Regression II
Vanderbilt University
Prof. Sean P. Corcoran

September 21, 2021

Lecture 5 In-Class Exercise

Part 1 This exercise will replicate the simple difference-in-differences result from chapter 8 of Murnane & Willett. This example comes from Dynarski (2003), who looked at the effects of Social Security survivor's benefits on college enrollment. The dataset used here is from the NLSY and consists of high school seniors in 1979-1983. The file is called *dynarski.dta* and is available via Github:

```
use https://github.com/spcorcor18/LPO-8852/raw/main/data/dynarski.dta, clear
```

1. Do a cross-tabulation of *yearsr* (the year in which the student is a senior) and *offer* (whether the student was a senior in a year in which Social Security survivor's benefits were available). What the "treatment period," and which students are "treated"?

2. Estimate a first difference model for the effect of survivor's benefits by limiting the analysis to the "ever treated," comparing outcomes in the treated and non-treated periods. The outcomes of interest are *coll* (whether the student was enrolled full time in college by age 23) and *hgc23* (the highest grade completed by age 23). You can do this with an OLS regression and be sure to use the sampling weights `weight=[wt88]`. Your results can be compared to Table 8.1 in Murnane & Willett.

3. Now estimate a difference-in-differences model of the effect of survivor's benefits by including the "never treated" group. Again be sure to use the sampling weights. Your results can be compared to Table 8.2 in Murnane & Willett.

4. For practice, plot the mean outcome *coll* by year for the two groups and note whether the parallel trends assumption appears to hold.