Advanced Corporate Finance - Assignment 2
Miguel A. Troppmair
Pedro Watuhã G. dos Santos

# 1 Introduction

This exercise investigates the financial characteristics and constraints of firms, employing a statistical analysis to understand their investment behavior. We utilise Capital IQ database and begin by presenting summary statistics of key financial variables, including leverage, cashflow, tangibility, size, Q, investment, and Z-score (Altman's Unleveraged Z-score).

The data reveals significant variability among these variables, with cashflow and Q showing the greatest dispersion. To address the presence of outliers, we apply a winsorization technique to trim extreme values, resulting in more stable statistical summaries. Our analysis is further supported by visual representations, including boxplots and heatmaps, which highlight the distribution and correlation among the variables.

We then replicate the seminal study by Fazzari, Hubbard, and Petersen (1988) on financial constraints, focusing on the relationship between investment and cashflow. Our regression analysis confirms that cashflow is a significant predictor of investment, indicating that firms may be financially constrained and reliant on internal funds for investment. Additionally, the positive and significant relationship between Q and investment suggests that firms invest more when they perceive their projects to be of higher value.

To test the robustness of our findings, we conduct a randomization check by assigning a treatment status to a subset of firms and re-estimating the regression model. The results indicate that our estimates are robust to this randomization process. We also utilize matching estimators to control for potential confounding factors, further validating our conclusions.

Overall, our findings contribute to the understanding of financial constraints and investment behavior in firms, providing insights into how internal cashflow and market valuations influence corporate investment decisions.

# 2 Empirical Results

## 2.1 Baseline Study

| | Mean | Standard Deviation | Min | Max | Count |
|---|---|---|---|---|---|
| Leverage | .3091255 | .1937526 | .0003163 | 2.892473 | 3774 |
| Cashflow | .5209227 | 3.283656 | -124.4266 | 30.14598 | 2606 |
| Tangibility | .4011973 | .2642268 | .0013629 | .957305 | 3774 |
| Size | 6.444676 | 2.02135 | .2851789 | 12.35342 | 3774 |
| Q | 1.91533 | 2.466815 | -48.19234 | 81.28764 | 3774 |
| Investment | .2397947 | .3692122 | -4.397313 | 6.5625 | 2606 |
| Z | 1.548434 | 1.84737 | -33.10587 | 20.04471 | 3774 |

Table 1: Summary of Statistics

The above table present statistical summaries of financial variables. Leverage has a mean of 0.31, and a std. deviation of 0.19. This indicates that the mean company relies little on debt, financing its activities with equity. Cashflow exhibits more variability than leverage, with a mean of 0.52, a standard deviation of 3.28, and a wide range (from -124 to 30). The variable Q is also very volatile and has multiple outliers. It has a mean of 1.91, which indicates that investors perceive the company's investments in assets to be productive. It also suggests that the market values the company's assets more highly than their current book value (or replacement cost). Investiment, which represent the amount spend in Capex over the fixed assets (discounting sales of property). The remaning variables are less volatile, comparing the ratio of standard deviation over the mean. Finally, Figure 2 complements our investigation by stating the low correlation among the variables that we will be focusing.

One concern that emerges from the data as seem in Figure 1 is the presence of outliers. The variables of investment, Q, Z, and cashflow present some considerable variation and thus are analysed after a winsoring in the $5^{th}$ and $95^{th}$ percentile for investment and cashflow, and in the $1^{st}$ and $99^{th}$ percentile for the others. We then utilise the statistics shown on table 2 for the rest of the analysis.

Looking at the second table, with winsoring, we can see a reduction of standard deviation for all the changed variables, as expected. Cashflow, for instance, has the largest decrease in standard deviation and shows a narrower distribution with a similar mean of 0.47. Q, Investment and Z show similar means, standard deviations, and ranges compared to their raw counterparts, suggesting little impact of outliers.

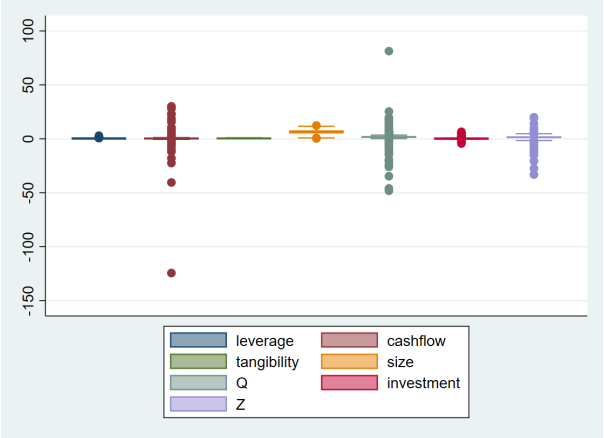


Figure 1: Boxplot of the variables

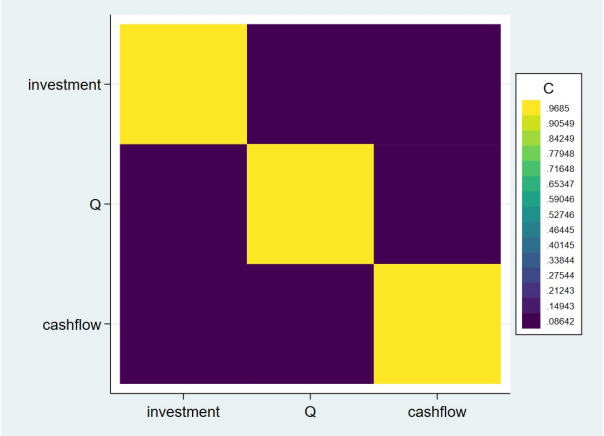


Figure 2: Heatmap for the Correlations

| | Mean | Standard Deviation | Min | Max | Count |
|---|---|---|---|---|---|
| Leverage | .3091255 | .1937526 | .0003163 | 2.892473 | 3774 |
| Cashflow | .4701951 | .6306156 | -.0859031 | 2.472222 | 2606 |
| Tangibility | .4011973 | .2642268 | .0013629 | .957305 | 3774 |
| Size | 6.444676 | 2.02135 | .2851789 | 12.35342 | 3774 |
| Q | 1.94814 | .9996803 | .1787255 | 6.710635 | 3774 |
| Investment | .2202766 | .1794362 | .0205 | .6893524 | 2606 |
| Z | 1.579032 | 1.31783 | -2.319769 | 5.969377 | 3774 |

Table 2: Summary of Winsored Statistics

In sequence, we replicate (Fazzari, Hubbard, and Petersenís, 1989) study on Financial Constraints by estimating the following equation:

$$Investment_{i,j} = \beta_0 + \beta_1 Q_{i,t} + \beta_2 CashFlow_{i,t} + \sum_i firm_i + \sum_t year_t + \epsilon_{i,t} \quad (1)$$

We seek to analyse if firms are financially constrained by testing if $\beta_2 > 0$. This test allow us to observe if a firm is utilising its own CashFlow to finance its investment instead of relying in outside options, such as debt.

In the absence of financial constraints, firms should invest in all projects where the expected return exceeds the cost of capital. Therefore, investment should be positively related to Q. This is captured by the parameter $\beta_1$.

Table 3: Regression Results

| | Standart | Fixed Effects | Clustered FE |
|---|---|---|---|
| | (1) | (2) | (3) |
| | Investment | Investment | Investment |
| Cashflow | 0.130*** | 0.108*** | 0.108*** |
| | (0.00504) | (0.00862) | (0.0172) |
| Q | 0.0279*** | 0.0233*** | 0.0233** |
| | (0.00341) | (0.00512) | (0.00851) |
| $N$ | 2606 | 2606 | 2606 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As one can see in the first column of Table 3, without FE nor clustering, Cashflow is highly significant, with a point estimate of 0.13. Adding FE and clustering, we have significant results as well, with an estimate of 0.108. In all specifications, this indicate that are indeed financially constrained and the presence of Cashflow seems to be essential to describe the investment process.

Looking now to $\beta_1$, in all three settings the estimates are positive and significant, confirming the prior that higher Q leads to higher investiment.

## 2.2 Matching Estimator

After this analysis, we randomly select a subsample of 1000 firms and attribute to them the status of treated. This is done to test the robustness of our estimation by selecting a random group, attributing to them the characteristic of the random control group through

a matching estimator and rerunning the original regression. If the estimates are indeed robust, we would expect small differences when comparing them to this check.

Table 4 shows the results of a random attribution to treated status. As observed, nor the control nor the treated group seem to be significantly different since most of the outliers were already cut from the sample through the winsoring process. This process, normally not observed in a natural experiment, just reafirms the randomization process of a substantial subsample from the data.

| | Investment | Z | Tangibility | Leverage | Size | Cashflow | Q |
|---|---|---|---|---|---|---|---|
| Treated | 0.2215 | 1.5742 | 0.4028 | 0.3113 | 6.4718 | 0.4663 | 1.9335 |
| Untreated | 0.2198 | 1.5923 | 0.3965 | 0.3030 | 6.3691 | 0.4808 | 1.9886 |
| Difference | 0.017 | -0.0181 | 0.0063 | 0.083 | 0.1027 | -0.0145 | -0.0551 |
| P-Value | 0.4748 | 0.7069 | 0.5173 | 0.2647 | 0.1683 | 0.6031 | 0.1436 |

Table 4: Average Treatment Effects before Matching

After observing the results, we utilise the matching estimator to build the matched group. The critearia used for the match was a perfect coincidence for the sector the firm works and a single nearest-neighbour proximity matching based on Firm Size, Leverage, Altman's Z-score and Tangibility under the Abadie-Imbens (2002) procedure. When doing the match, we observed a sample average treatment effect for the subpopulation of the treated (SATT) of 0.01 with a P-value of 0.572, indicating that the effect of treatment is not significant as expected due to the randomness. The results are presented on Table 5 and most of the differences were mitigated.

| | Investment | Z | Tangibility | Leverage | Size |
|---|---|---|---|---|---|
| Treated | 0.2460 | 1.6267 | 0.4063 | 0.2959 | 6.4757 |
| Matched | 0.2335 | 1.6711 | 0.4099 | 0.2955 | 6.5087 |
| Difference | 0.0125 | -0.0444 | -0.0036 | 0.0004 | -0.033 |
| P-Value | 0.5832 | 0.5973 | 0.7956 | 0.9773 | 0.7580 |

Table 5: Average Treatment Effects after Matching M=1

Comparing the matching results when no FE are added and there is not clustering to the first OLS we can similar results in terms of significance and magnitude, both coefficients are positive. However, adding fixed effects decreases by about halt the point estimate of Cashflow (why?). Using clustered errors also changes the results: both coefficients are no longer significant, due to an increase in the standard error. This maybe due to high correlaction among individuals in the same group cluster. This lack of robustness for estimates in different specifications can be seen in table 6 below.

In the results presented in the table 7 bellow, we changed the treatment group by the control group, but since there were no individuals in the treatment group in the regression, the results are identical to those present in table 6.

Table 6: Matching M=1 Regression Results

| | Standart | Fixed Effects | Clustered FE |
|---|---|---|---|
| | (1) | (2) | (3) |
| | Investment | Investment | Investment |
| Cashflow | 0.0971*** | 0.0532*** | 0.0532 |
| | (0.00658) | (0.0153) | (0.0344) |
| Q | 0.0297*** | 0.0360*** | 0.0360 |
| | (0.00429) | (0.00837) | (0.0189) |
| $N$ | 1246 | 1246 | 1246 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Swapped M=1 Regression Results

| | Standart | Fixed Effects |
|---|---|---|
| | (1) | (2) |
| | investment_w | investment_w |
| cashflow_w | 0.0971*** | 0.0532 |
| | (0.00658) | (0.0344) |
| Q_w | 0.0297*** | 0.0360 |
| | (0.00429) | (0.0189) |
| $N$ | 1246 | 1246 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Analizing the match with $m = 4$, we see a SATT of 0.03 with a P-value of 0.8414, which again is not significant. The SATT for the remainder variables is presented in table 8. All the differences are not statistically significant.

Repeating the exercise of matching but now with $m = 4$, i.e., matching each firm to the 4 closest to it, we have distinct results. Although the first column is similar to the OLS (as was the first column of the matching with $m = 1$), the other two columns, FE and clustering, have oppositive Cashflow sign, and is not significant when we cluster the errors. Q however is in line with the previous estimations. The results are presented in table 9.

Now we change the treatment for the control group with $m = 4$. This time however, we can see the estimates are not identical, although extremily similar. This is due to some observations being dropped (from 1238 in table 9 to 1140 in table 10).

| | Investment | Z | Tangibility | Leverage | Size |
|---|---|---|---|---|---|
| Treated | 0.2460 | 1.6267 | 0.4063 | 0.2959 | 6.4757 |
| Matched | 0.2430 | 1.6668 | 0.4087 | 0.2924 | 6.5096 |
| Difference | 0.03 | -0.0401 | -0.024 | 0.0035 | -0.0339 |
| P-Value | 0.8414 | 0.6127 | 0.8614 | 0.7098 | 0.7475 |

Table 8: Average Treatment Effects after Matching M=4

Table 9: Matching M=4 Regression Results

| | Standart | Fixed Effects | Clustered FE |
|---|---|---|---|
| | (1) | (2) | (3) |
| | investment_w_mean | investment_w_mean | investment_w_mean |
| cashflow_w | 0.102*** | -0.0410* | -0.0410 |
| | (0.00825) | (0.0184) | (0.0707) |
| Q_w | 0.0313*** | 0.0470*** | 0.0470** |
| | (0.00486) | (0.00847) | (0.0147) |
| $N$ | 1238 | 1238 | 1238 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Swapped M=4 Regression Results

| | Standart | Fixed Effects |
|---|---|---|
| | (1) | (2) |
| | investment_w_mean | investment_w_mean |
| cashflow_w | 0.109*** | 0.0410 |
| | (0.00842) | (0.0679) |
| Q_w | 0.0395*** | 0.0470*** |
| | (0.00474) | (0.0141) |
| $N$ | 1140 | 1140 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# 3 Final Comments

For all specification, when the regression is done with clustering or FE, the coefficients estimated are about the same (around 0.10 for Cashflow, and 0.30 for Q). Nevertheless, the matching estimator presented quite distinct results when we added FE and clustered the errors: the point estimates for Cashflow decreased and even became negative, which would lead one to think that firms are not financially contrained. It is difficult to provide an economic explation of a negative $\beta_2$. This would imply that a higher Cashflow would lead to lower levels of investiment which sounds unreasonable.

The large difference in results for different specifications is a sign that perhaps the causality between these variables is not so easy to estimate. The lack of robustness also puts in check a simple relationship between Cashflow, Investment and financial contraint.