FIN 971: Corporate finance
Problem set 2

This problem set asks you to test the Q-theory of investment and estimate investment-cash flow sensitivities using multivariate linear regressions on unbalanced panel data from Compustat. Using data beginning in 1980, run the following panel regressions:

$$Investment_{i,t} = \beta_0 + \beta_1 Q_{i,t-1} + u_{i,t} \tag{1}$$

$$Investment_{i,t} = \beta_0 + \beta_1 Q_{i,t-1} + \beta_2 CashFlow_{i,t} + u_{i,t} \tag{2}$$

where investment is the ratio of capital expenditures to lagged PP&E, and Q is the market to book value of assets, and cash flow is the ratio of cash flow to lagged book assets. Winsorize the data at the 99/1 percentiles. Filter the data as described at the end of this document. Use the variable definitions given by Erickson and Whited (2012) in their appendix, page 1325. Use the definitions given for their *second* set of regressions. Keep in mind that some of these are not canonical in that there are alternative ways to reasonably measure some of these objects. Run the panel regressions both with and without the inclusion of firm fixed effects.

1. Estimate the models and interpret your results. How do they compare to Fazzari, Hubbard, and Petersen (1988) (FHP)?

2. Rerun using lagged cash flow. How do the results change?

3. Run regression (1) with the addition of a quadratic term for Q. Interpret your findings. Is there evidence of a functional form mispecification?

4. Rerun the regressions using standard errors which are robust to heteroskedasticity. How does this change the standard errors and inference from part 1?

5. Using regression specification (2) and heteroskedasticity-robust standard errors, test the null hypothesis that the two slope coefficients are equal.

6. Sort the firm-year observations into quintiles based on book leverage. Do the sort conditional on year such that there will be an (roughly) equal number of observations in each bin across years.[1] Run specification (2) on each of the quintiles. How do the cash flow coefficients compare across leverage bins? What are potential reasons for this?

7. Repeat part 6 using the financial constraints index of Hadlock and Pierce (2010).[2] For this exercise, sort unconditionally (not within each year). This means the number of observations in each bin will vary across years. Compare the pattern of the coefficient estimates with that from part 6 and FHP. What can we learn from this exercise?

8. Run the regressions controlling for measurement error in Q using the high order cumulants method of Erickson and Whited. The Stata library created by the authors will be useful: `http://toni.marginalq.com/ewestimators.html`. How do the estimates differ from part 4?

9. Estimate Eq. (1) using the Total Q measure from Peters and Taylor (2017).[3] How does this compare (both coefficients and R-square) to the baseline results? Does total Q work better to explain investment than traditional Q?

Data construction

Use Compustat Annual data. Use the following filters (each observation must satisfy all of the following criteria):

1. `(indfmt == "INDL" & datafmt == "STD" & popsrc == "D" & consol == "C")`

   These conditions ensure that gvkey-fyear uniquely identify each observation.

---

[1]Stata package `egenmore` will be useful for this because it allows you to construct quintiles within a `by` group (combine `xtile` with `by`), e.g. `egen lev_bin = xtile(book_lev), by (year) nq(5)`.

[2]See page 1929 of the paper. Be sure to properly deflate size to real 2004 dollars.

[3]Note that WRDS provides this measure under the dataset "Peters and Taylor Total Q".

2. (`at>0`)

   Positive book assets.

3. (`sic < 6000 | sic > 6999`)

   No financial firms.

4. (`sic < 4900 | sic > 4999`)

   No utilities.

5. (`gvkey != "" & fyear != .`)

   No missing identifiers.

6. (`fic == "USA"`)

   Only US firms. Note that some papers use US and Canada; others use only US.

7. (`prcc_f != .  & csho != .`)

   Drop missing stock market data. Note that the CRSP data may be more accurate than Compustat for stock price data. It's a good idea to use CRSP when possible, but many in the corporate finance papers use the Compustat to construct market cap. You can use the CRSP-Compustat merged database to combine the two: the matching has been done.

References

Erickson, T., Whited, T. M., 2012. Treating measurement error in tobin's q. The Review of Financial Studies 25, 1286–1329.

Fazzari, S. M., Hubbard, R. G., Petersen, B. C., 1988. Financing constraints and corporate investment. Brookings papers on economic activity 1988, 141–206.

Peters, R. H., Taylor, L. A., 2017. Intangible capital and the investment-q relation. Journal of Financial Economics 123, 251–272.