

MGTF-415 Homework 3: Private Equity

This assignment builds off of Lectures #6 and #7. Students should upload their assignments to Canvas by 11:59pm on November 24. Save your assignment as a .pdf file, and include your code in the same file as your homework assignment. If you are using Python with a Jupyter Script, recall that you can install new modules by inserting the following commands at the beginning of your script:

```
import sys
!{sys.executable} -m pip install your_module_name
```

1 Venture Capital Deal Selection

In this exercise, you will study the probability a venture capital investment ends in an IPO based on (a) the company's industry; and (b) the company's geographic market. As in the lecture, you will perform this analysis using publicly-available data scraped from Crunchbase. You will be provided with already-scraped data as a .csv file, but you are encouraged to replicate part of the web scraping algorithm from lecture on your own. If you are using Python, you will likely need to import the *pandas* module to complete this problem.

- (a) Read the file *scraped-data.csv* as a dataframe. This file contains the output of the web scraping procedure described in lecture.
- (b) As in lecture, create a variable which encodes whether the company's headquarters is in the San Francisco Bay Area, Greater Boston, Greater New York, Greater Seattle, or some other metro area.
- (c) Create a variable which encodes whether the company has experienced an IPO.
- (d) Create a variable which equals one if the company's category contains the substring "Health", which captures whether the company is in the healthcare sector. If you are using Python, this can be accomplished by following a similar method used to produce the *IsPublic* variable in lecture.
- (e) Similarly, create a variable which equals one if the company's category contains the substring "Fin", which captures whether the company is in the financial sector. What are the advantages and disadvantages to using the the substring "Fin" instead of, say, "Financial Services"?

- (f) Filter out companies founded after 2005 or before 2017. If you are using Python, this can be accomplished through the commands

```
D["Year Founded"] = [f[-4:] for f in D["Founded"].values.tolist()]
D["Year Founded"] = np.where(D["Year Founded"].str.contains("20"),
                             D["Year Founded"].values, np.nan).astype(float)
D = D[(D['Year Founded'] >= 2005) & (D['Year Founded'] <= 2017)]
```

- (g) Calculate the share of companies founded between 2005-17 in each of the five metro areas from (b) that have experienced an IPO. If you are using Python, this can be accomplished through the commands

```
print(D.groupby(['City'])['IsPublic'].mean().reset_index())
```

- (h) What share of such companies headquartered in San Francisco experienced an IPO, and how does this value compare with the share for the other four metro areas? In lecture, we considered the total number of IPOs in a metro area, not the share of companies that experience an IPO: which of these two statistics is more valuable for making a venture capital investment?
- (i) What share of companies in the healthcare sector experienced an IPO over 2005-17, and how does this compare with the share of companies from non-healthcare sectors? Answer the same question with respect to companies in the financial sector versus non-financial sectors. If you are using Python, you can calculate these shares through the commands

```
print(D.groupby(['IsHealth'])['IsPublic'].mean().reset_index())
print(D.groupby(['IsFin'])['IsPublic'].mean().reset_index())
```

- (j) One reason companies go public is that they can typically raise more money on public versus private equity markets. How might the result you found in (i) change if venture capital funds started to invest less in less in FinTech companies and more in healthcare companies?

2 Leveraged Buyout Deal Selection

In this exercise, you will study the relationship between a leveraged buyout (LBO) fund's choice of debt and the profitability of the target company it is purchasing. As in the lecture, you will perform

this analysis using data from Preqin, provided as a .csv file. If you are using Python, you will likely need to import the following modules: *pandas* and *matplotlib.pyplot*

- (a) Read the .csv file *preqin-lbo-data.csv* provided on Canvas. What is the unit of observation?
- (b) Create a profit ratio variable and a leverage ratio variable. Use the same definitions as in lecture. Drop observations with a leverage ratio above 100%.
- (c) Partition the data into 30 bins based on profit ratio. Create a variable which equals the bin associated with each observation. If you are using Python, use the same code as in lecture but change the number of bins from 15 to 30.
- (d) Produce a binned scatterplot of the relationship between a fund's leverage ratio and the profit ratio of its target company.
- (e) List one reason why an LBO fund might want to borrow more to purchase an unprofitable target company.