# Finance II: Assignment 2 Report

Giovanni Villa Salas

The figures for the paper replication are shown at the end of this report.

In first place, I will answer the questions and then explain my results. The questions were:

* What is the economic model underlying the relationship between passive ownership and governance?
  + There is many underlying economic models behind the relationship between passive ownership and governance. One can think of the idea of screening, where managers of passive funds invest in certain types of firms (less risky, including corporate governance under that risk measure) while managers of active funds invest in more risky portfolios. Another economic model that links both topics is the monitoring/incentives of mutual funds managers while using their seats on the board of the companies that they invest in. Are they representing the incentives of the mutual funds clients or the interests of the firm.
* In the specific setting of this paper, what is the endogeneity concern that the authors face? How might this concern bias the estimation in a setting that is not endogeneity-proof? You can address this issue using formulas as in basic econometrics textbooks. Explain what are the two conditions that a valid IV should satisfy and how the authors propose to deal with these two conditions.
  + In this specific paper, the main challenge that authors face is the problem that passive ownership can affect corporate governance through omitted/unobserved variables, such as the access to capital of the firm or even the ownership of active investors.   
    To deal with this problem, the authors proposed an IV model, using the threshold between the Russell 2000 and Russell 1000 as instrument to account for passive ownership. The relevance assumption is given by the fact that by jumping from the lower tail of the Russell 1000 to the higher tail of the Russell 2000, you are effectively receiving a bigger amount of passive funds through the change from low weight to high weight in the new index. The exogeneity assumption comes from the fact that there shouldn’t be a big change in corporate governance after dropping the Russell 2000.
* What are the economic rationale behind using the inclusion into Russell-1000 or Russell 2000 indices as one of the determinants of ownership structure of firms?
  + As we discussed before, there are many reasons that could link corporate governance with passive investment (going from 1 index to another is really relevant in terms of investment received). In this case, using Russell 2000/1000 affects directly the amount of passive investment received, thus affecting Passive investment, which makes it a perfect candidate to be an instrument.
* How do Table 2 and Table 3 help to justify this rationale?
  + As we can see on table 2, it provides evidence that being a member of Russell 2000 affects the ownership structure of the firm, having more passive owners between its investors. Table 3 on its side, provide evidence that ownership structure affects corporate governance, justifying the previous rationale.
* How does the econometric setup in this paper compare to other related recent work? Focus, in particular on the following two papers: (i) Crane, Michenaud, and Weston, “The effect of institutional ownership on payout policy: Evidence from index thresholds,” 2016, Review of Financial Studies 29(6), 1377-1408, and (ii) Schmidt and Fahlenbrach, “Do exogeneous changes in passive institutional ownership affect corporate governance and firm value?,” 2017, Journal of Financial Economics 124, 285-306. Specifically,
  + a. What other datasets do they use? What are the advantages of these datasets?

Both papers exploit the same Russell 1000/2000 discontinuity, relying on the Russell reconstitution. The advantages of this method is a neat, and clever way to identify changes in ownership structure and any problem related to this phenomenon.

* + b. Define, compare, and discuss the different empirical approaches used in these papers. What are the potential weaknesses in the empirical designs?

The three papers do the same Russell- 2000/1000 decomposition, following the IV method to perform the identification strategy.

The main potential weakness in my opinion is the relevance of the estimator. In my opinion, the instrument will not be strong for too many variables. There is also the idea that investors know that a certain firm could cross the threshold with anticipation, thus it can have some relations with unobserved variables.

Now I will go through my replication results.

For the table 1, we have:

**Summary statistics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | N | Mean | Median | SD |
| MF OWN | 4325 | 33.031 | 30.490 | 18.861 |
| PF OWN | 4325 | 2.557 | 2.296 | 1.717 |
| AF OWN | 4325 | 25.176 | 23.075 | 15.283 |
| U OWN | 4325 | 5.757 | 4.400 | 4.913 |
| indpt pct | 2696 | .651 | 0.667 | .174 |
| ppill rem | 3097 | .013 | 0.000 | .112 |
| ab sm | 3097 | .005 | 0.000 | .072 |
| DUALCLASS | 3097 | .118 | 0.000 | .322 |
| ROA W | 4324 | .04 | 0.042 | .089 |

Which is close to the original table presented in the paper. In particular, there is approximately 100 observations that I didn’t match with the paper. Now, the big difference in my replication, that will cause problems onwards, is the amount of links that I found in my replication. In particular, I got a 33% of total mutual fund shares holdings while the paper just found 25%. That is mainly explained by the active and unobserved funds, and thus, by construction of the variables, to our Passive Fund definition in the replication.

In general, the results that I got were really sensitive to changes in the way to classify the funds names (using algorithms such as matchit had different results than the approach that I followed (complete word in name).

Another problem that is not in the table was with the market value timing in the paper and the prices used to generate the Market Cap/Value variables. Depending on the date used (and source), the variables were really different from each other.

Other problem that I faced was that the MKT Cap variable in the algorithm was so big in the 3rd order polynomial that was collinear and thus eliminated from the regression, as we can see in the table 7.

As a summary for the tables 2 to 7, the magnitudes generally don’t match the ones on the paper, but they have a decent performance in the sense that they have the same sign and closer if not best significance.

The results are presented below:

Table 2:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) |
| VARIABLES | All Mutual Funds | All Mutual Funds | Passive | Active | Unclassified |
|  |  |  |  |  |  |
| top250 | 4.532\*\* | 4.561\*\* | 0.633\*\* | 3.253\*\* | 0.882\*\* |
|  | (0.689) | (0.690) | (0.048) | (0.582) | (0.183) |
|  |  |  |  |  |  |
| R-squared | 0.272 | 0.270 | 0.569 | 0.225 | 0.229 |

Robust standard errors in parentheses

\*\* p<0.01, \* p<0.10

Table 3:

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
| VARIABLES | Polynomial N=1 | Polynomial N=2 | Polynomial N=3 |
|  |  |  |  |
| top250 | 0.414\*\* | 0.409\*\* | 0.368\*\* |
|  | (0.028) | (0.029) | (0.028) |
|  |  |  |  |
| R-squared | 0.562 | 0.563 | 0.569 |

Robust standard errors in parentheses

\*\* p<0.01, \* p<0.10

Table 4:

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
| VARIABLES | Polynomial N=1 | Polynomial N=2 | Polynomial N=3 |
|  |  |  |  |
| pas | 0.112\*\* | 0.112\*\* | 0.104\*\* |
|  | (0.028) | (0.028) | (0.030) |
|  |  |  |  |
|  |  |  |  |
| R-squared | 0.106 | 0.105 | 0.120 |
|  |  |  |  |

Robust standard errors in parentheses

\*\* p<0.01, \* p<0.10

Table 5:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| VARIABLES | Polynomial N=1 | Polynomial N=2 | Polynomial N=3 | Polynomial N=1 | Polynomial N=2 | Polynomial N=3 |
|  |  |  |  |  |  |  |
| pas | 0.177\*\* | 0.175\*\* | 0.161\* | 0.088\*\* | 0.079\*\* | 0.079\*\* |
|  | (0.061) | (0.061) | (0.063) | (0.026) | (0.025) | (0.025) |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| R-squared | 0.070 | 0.072 | 0.091 | 0.014 | 0.023 | 0.023 |

Robust standard errors in parentheses

\*\* p<0.01, \* p<0.10

Table 6:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | | (2) | | (3) | | (4) | | (5) | | (6) | |
| VARIABLES | Polynomial N=1 | | Polynomial N=2 | | Polynomial N=3 | | Polynomial N=1 | | Polynomial N=2 | | Polynomial N=3 | |
|  |  | |  | |  | |  | |  | |  | |
| pas | 0.023\* | | 0.024\* | | 0.011 | | 0.017\* | | 0.016\* | | 0.017\* | |
|  | (0.011) | | (0.011) | | (0.011) | | (0.007) | | (0.007) | | (0.007) | |
|  | |  | |  | |  | |  | |  | |

Table 7:

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
| VARIABLES | Polynomial N=1 | Polynomial N=2 | Polynomial N=3 |
|  |  |  |  |
| pas | -0.093\* | -0.092\* | -0.062 |
|  | (0.054) | (0.055) | (0.060) |
|  |  |  |  |
| R-squared | 0.058 | 0.058 | 0.076 |

Robust standard errors in parentheses

\*\* p<0.01, \* p<0.10