: Chile

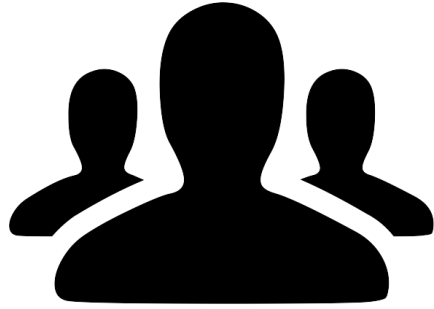
Centralized admission system





send number of spots to

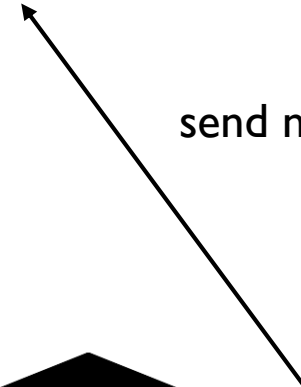


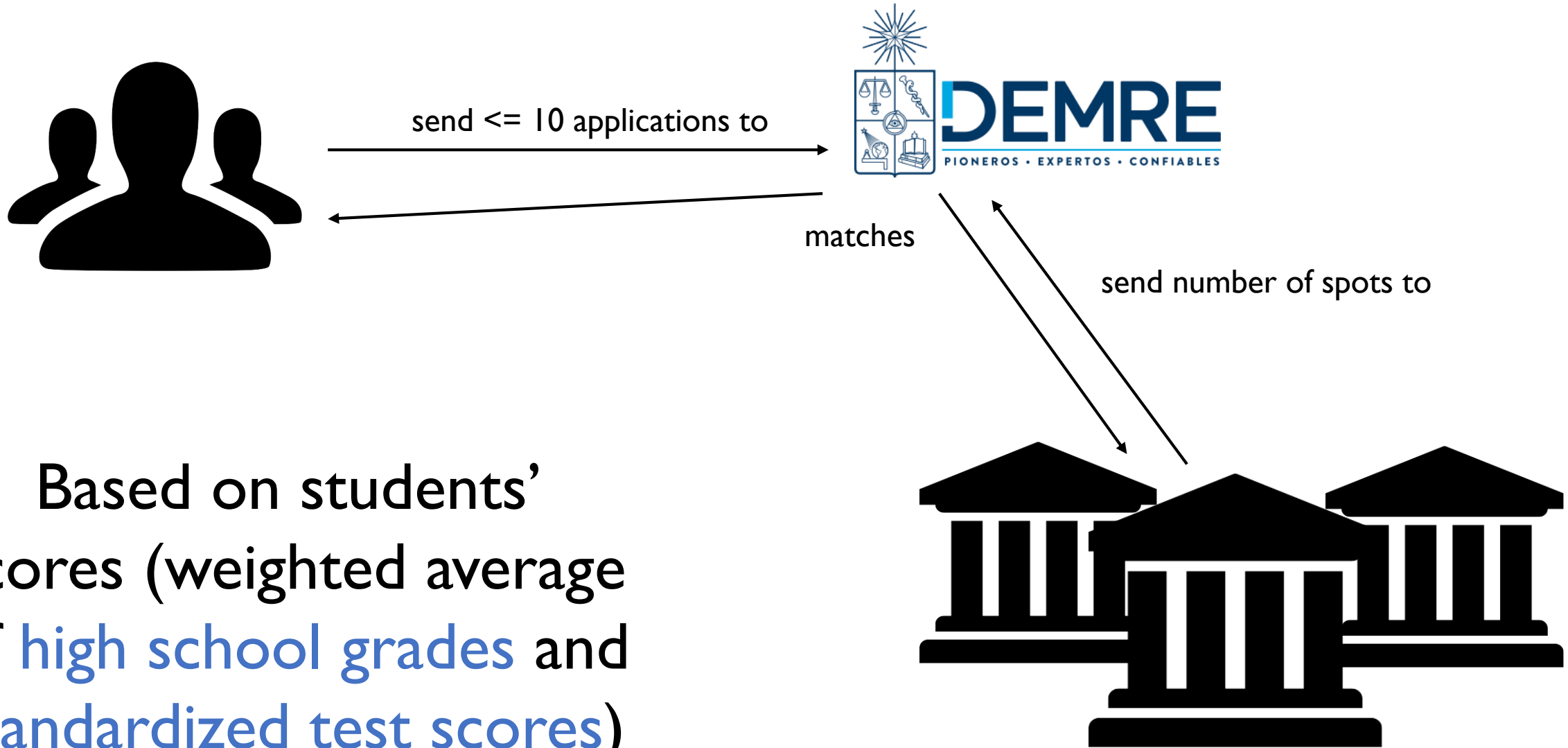


send ≤ 10 applications to

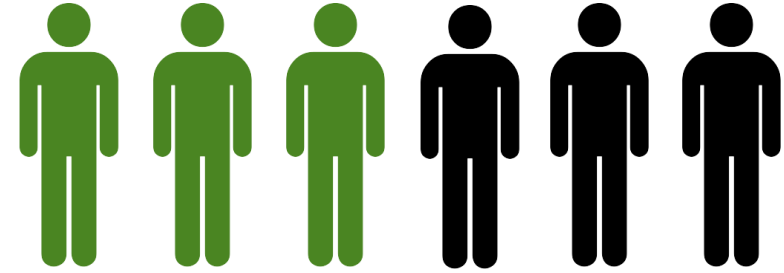
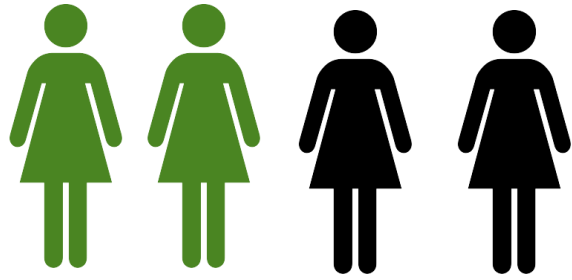


send number of spots to





Based on students' scores (weighted average of **high school grades** and **standardized test scores**)



Statistical parity difference (SPD):

$$SPD = P(\text{accepted} | A = a) - P(\text{accepted} | A \neq a)$$

0: ideal

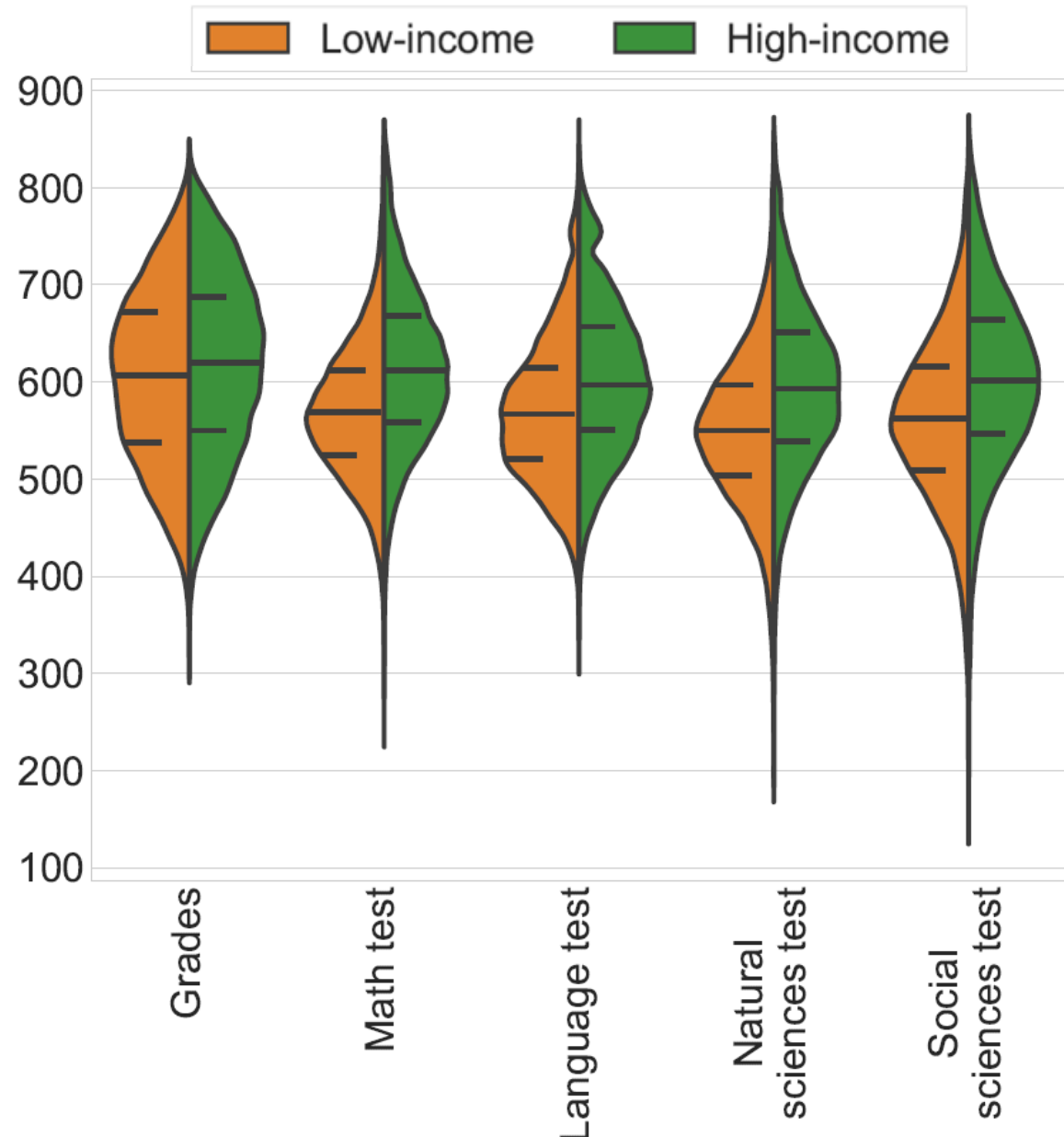
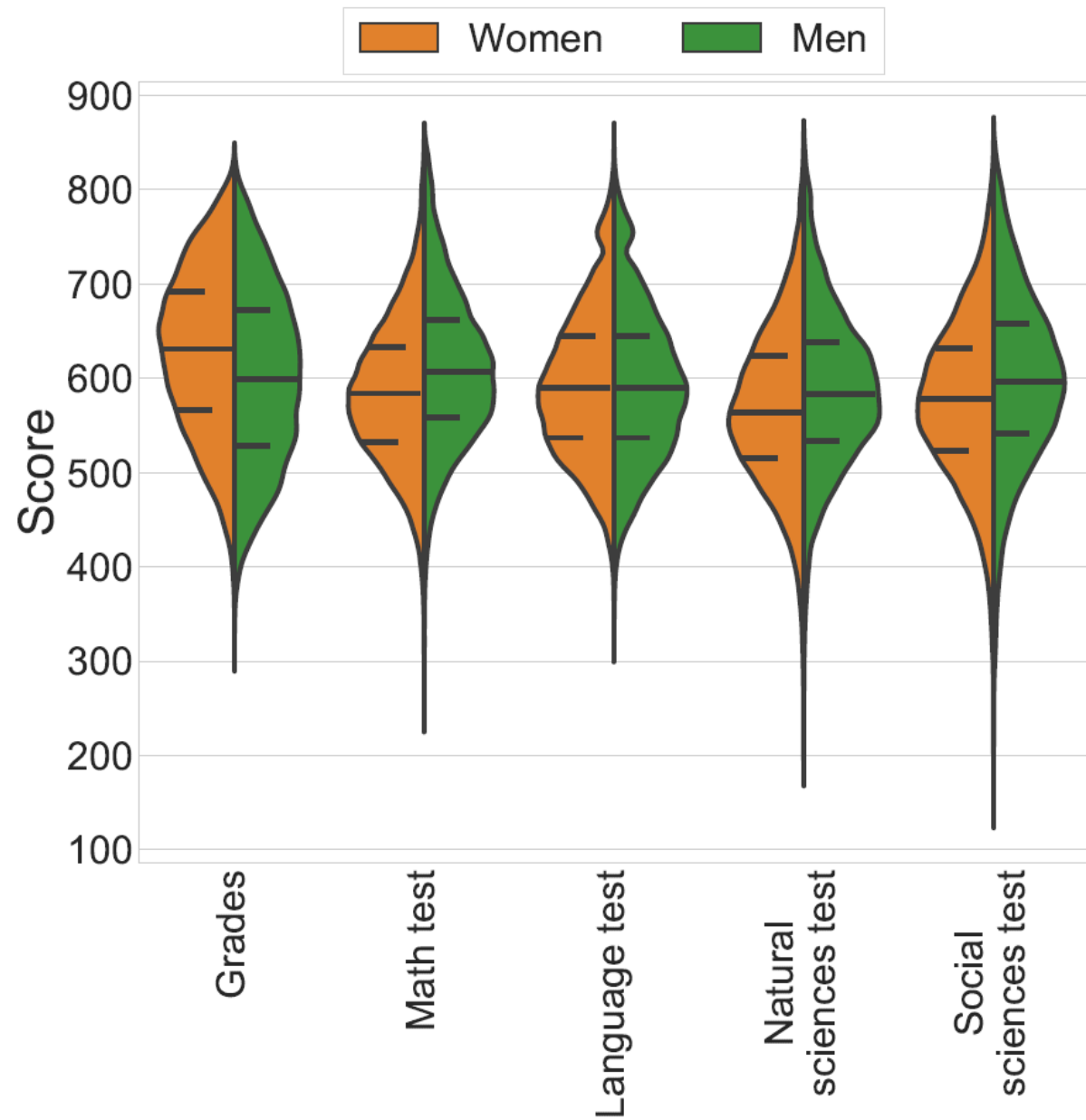
-0.1 – 0.1: “fair”

A: sensitive attribute (e.g.
gender, race, income)

gender: 8%

income: 10%

Programs with **strongly unequal admission rates**,
i.e., $|\text{SPD}| > 0.1$



Affirmative Action Policies

The background of the slide is a photograph of a large, historic stone building, likely a university campus. The building features a prominent central tower with a clock face and a crenellated roofline. The courtyard in front of the building is paved and has a central fountain. A man in a suit and hat is standing in the foreground on the left, and another person is visible in the distance on the right. The sky is blue with scattered clouds.

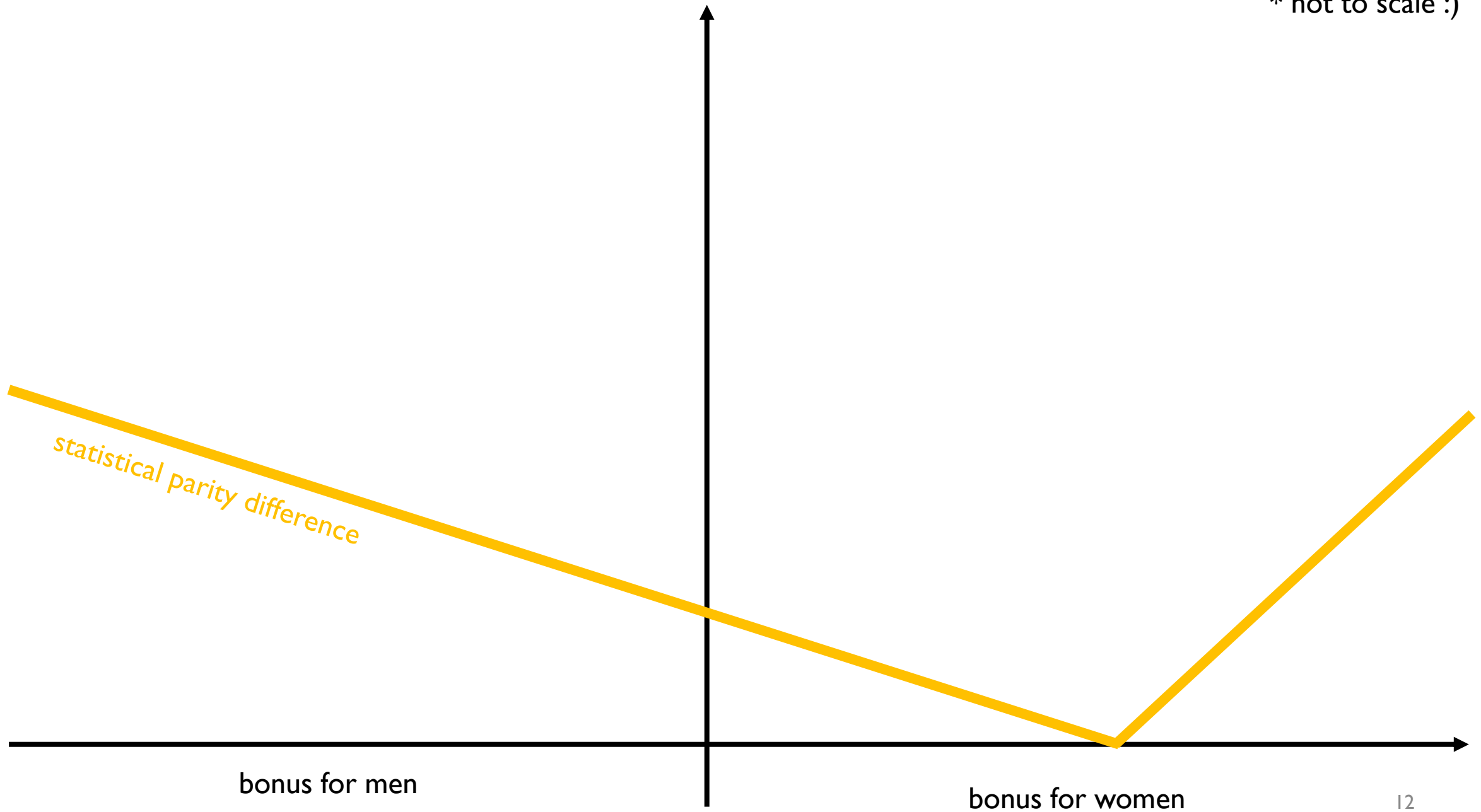
Favor disadvantaged groups

- Quota
- **Bonus points**
- Etc.



What's a good number of bonus points?

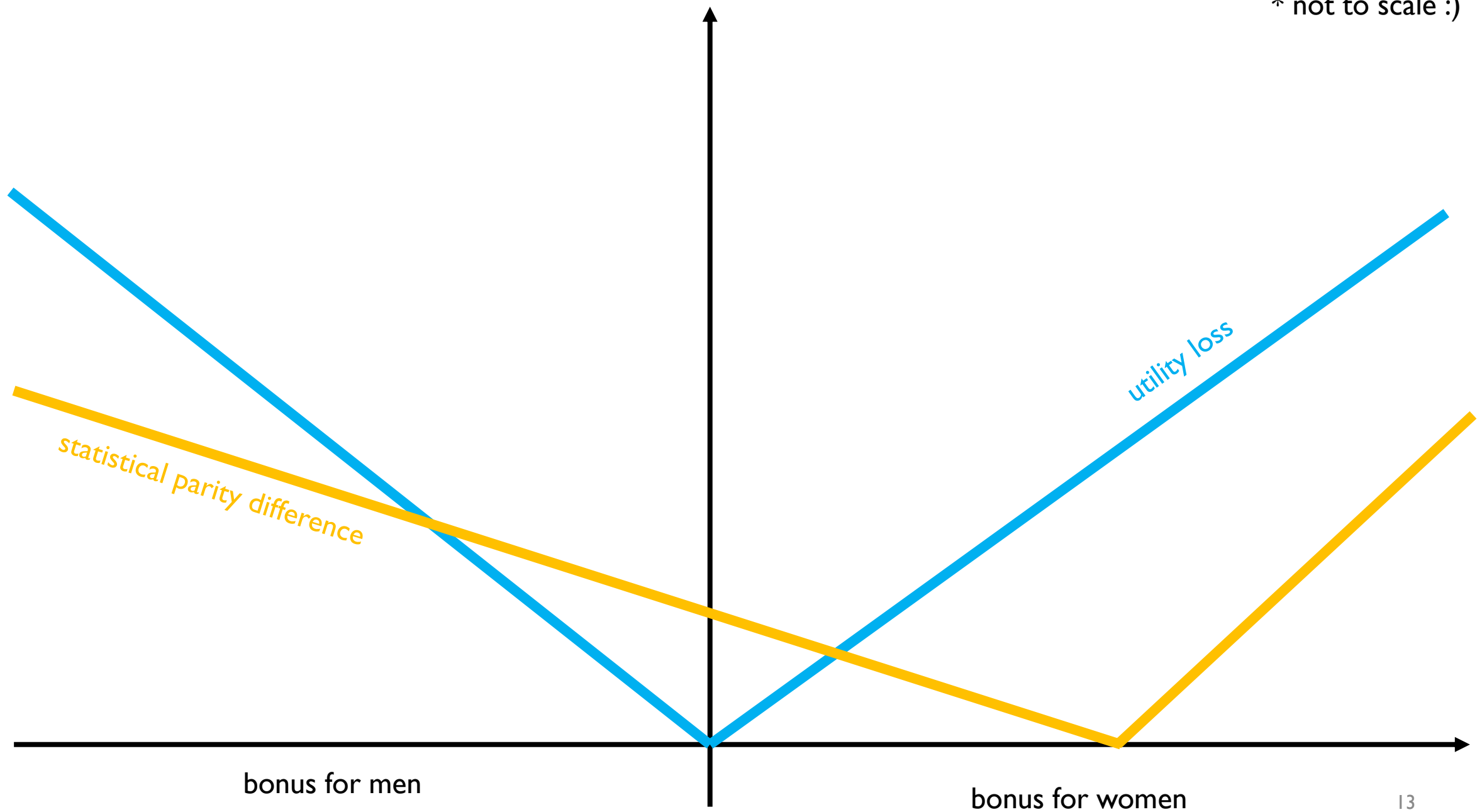
* not to scale :)



bonus for men

bonus for women

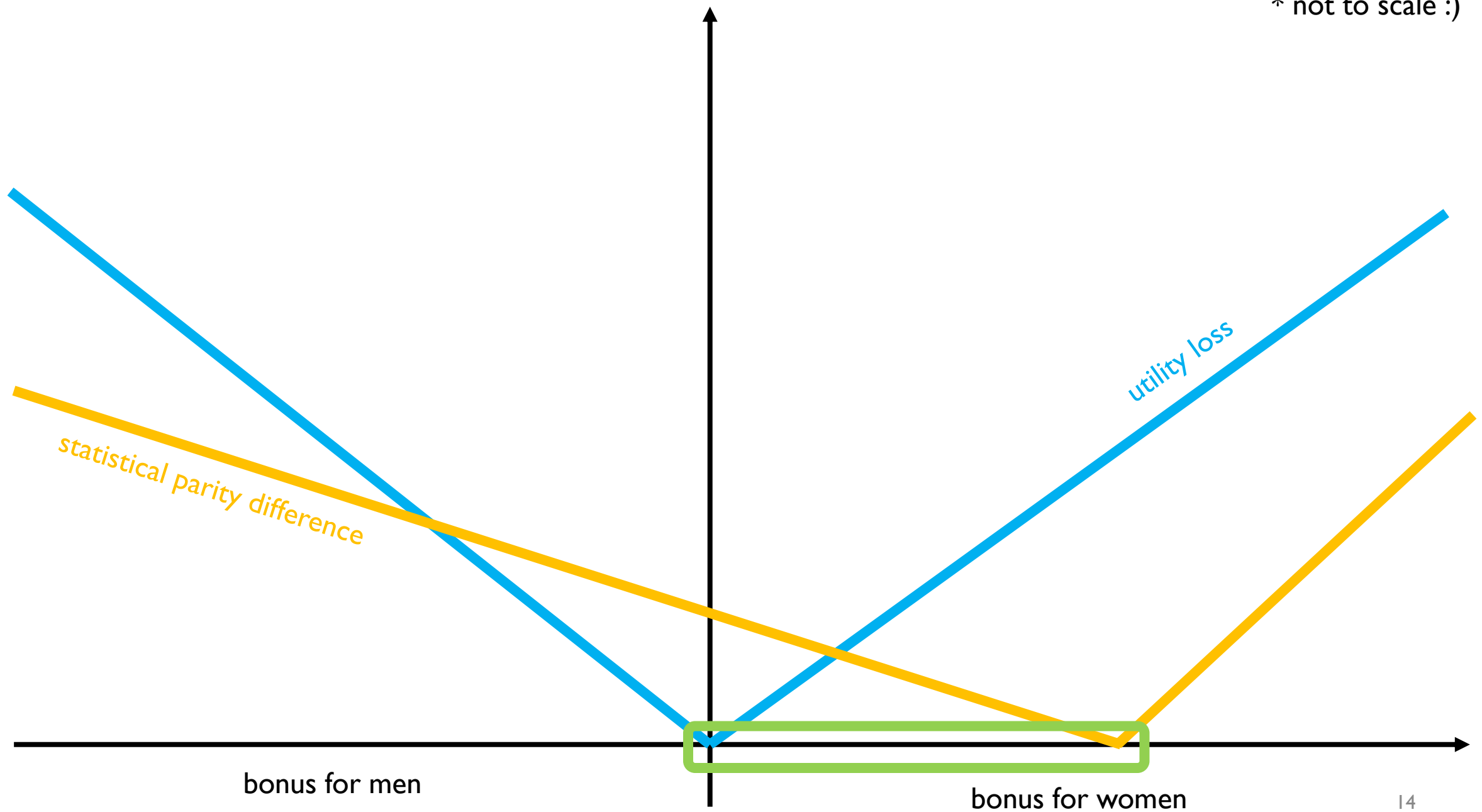
* not to scale :)



bonus for men

bonus for women

* not to scale :)



statistical parity difference

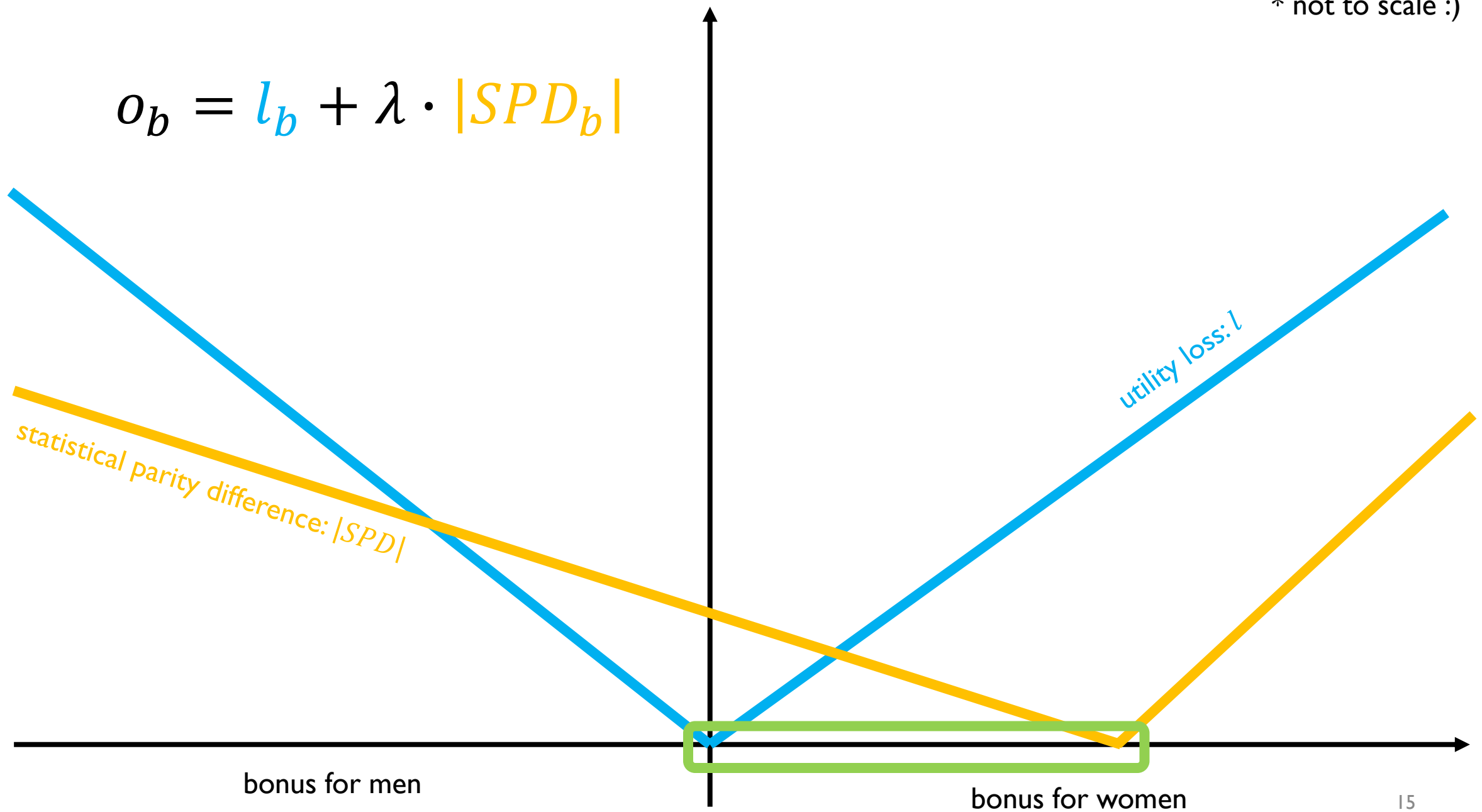
utility loss

bonus for men

bonus for women

* not to scale :)

$$o_b = l_b + \lambda \cdot |SPD_b|$$



bonus for men

bonus for women



**UNCERTAINTY
AHEAD**



Data

2004 – 2017:

- Student features (e.g. 2016: 60 000)
- Program features (e.g. 2016: 1 500)
- Applications

Last year's ideal policy

Pro: Simple, fast, transparent

Con: Risky (outlier)



Average policies of multiple years

Decreases risk

Limited number of years





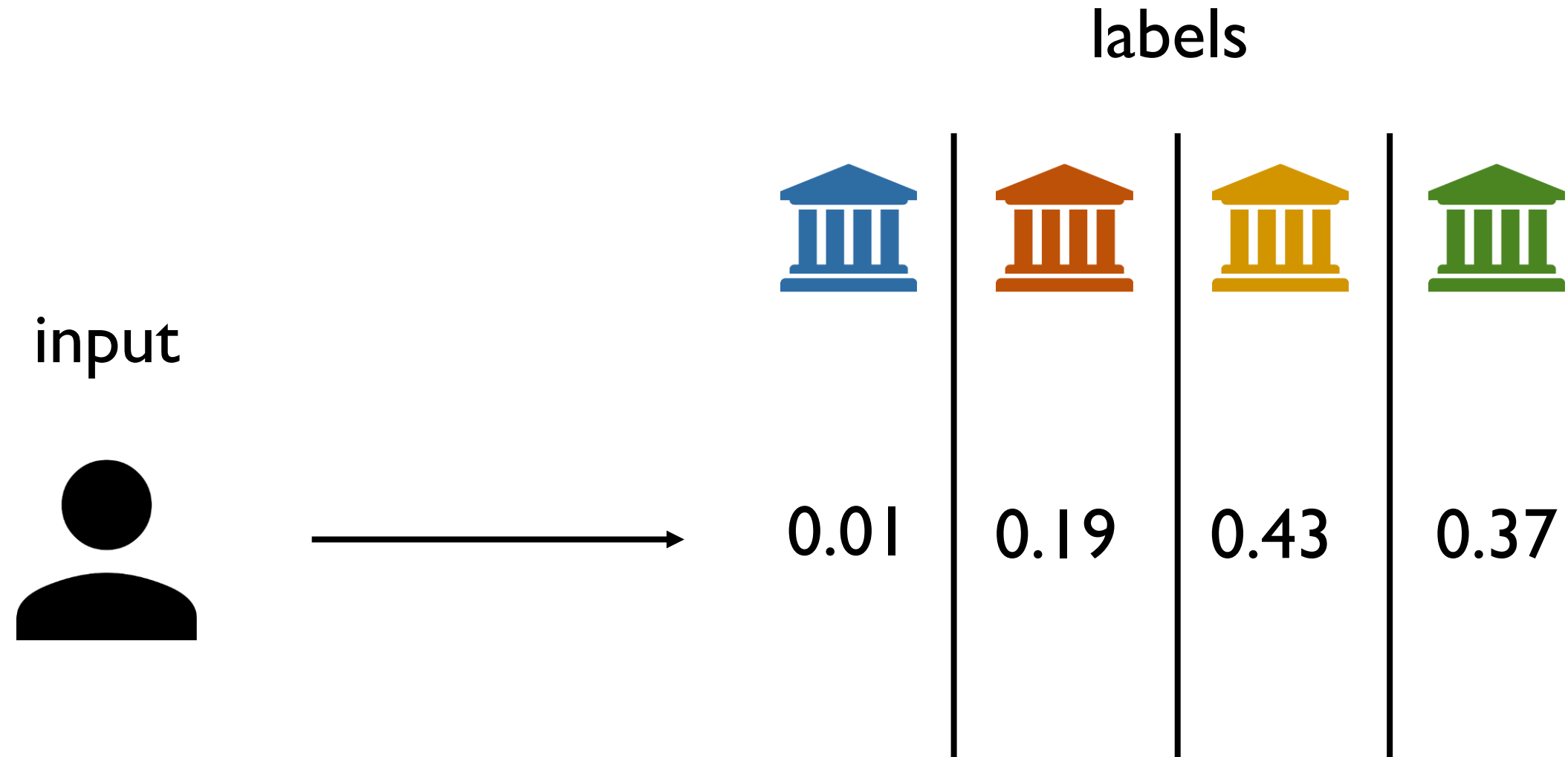
Simulate application sets!

Sample students from past years and apply matching

However: Programs change from year to year

Idea: Learn last year's application behavior

Multi-label probabilistic classifier





Strategy comparison

- (i) All programs
- (ii) Only programs that consistently disadvantage the same demographic group

Error in objective function relative to ideal policies. Smaller values are better.

| | All programs | | | | Consistently unequal | | | |
|-----------------------|--------------|-------------|-------------|-------------|----------------------|-------------|-------------|-------------|
| | Gender | | Income | | Gender | | Income | |
| Strategy | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Historical – 1 year | 0.37 | 1.11 | 0.44 | 1.31 | 1.31 | 1.88 | 1.79 | 1.63 |
| Historical – 3 years | 0.30 | 0.94 | 0.34 | 1.06 | 1.02 | 1.81 | 1.36 | 1.62 |
| Historical – 5 years | 0.32 | 0.99 | 0.33 | 0.98 | 1.22 | 1.20 | 1.63 | 1.57 |
| Predictive – 50 sets | 0.28 | 0.90 | 0.37 | 1.13 | 0.84 | 1.16 | 2.28 | 2.12 |
| Predictive – 200 sets | 0.29 | 0.90 | 0.36 | 1.11 | 0.91 | 1.15 | 2.14 | 1.97 |

Difference in absolute SPD relative to no intervention. Lower values are better.

| | All programs | | | | Consistently unequal | | | |
|-----------------------|----------------|---------------|----------------|---------------|----------------------|---------------|----------------|---------------|
| | Gender | | Income | | Gender | | Income | |
| Strategy | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Historical – 1 year | 0.0014 | 0.0339 | -0.0013 | 0.0383 | 0.0028 | 0.0922 | -0.0774 | 0.1000 |
| Historical – 3 years | -0.0005 | 0.0269 | -0.0034 | 0.0266 | -0.0153 | 0.1069 | 0.0983 | 0.0704 |
| Historical – 5 years | 0.0005 | 0.0234 | -0.0037 | 0.0270 | -0.0037 | 0.0858 | -0.0849 | 0.0827 |
| Predictive – 50 sets | -0.0003 | 0.0124 | -0.0020 | 0.0210 | -0.0156 | 0.0586 | -0.0625 | 0.0822 |
| Predictive – 200 sets | -0.0002 | 0.0126 | -0.0023 | 0.0215 | -0.0135 | 0.0575 | -0.0670 | 0.0817 |

A hand holding a glass sphere over a sunset over the ocean. The sphere reflects the sunset and the ocean. The background is a blurred sunset over the ocean.

Conclusion & future work

- More application sets increase robustness
- Historical approach more practical
- Predictive approach advantageous if historical data is limited
- Future research: influence of policies on application behavior