

Goals for the week:

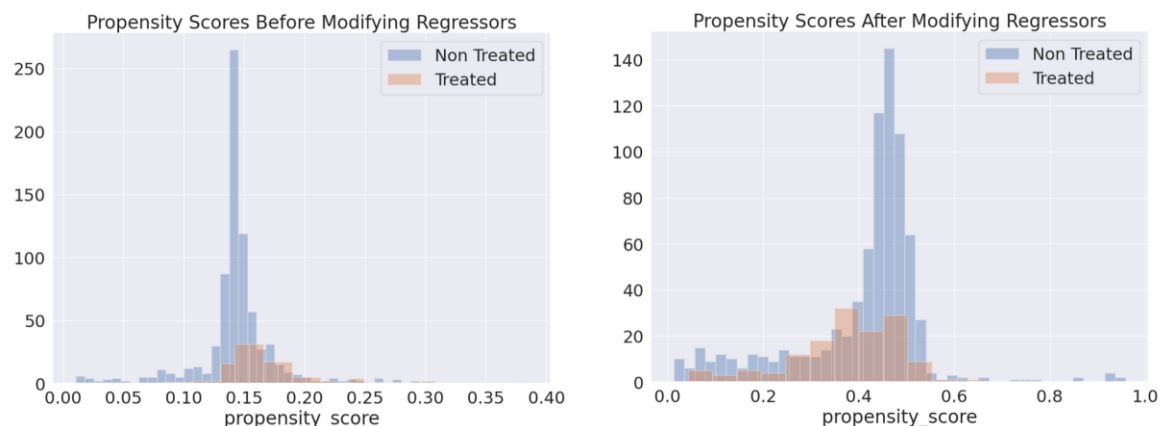
- Create event study model and review methodology for analysis 1

Excluding Dependent Variables for Analysis 1

I experimented with trying out different dependent variables for my propensity score estimations for my first analysis comparing the difference in financial performance between backed and non-backed companies. I did this primarily because I was concerned about how low my propensity scores were in the model that I was previously using. As a result, my new analysis uses the following as dependent variables to regress upon:

```
['Valuation/Ebitda', 'Valuation/EBIT', 'Valuation/Net Income',  
 'Valuation/Revenue', 'Valuation / Cash Flow', 'Deal Size/EBITDA',  
 'Deal Size/EBIT', 'Deal Size/Net Income', 'Deal Size/Revenue',  
 'Deal Size/Cash Flow', 'EBITDA Margin %', 'Years Since Founding']
```

Previously, I was also using Deal Year, Total Invested Equity, Deal Size, and Fiscal Year as dependent variables to regress upon. Shown below is the change in propensity scores after modifying the regressors used.

**What I'd like to focus on this week with Analysis 1**

I would like to investigate how using data regularization on my dependent variables would affect the propensity scores. Data regularization techniques are essentially used as means to select which dependent variables are useful to determine regression estimates.

Creating the event study

For my event study model, I am largely relying on [this guide](#) from the Princeton Library on how to construct one. As with analysis 1, my goal is to see if EBITDA, Net Income, Liquidity, and Return on Equity growth is sustained or statistically significant pre and post IPO on a per company basis.

Steps:

Data Preparation

Company data was already made available given the earlier model that I built, and so what was left now was to split up the financials on a yearly basis based on the years out from IPO. Given the limited timeline, I rely on the interval of years of $-3 < \text{year}_i < 5$ where year_i represents the years since the IPO.

Cleaning the Data and Calculating the Event Window

Cleaning the Data was pretty straightforward. Because I already had methods to extract rows of company financials based on the amount of years from IPO, all I really had to do in this step was to create a new data table for each company ranging from three years before IPO to five years after. The amount of datapoints ranges from company to company.

For the initial concept of the model, I use the company, GoDaddy, as the company to build the model upon.

Estimating Normal Performance

Given the success I had with using the dependent variables in my propensity score estimations in Analysis 1, I decided to use the same regressors to construct the initial estimates of normal performance. Using the Linear Regression extension from the scikit-learn python package, I constructed my regression.

This process is taking longer than expected due to the amount of data manipulation I have to perform in order to get the necessary rows to run the estimations up.

What I will focus on this week with the Event Study

This upcoming week I will be working on calculating abnormal and cumulative abnormal returns, testing for significance, and then abstracting my model to work for all portfolio companies that I have access to.