

Research Check-in

Leonardo Straccia
11/20/2023

Goals for the week:

- Review Feedback and implement changes into next draft for review
- Further modularizing the code so that others may copy my results
- Eliminating any potential multicollinearities in the model

As shown below in the summary table, the initial predictors that I have used to calculate propensity scores could potentially have some multicollinearity issues in the model. For example, I use columns 'Revenue' and 'Valuation/Revenue' as two of the predictors that I use to calculate propensity scores. As we can see in the covariance matrix, this doesn't seem to be too much of an issue given the fact that on initial view, none of the variables are perfectly collinear.

Of note as well in the table below is the fact that some of the 50th percentile variable data is equivalent to the 75th percentile data. This was because many of the variables used in the deal table with IPO information had several missing datapoints. To account for this, I filled in missing datapoints with median values. Upon inspecting the summary table it has become evident that some of these columns are missing so much data that it is practically infeasible to use them as accurate predictors.

My solution to this, which I will have ready in a few days time, is to exclusively rely on company financial information from the timeseries table that I also have access to and use that information to perform the logistic regression on.

Summary Table of predictor variables used in the propensity score regression

	count	mean	std	min	25%	50%	75%	max
Revenue	953.0	1622.255467	1.155145e+04	-1.000000e-02	11.21	95.570	420.57	201487.00
Gross Profit	953.0	423.973127	3.685990e+03	-3.331200e+02	62.33	91.970	141.54	107400.00
Net Income	953.0	245.856747	6.858894e+03	-3.969000e+03	-23.05	-2.790	11.26	211400.00
EBITDA	953.0	107.483033	5.984562e+02	-3.129000e+03	-12.88	1.200	58.28	10539.00
EBIT	953.0	66.582802	5.646882e+02	-3.129000e+03	-17.88	-0.790	36.57	9742.00
Total Debt	953.0	828.742581	6.352034e+03	0.000000e+00	1.88	20.000	275.00	145670.00
Valuation/Ebitda	953.0	113.281920	5.022078e+03	-6.000000e+04	-11.82	-4.600	6.73	139081.50
Valuation/EBIT	953.0	-177.679916	4.210747e+03	-1.144972e+05	-14.69	-6.995	7.63	9485.89
Valuation/Net Income	953.0	508.281186	1.826203e+04	-1.144972e+05	-17.37	-8.870	2.81	551156.00
Valuation/Revenue	953.0	-371593.124292	1.030262e+07	-3.156963e+08	3.25	4.600	6.55	446446.00
Valuation / Cash Flow	953.0	483.662823	1.671728e+04	-8.708300e+04	-11.30	12.390	34.33	501842.00
EBITDA Margin %	953.0	-487.575792	4.426103e+03	-8.378402e+04	-8.43	0.000	12.70	1406.00
Employees	953.0	782.421826	5.101035e+03	3.000000e+00	275.00	275.000	275.00	93000.00
Years Since Founding	953.0	21.464848	3.269173e+01	0.000000e+00	5.00	10.000	19.00	191.00

Covariance Matrix of predictor variables

	Revenue	Gross Profit	Net Income	EBITDA	EBIT	...	Valuation/Ebitda	Valuation/EBIT	Valuation/Net Income	Valuation/Revenue
Revenue	1.334361e+08	9.384453e+06	3.756837e+06	2.838350e+06	2.233203e+06	...	-1.707869e+05	2.772261e+05	-7.265106e+05	6.034651e+08
Gross Profit	9.384453e+06	1.358652e+07	2.463362e+05	6.646144e+05	4.467949e+05	...	-4.386152e+04	5.886197e+04	-1.625285e+05	1.234982e+08
Net Income	3.756837e+06	2.463362e+05	4.704443e+07	1.805536e+05	1.632355e+05	...	-2.949549e+04	4.261536e+04	-1.191510e+05	1.007039e+08
EBITDA	2.838350e+06	6.646144e+05	1.805536e+05	3.581498e+05	2.811920e+05	...	-1.146574e+04	1.953975e+04	-3.099142e+04	4.085379e+07
EBIT	2.233203e+06	4.467949e+05	1.632355e+05	2.811920e+05	3.188727e+05	...	-7.265976e+03	1.475043e+04	-1.092121e+04	3.388911e+07
Total Debt	7.388893e+06	8.715071e+05	1.219026e+07	5.153300e+05	3.735128e+05	...	-9.458541e+04	1.457274e+05	-3.340506e+05	3.016509e+08
Valuation/Ebitda	-1.707869e+05	-4.386152e+04	-2.949549e+04	-1.146574e+04	-7.265976e+03	...	2.522127e+07	3.611236e+06	-4.362541e+04	4.655747e+07
Valuation/EBIT	2.772261e+05	5.886197e+04	4.261536e+04	1.953975e+04	1.475043e+04	...	3.611236e+06	1.773039e+07	1.390835e+07	-6.180117e+07
Valuation/Net Income	-7.265106e+05	-1.625285e+05	-1.191510e+05	-3.099142e+04	-1.092121e+04	...	-4.362541e+04	1.390835e+07	3.335016e+08	1.932467e+08
Valuation/Revenue	6.034651e+08	1.234982e+08	1.007039e+08	4.085379e+07	3.388911e+07	...	4.655747e+07	-6.180117e+07	1.932467e+08	1.061441e+14
Valuation / Cash	1.164774e+06	-1.219959e+05	-1.192249e+05	-6.281453e+04	-4.293408e+04	...	-6.555782e+04	-1.881250e+05	-5.021954e+05	1.755910e+08

Questions to address during the check-in

- In the notebook I provided, I tried to modularize my code to make the logic easier to understand for others to review and replicate my results. Would it be helpful to include a section for each of the analyses regarding the different functions that I define and what each of them is supposed to do?
- My literary review is heavily tailored toward how I can add to the existing dialogue of private equity outcomes through my first analysis. Should I update my review to further include existing literature relating to the event study (analysis 2) as well?
- Would it be helpful to go into more detail regarding the covariates (X's) that I use for each one of the models I build in the paper?
- Do you have any time this week to meet? I have a few more questions I would like to ask you in person so that I can turn around another draft for your review quickly while you review the first.
- How will you review the paper once it is due? How will I be graded?

What I will focus on this upcoming week: Finalizing the conclusion / formatting and implementing advisor feedback into the paper