Financial econometrics project 2013000215 shin dong chan

Q) Does the presence of a female independent director have an impact on CEO compensation, particularly with respect to pay-performance sensitivity?

--------------------------------------------------------------------------------------------------------------

A1) I have a purpose to confirm that whether the CEO compensation & pay performance sensitivity relation is different or not by existence of female director.

In other words, I want to see that the performance of a firm is well reflected in the pay (CEO compensation) by female independent director are present or not. So, I should choose a measure it can determine firm’s performance.

I choose a measure that can check performance of CEO. It is ROA and 1 YEAR STOCK RETURN(RET1Y). First. ROA is financial factors that can tells return of firms divided by assets of a firm. In generally, return of firms is affected by size of firms so ROA is a good measure for scaling objective performance of a firm.

But scaling performance a firm using only ROA measure is weak by this claim. “Because of America economics be prosperous in 2006 so that most of firms earns much money no matter of firm performance?”

So, we need another variable that can measure correct firm performance no matter of economics circumstances in America. That is 1 YEAR STOCK RETURN. This variable is stock return of a firm minus benchmark return in this year. So, it can measure performance well follow benchmark growth.

I established a hypothesis for solving this question

HYPOTHESIS : the firm that having a female director is significantly influence of sensitivity of CEO compensation by changing ROA and RET1Y.

A2) To check HYPOTHESIS that established on A1, we use linear regression.

To try linear regression, population should follow linear relation of control variables. But if I use REALCEOCOMP in dependent variable, it can make exponential relationship. So, I use log(CEOCOMP) instead REALCEOCOMP.

To check hypothesis, I should put in control variables ROA,RET1Y and FEMALEID. But these 3 variables cannot control overall relationship of log(CEOCOMP), so it can be omitted value bias problem. To prevent this problem, we should put in another control variables that it has a relationship of log(CEOCOMP).

But it can be make opposite effect if I put in all of control variables in data. Regression coefficient estimator to be directly proportional to variation inflation factor(VIF) and variation inflation factor will continuously increase whenever put in any control variables. So, I should have proper criteria in variable selection.

1] removed variables

- GVKEY : I couldn’t find a special pattern of GVKEY. So, I remove this variable.

- COMPNAME : the COMPNAME itself don’t have any relationship of CEO compensation.

- CEOAGE : generally, CEOAGE are a bit correlated of CEOTENURE, but we have CEOTENURE data. So, I don’t have necessity need to put in this variable.

- CEOGENDER : the proportion of firms that CEOGENDER is 1 is lower than 3% and CEOGENDER is irrelevant object of this experiment.

- BOARDSIZE : I didn’t put in this variable in regression model but I divided data into big BOARDSIZE firms and small BOARDSIZE firms.

Because, it is obvious that company that has a large board of directors have higher probability about exist female independent director. So, If I don’t consider BOARDSIZE in this experiment, it can be concluded not “because a firm have female independent director, so CEO’s pay-performance sensitivity is changed.” but “because BOARDSIZE is high (then female independent directors are existed), so pay-performance sensitivity is changed.”

Also, BOARDSIZE is affected by size of firm, then if you put in BOARDSIZE and SIZE in the model together, it can be multicollinearity problem.

2] addition variables

- SIC2D : this variable is represent kind of job classification. Depending on industry trend and policy, some industry is boomed but another industry is affected by recessions. So, I think it can affect CEO compensation. But its value is not numeric so I SIC2D divide into 1digit(train\_1\_n) and 2digit(train\_2\_n) and make its dummy variables. moreover, to prevent perfect collinearity problem, I reduce train\_1\_1 dummy and train\_2\_0 dummy.

- DUALITY : if a firm have different CEO and chairman, board of director well perform a function such that monitoring, control, supervision about CEO. So, investor have trust in this firm it can be positive effect of finance of firm. however, if a firm’s CEO is same of chairman, it means that CEO’s authority

So, these complex factors can affect CEO compensation.

- CEOTEUNRE : likewise DUALITY, veteran CEO can have more authority than rookie CEO.

- IDPC : Contrariwise, if IDPC is higher, it means that independent director’s influence is higher. Then, board of director can perform original function.

- SIZE : total asset of a firm is very important factor of CEO compensation.

- SALESGROWTH : if the company is on an upward trajectory, then CEO is a bit advantageous about negotiation salary.

- FCF, MKTLEV: these are popular financial variable in corporation finance. The former means flexible money flow, the latter represent stability of a firm. So, I put in these variables in model.

- BUSYBOARD : When main independent directors participate many board, main director maybe less contribute of corporate governance of each firms. So, board of director may less monitor about CEO’S management.

-BM : it represent that whether this firm is overrated or not. If some firm has overrated than its book value, maybe this firm should pay high compensation to CEO.

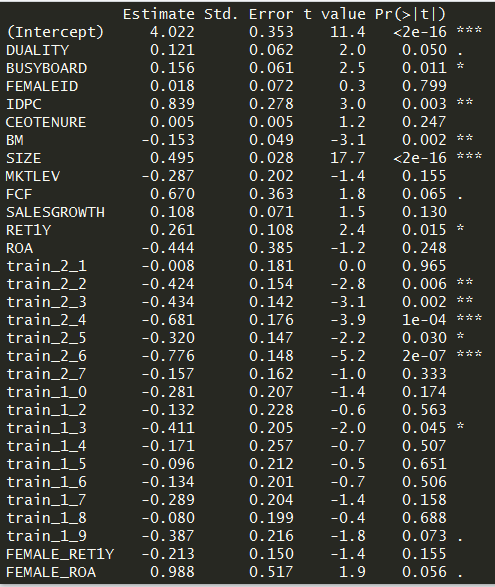
So, my regression model and hypothesis are like this.

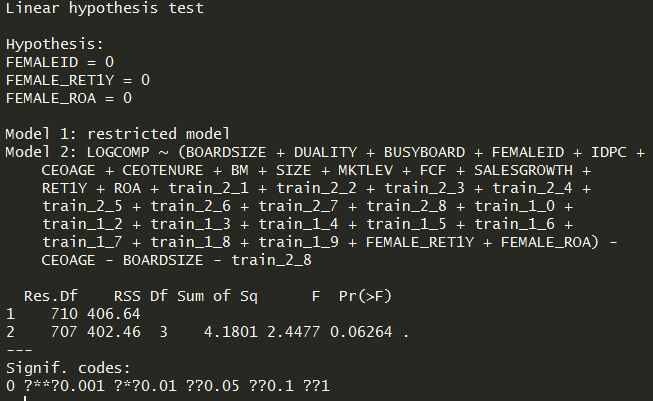
Data1) BOARDSIZE < 10

Data2) BOARDSIZE >= 10

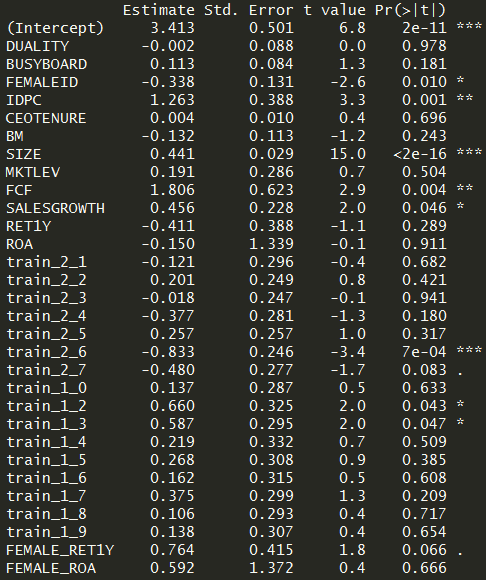
I will select significant level is 0.05 before run regression.

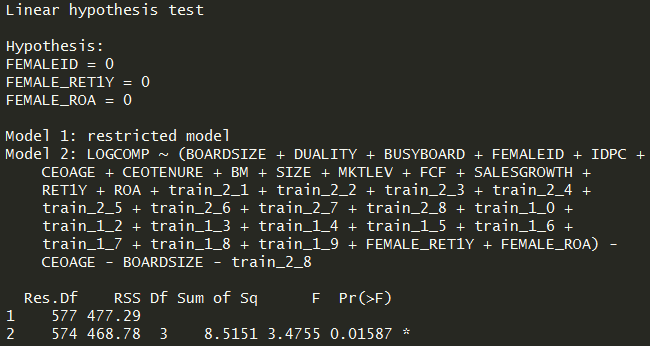
A3) result of multiple regression and hypothesis test of DATA1(Adjusted R-square = 0.4004)





Result of multiple regression and hypothesis test in DATA2 (Adjusted R-square = 0.3954)





DATA 1 : p-value = 0.062 > 0.05 & DATA 2 : p-value = 0.015 < 0.05

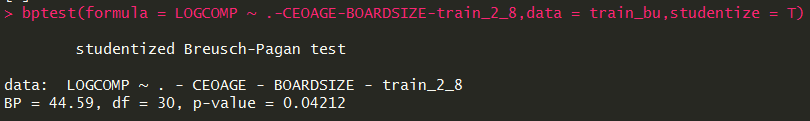
So, firms that have high BOARDSIZE, existence of female independent director is significantly affect the relation of pay performance sensitivity of CEO.

And regression coefficients are ,

1. the firms that have female independent director gave 0.45 times of ceo compensation than not. (log(-0.338) = 0.45)

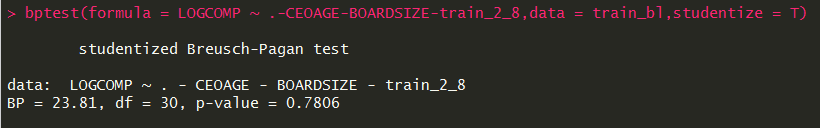
2. but, firms that have female independent director, if ROA increase 1, the ceo compensation increase 3.9 times and if RET1Y increase 1, the ceo compensation increase 5.8 times.

Then, to check this regression estimator can satisfying homoskedasticity hypothesis, I run bp-test.



The firms that BOARDSIZE are higher, regression estimator reject null hypothesis of bp-test.

I can tell that this regression estimator is satisfy homoskedasticity so I don’t need to Substitute to heteroskedasticity robust standardized error.

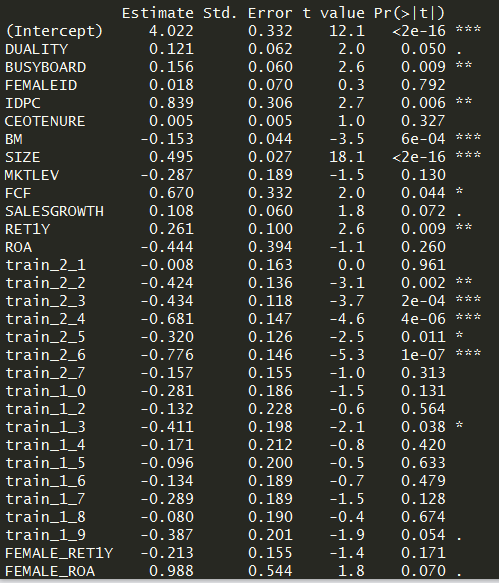


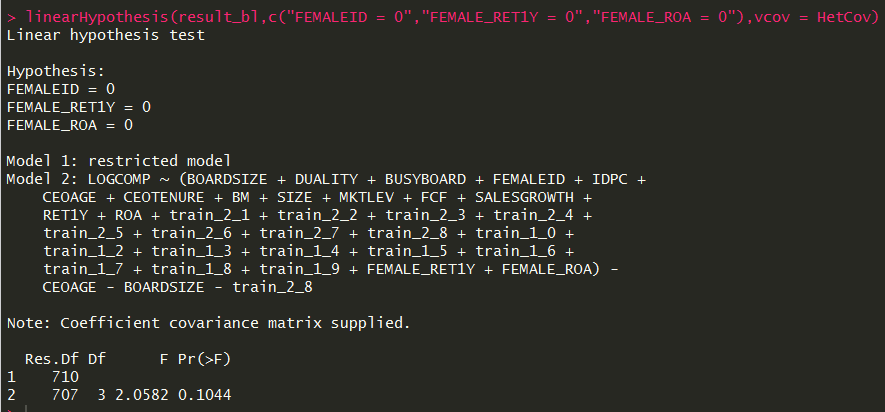
But, the firms that BOARDSIZE are less, I cannot tell that satisfying homoskedasticity.

So, I should use heteroskedasticity robust standardized error instead least square estimator.

These following results are regression and hypothesis test about

Using heteroskedasticity robust standardized error.





But, p-value is 0.1044 > 0.05

The firms that BOARDSIZE are less. So, I can’t tell about existence of female independent director is significantly affect the relation of pay performance sensitivity of CEO.

A4) Remaining issues

1. Gauss Markov assumption

- linearity : in generally, CEO compensation increase geometrically. So, I use log(CEOCOMP) instead CEOCOMP. then the relation of log(CEOCOMP) almost not have high polynomial relationship.

- Randomness : in fact, this data in the S&P 1500 firms so its cannot satisfying randomness of overall firms. But, S&P 1500 firms have most of market capitalization in America, to analyze hypothesis of these firms are useful.

To satisfy Randomness hypothesis, the solution is only collect more data.

- no perfect collinearity : because regression computations perform well, there isn’t perfect collinearity.

-zero conditional mean : in fact, I don’t tell that this regression satisfying zero conditional mean assumptions. Because the factors that it determines CEO compensation are infinitely many. So, many omitted values are underlying then also exist correlation of omitted values and some control variables.

To reduce omitted value bias as possible, we should collect more feature that have relationship about CEO compensation.

For example, institutional investor may be concerned in CEO compensation. Because if CEO compensation is higher, then the dividend of institutional investor is decreasing. So, features like as proportion of institutional investor can contribute to reduce omitted value bias problem.

- Homoskedasticity : to check Homoskedasticity assumption, I use bp-test.

Data1 (many board size) is satisfied Homoskedasticity, but Data2(few board size) cannot reject null hypothesis of bp-test. So, I use heteroskedasticity robust standardized error in Data2.

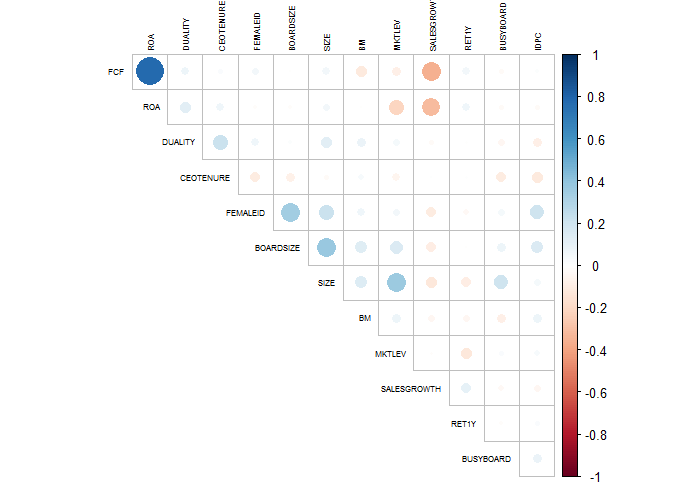
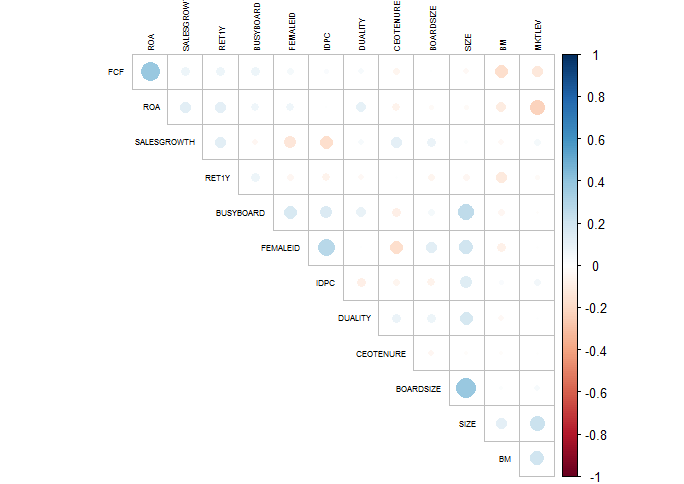
-error normality : because I use log transformation in dependent variables, then the possibility of existence outlier( +4 std, +5 std …) in error term is decline than before.

So, it is can basis about a claim “error terms follow normal distributions”

Moreover, if error terms didn’t follow normal distributions, if my data satisfied assumption 1~5, my regression estimator is consistent estimator so that can be advantage when run hypothesis test using t-test and f-test.

2. multicollinearity

To check multicollinearity, I visualize correlation matrix in these data ( up : Data1 – many board size, down : Data2 – few board size)



In totality, Data 2 have more correlated control variables than Data 1, especially, ROA and FCF have high correlation in common.

if regression model has high correlated control variables, the variance of regression estimator is higher. to solve multicollinearity problem, I think these two solutions.

1) Ridge regression : the difference of Ridge regression and least square regression is cost function. Ridge regression optimize least square error plus l2 decay(sum of square of weight terms). it generates bias estimator but, the variances of estimation are sharply decreasing.

2)feature preprocessing : since ROA and FCF have high correlation, we can reduce correlation through feature preprocessing. For example, FCF affected by maintenance costs like as long-term debt. But ROA affected by long-term debt because total asset contained long-term debt So, if I use ROE ( ROA\* financial leverage), it calculated regardless any debt. so I expect the correlation of FCF and ROE are less than the correlation of FCF and ROA.

A5) reference

원구환(2008) 지방공기업 내부지배구조로서의 이사회 구성과 재무성과간의 상관성 분석, 지방정부연구, 11(4) 111-124

한국상장회사협의회 – 이사회의 특성이 기업가치에 미치는 영향

차운아, 정태훈 규제연구 제23권 제1호, 2014.04, 131-163 이사회 다양성과 기업성과