# Homework 5

**NAME: Your Name**
**DUE DATE: May 3rd, 11:59 pm**

## Problem 1 (GH 9.9) [4 points]

Consider a study with an outcome, $y$, a treatment indicator, $T$, and a single confounding covariate, $x$. Draw a scatterplot of treatment and control observations that demonstrates each of the following:

a) A scenario where the difference in means estimate would not capture the true treatment effect but a regression of $y$ on $x$ and $T$ would yield the correct estimate.

b) A scenario where a linear regression would yield the wrong estimate but a nonlinear regression would yield the correct estimate.

## Problem 2 (Adapted from GH 9.13) [8 points]

The file **congress90.csv** on Canvas has election outcomes and incumbency for U.S. congressional election races in 1990.

a) Estimate the effect of incumbency by fitting a regression of vote90, the Democratic share of the two-party vote in district $i$ in 1990, on vote88 (the outcome in the previous election), inc.stat (the incumbency status in district $i$ in the 1990 election, coded as 1 for Democratic incumbents, 0 for open seats, and -1 for Republican incumbents), and sitting (the incumbent party, coded as 1 if the sitting congressmember is Democrat and 0 if he or she is Republican).

b) Discuss the political interpretation of the estimated coefficients.

c) What assumptions are needed for this regression to give a valid estimate of the causal effect of incumbency? In answering this question, define clearly what is meant by incumbency as a "treatment variable."