The objective of the matching estimator is to compare the effects of treated and untreated observations that have the same *propensity* for being treated, if treatment is random after conditioning on a set of observables. If for every treated observation, we are able to find a similar observation that was not treated, then we will be able to find an estimate of $\tau(x) = \mathbb{E}(Y_i(1) - Y_i(0)|X_i = x)$. If possible, we could potentially compare the outcomes of two similar individual observations, one with treatment and one without; but there may be a way to make use of more information by compositing the effects of different groups. The method of compositing is the distinguishing feature of the weighted and blocked propensity scores presented in lecture. The choice of method may be a matter of style or it could be driven by sparse data, and the need to interpolate the scores. The objective of this section is to review a few of the alternatives.

First, we should probably create the data that will be used in each of the methods. Let D_i be the indicator of treatment for observation $i = 1, 2, \ldots, N$; let Y_i be the outcome variable; and let X_i be the vector of observable characteristics, which affect the propensity for receiving treatment:

$$Y_i = \alpha + \delta D_i + \beta X_i + \epsilon_i, \quad \text{with} \;\; \epsilon_i \sim N(0, 1) \tag{1}$$

Regression adjusting on the propensity score