test

Mind the Gap

When faced with a gap in mean outcomes between two groups, researchers frequently examine how much of the gap can be explained by differences in observable characteristics.

The simple approach is to estimate the pooled regression including an indicator variable for group membership as well as the other observable characteristics, interpreting the coefficient on the group indicator as the unexplained component.

The Oaxaca-Blinder (O-B) decomposition represents an alternative approach

Oaxaca-Blinder decomposition

Consider a categorical (or dummy) variable d that splits our dataset into two groups.

In this case we can run regressions of the form $y = X\beta + \epsilon$ to estimate the the mean difference between groups, as follows

$$\mathbb{E}\left[y^0\right] = \mathbb{E}\left[X^{(0)}\right]\beta_0;\ \text{group } d = 0$$
$$\mathbb{E}\left[y^1\right] = \mathbb{E}\left[X^{(1)}\right]\beta_1;\ \text{group } d = 1$$

Alternatively