CEO Characteristics and Firm R&D Spending

Math Capstone PBL (Data Analysis) - Project 1

Jaeseon Lee[1] Junwoo Yang[2]

[1]Department of Economics & Finance
Hanyang University

[2]Department of Finance
Hanyang University

October 27, 2020

Outline

Outline

Topic

Does the CEO's nature affect the company's R&D investment?

- What financial variables are relevant to the extent of the firm's R&D investment?
- Does the variable that characterizes the CEO lack explanatory power for R&D investments?
- If so, what characteristics correlate?

Outline

Data description

Dimension of raw data

1726 rows: U.S. S&P 1,500 + other Execucomp-reporting firms

27 columns: financial variables + CEO characteristics

Data source

Wharton Research Data Services (WRDS) Compustat Data / Execucomp
[1] [2]

[1] http://wrds-www.wharton.upenn.edu.ssl.access.hanyang.ac.kr/

[2] https://lib.hanyang.ac.kr/#/er/web

Variable list

Variables for identifying firms

GVKEY: A firm's identifier used by S&P Capital IQ / Compustat / Execucomp

SIC: A four-digit number that identifies a firm's primary industry of operation [3]

STATE: A two-letter code denoting the state in which the firm is headquartered

[3] Variables SIC2D and SIC3D denote the first two or three letters of the SIC code – the first two letters denote the broad classification, the third letter the middle classification

Variable list

Financial variables (1)

SIZE:	Firm size (total assets)
BM:	Book-to-market
FCF:	Free cash flow (divided by total assets)
HHI:	Herfindhal-Hirschman Index (the more competitive the industry a firm operates in, the lower this number. It is always between 0 and 1)
OPPERF:	Operating performance (divided by total assets)
LEVERAGE:	Market value of a firm's leverage (divided by total mark-to-market assets)
TOBINSQ:	Tobin's Q
RNDMISSING:	A dummy variable that equals 1 if a firm's R&D expenditure is missing in the financial statement
RNDRATIO:	A firm's R&D expenditure (divided by total assets)

Variable list

Financial variables (2)

ROA: Return on assets (divided by total assets)

SALESGROWTH: Year on year growth rate of a firm's sales

CASHRATIO: A firm's cash holdings (divided by total assets)

DIVPAY: A dummy variable that equals 1 if and only if a firm has paid out dividends in the same fiscal year.

INTAN: Intangibility measure. The higher this number, the more intangible a firm's assets ((total assets – property, plants, and equipments value)/total assets)

INVEST: A firm's investment ratio (investment / property, plants, and equipments value)

Variable list

CEO characteristics

CEOAGE: CEO's age

CEOCOMP: CEO's total compensation, in $1000s

INSIDERCEO: CEO who was promoted from inside the company (as opposed to an outsider)

FEMALECEO: A dummy variable that equals 1 if the CEO is female

CEOPAYSLICE: CEO's compensation divided by the firm's top 5- earning directors' total compensation (the higher this number, the more disproportionately highly a CEO is compensated within the top management team) [4]

CEOEQUITY: CEO's holdings of the firm's stock

CEOTENURE: CEO's tenure in his or her current position, in years

[4] Bebchuk, Cremers, and Peyer (2011) argue this is a good measure of how powerful a CEO is within the company.

Outline

2. Body
   - Handling missing values
   - Data Transforamtion
   - Summary statistics
   - Data Analysis: R&D ratio - Continuous variables
   - Data Analysis: R&D ratio - Categorical variables
   - Data Analysis: Between Continuous variables
   - Data Analysis: Between Categorical variable and Continuous variable

Handling missing values

Variable	N	%
GVKEY	0	0
SIC	0	0
SIC2D	0	0
SIC3D	0	0
STATE	56	3.24
SIZE	1	0.06
BM	18	1.04
FCF	104	6.03
HHI	0	0
OPPERF	104	6.03
LEVERAGE	36	2.09
TOBINSQ	18	1.04
RNDMISSING	0	0
RNDRATIO	779	45.13
ROA	2	0.12
SALESGROWTH	5	0.29
DIVPAY	0	0
CASHRATIO	2	0.12
INTAN	101	5.85
INVEST	120	6.95
CEOAGE	1	0.06
CEOCOMP	1	0.06
INSIDERCEO	0	0
FEMALECEO	0	0
CEOPAYSLICE	266	15.41
CEOEQUITY	29	1.68
CEOTENURE	21	1.22

Table: Missing values of raw data

Mean of **SIC**

3672: Printed Circuit Boards

↓

Mean of **SIC3D**

(middle classification)

367: Electronic, Component and Accessories

↓

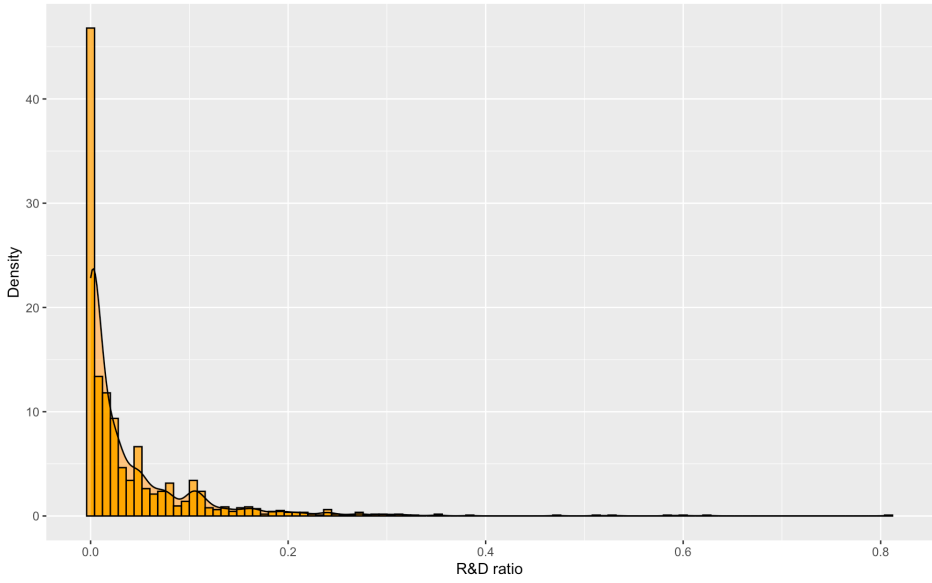
Mean of **SIC2D**

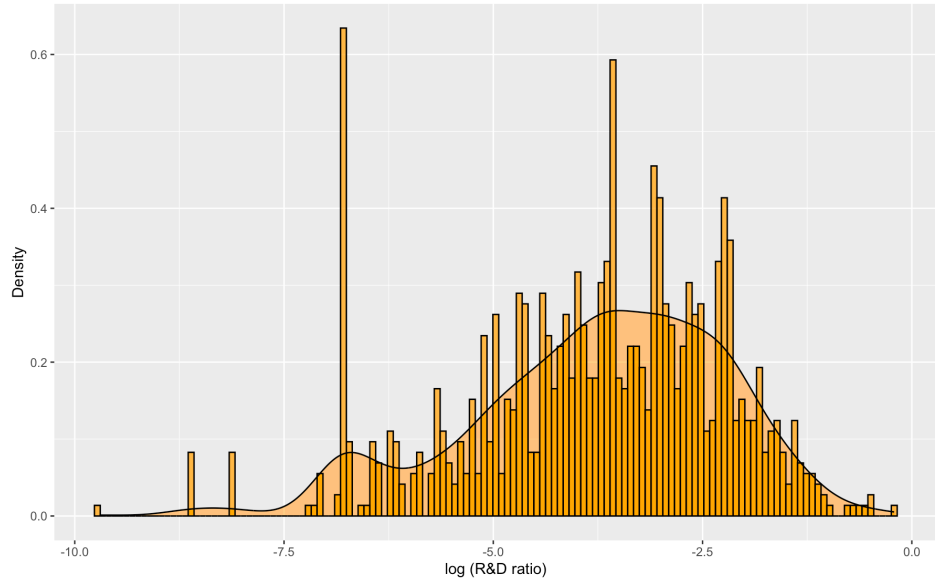
(broad classification)

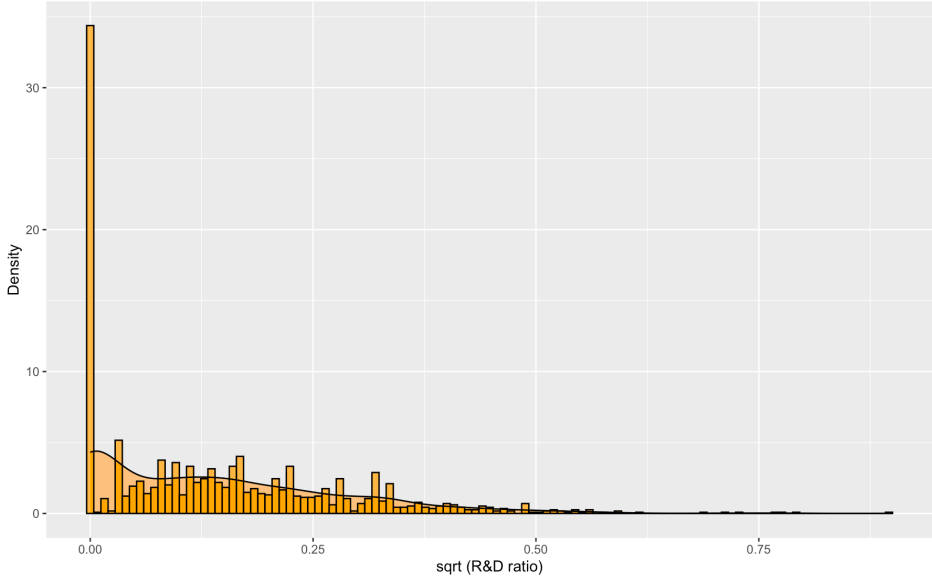
36: Electronic and Other Equipment

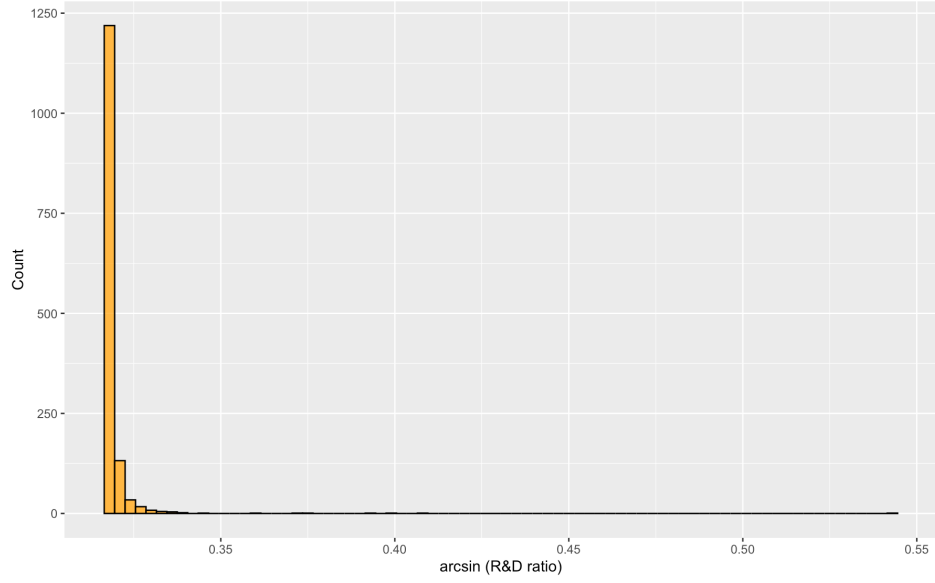
Outline

Transformation of R&D ratio

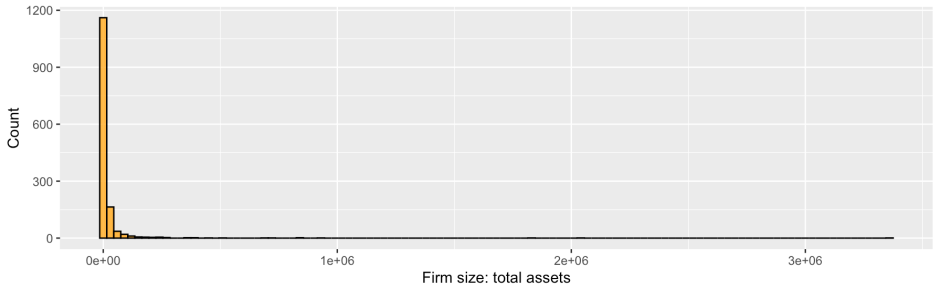


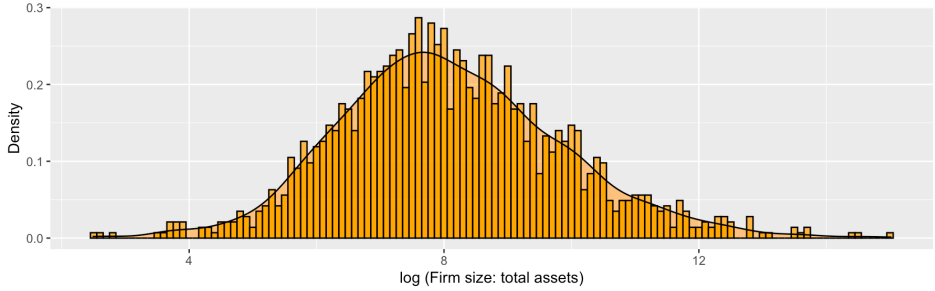


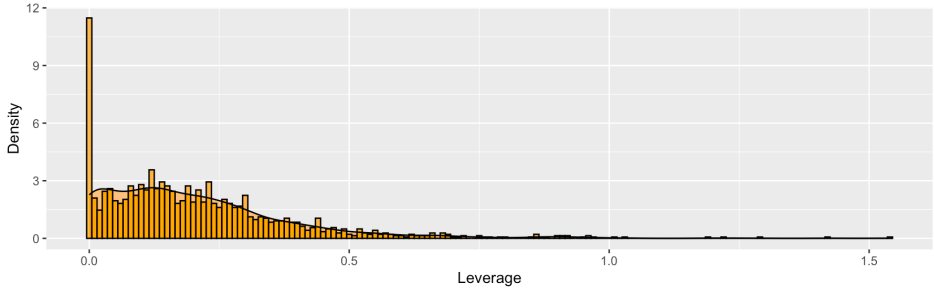


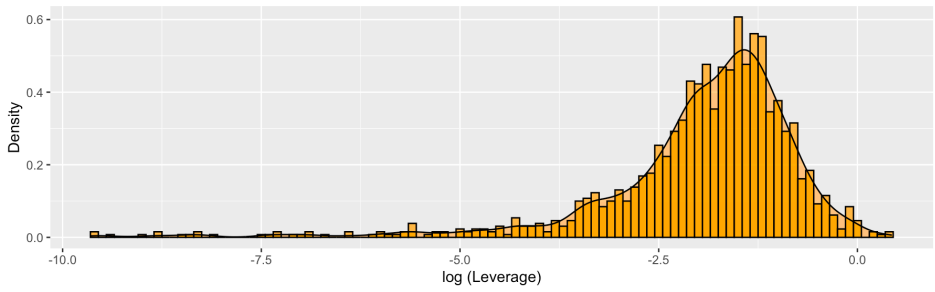


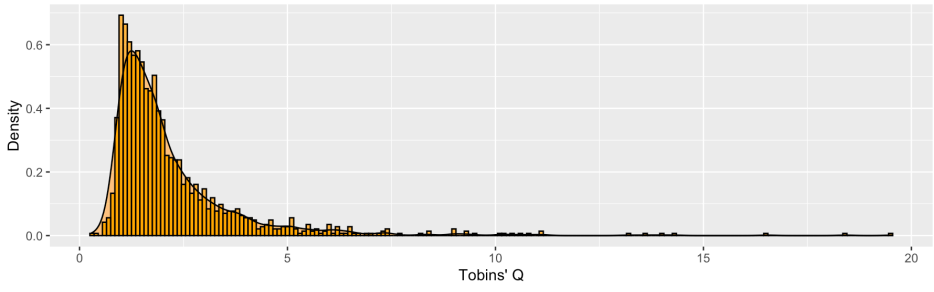
Log transformation (1)

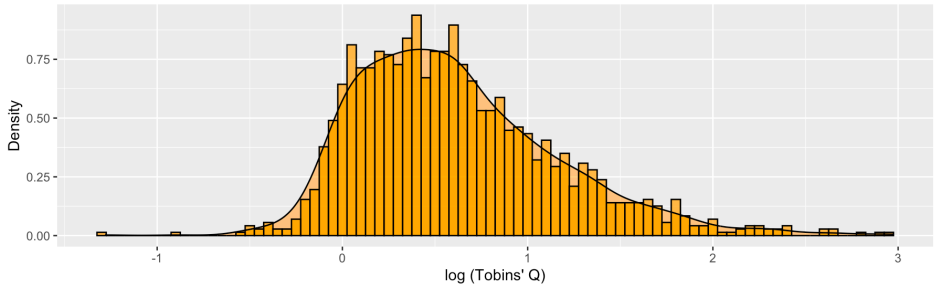


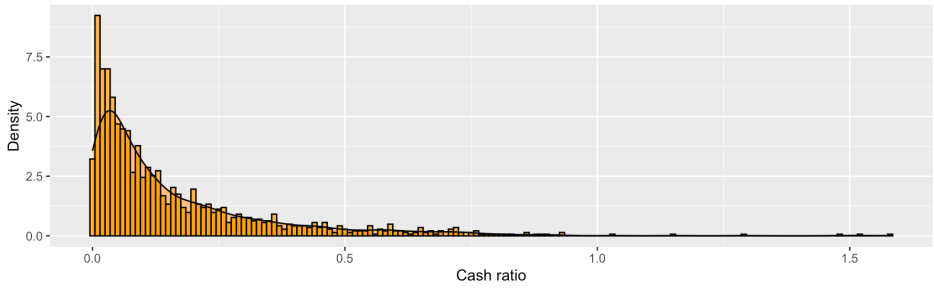


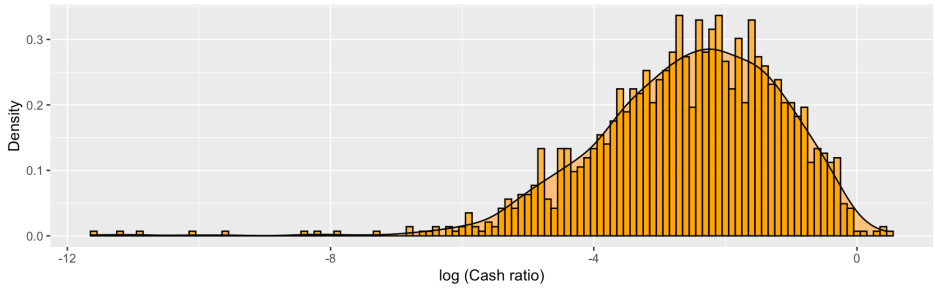




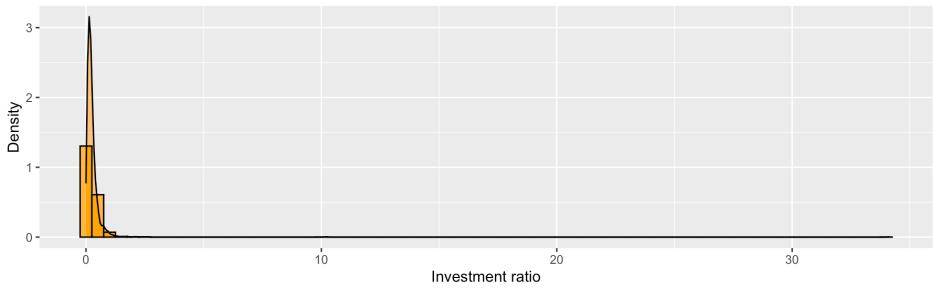


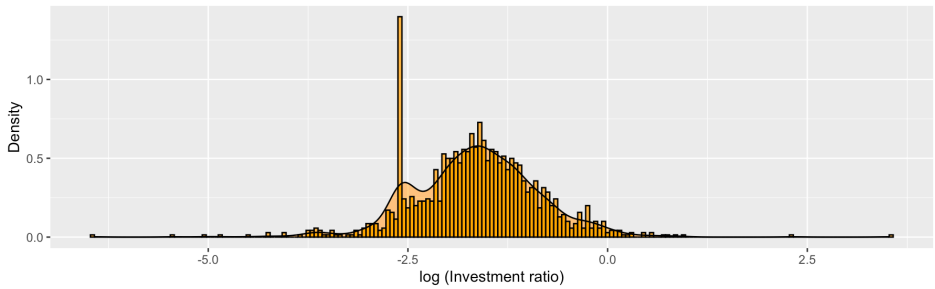


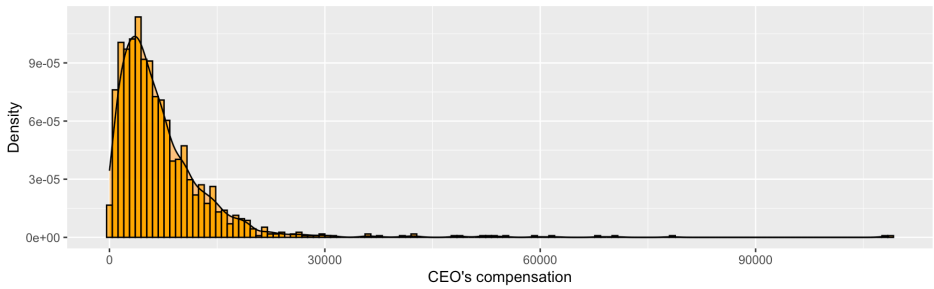


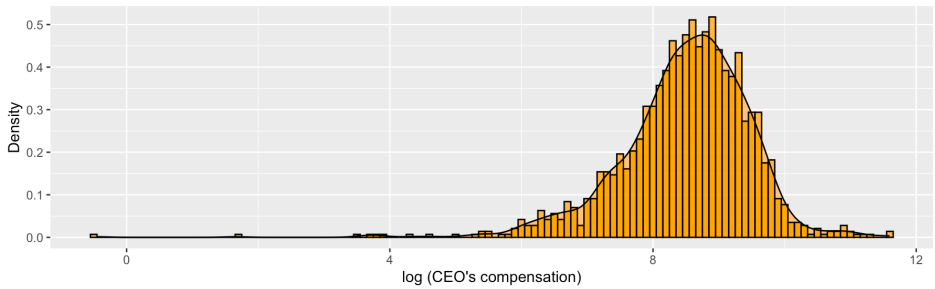


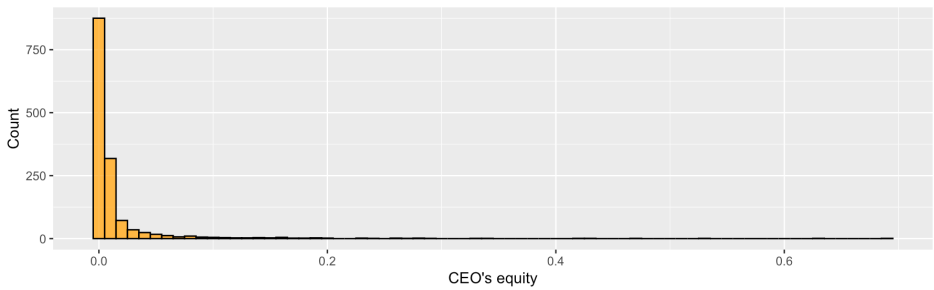
Log transformation (2)

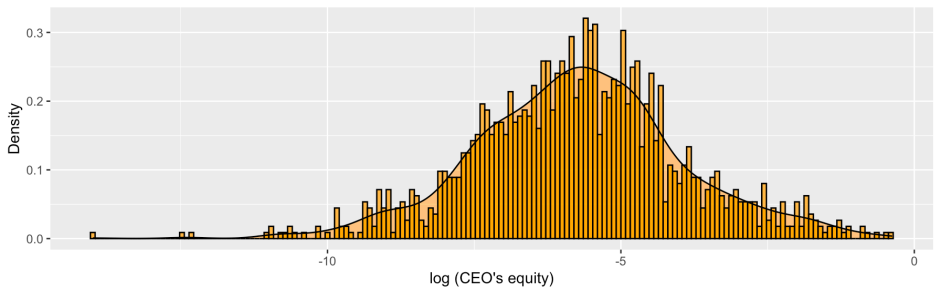


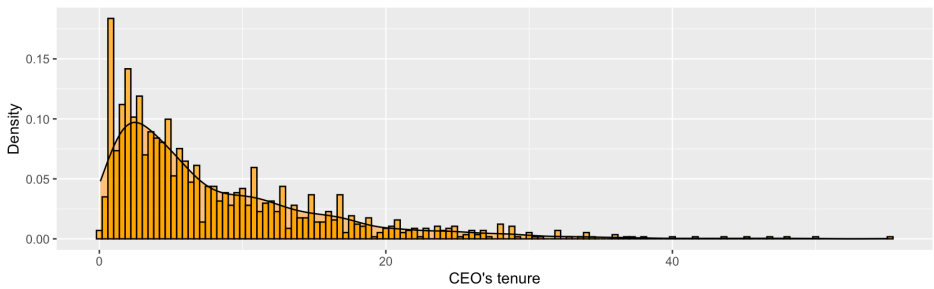


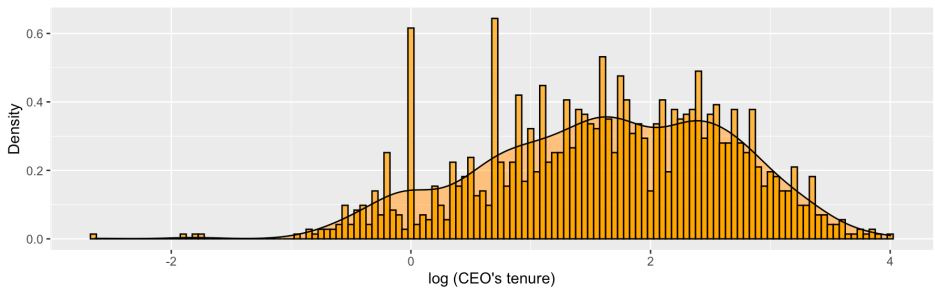




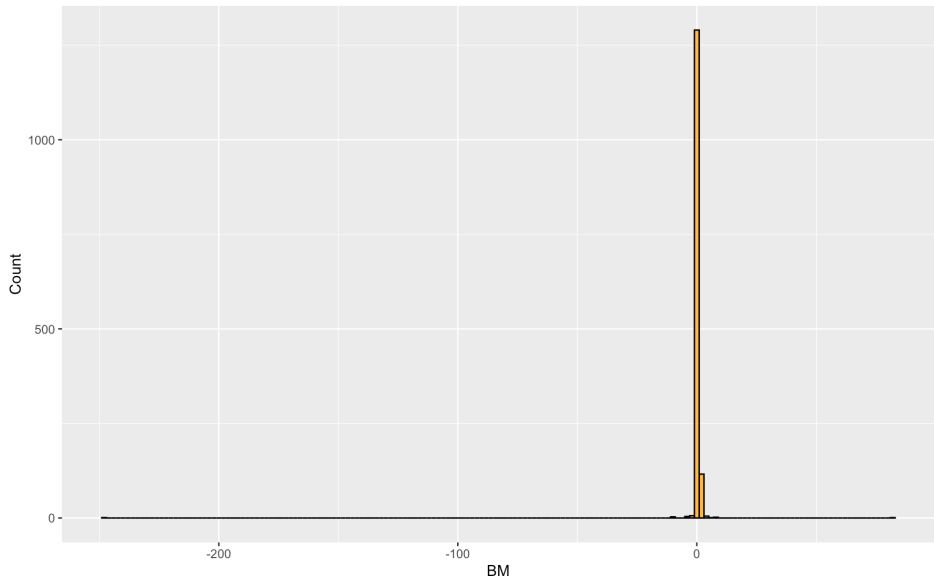


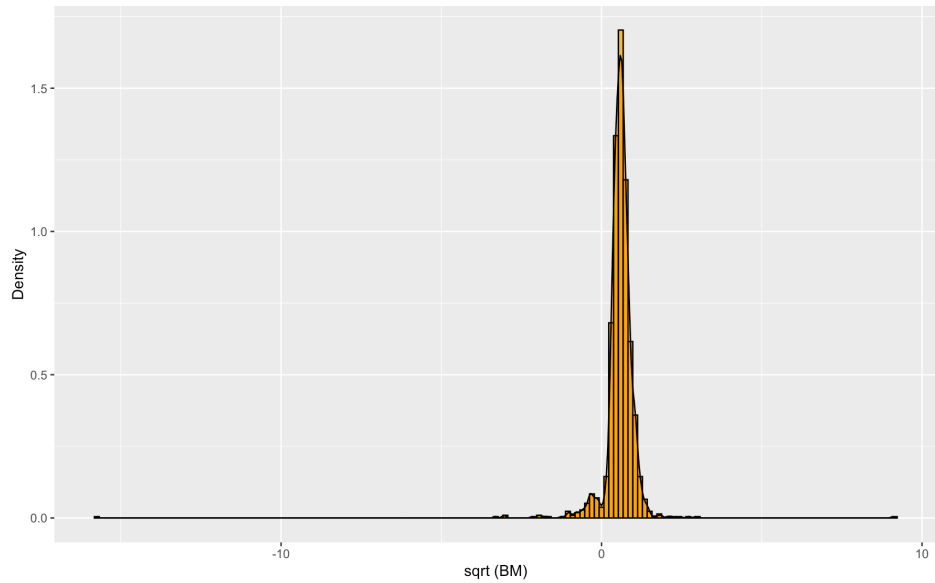


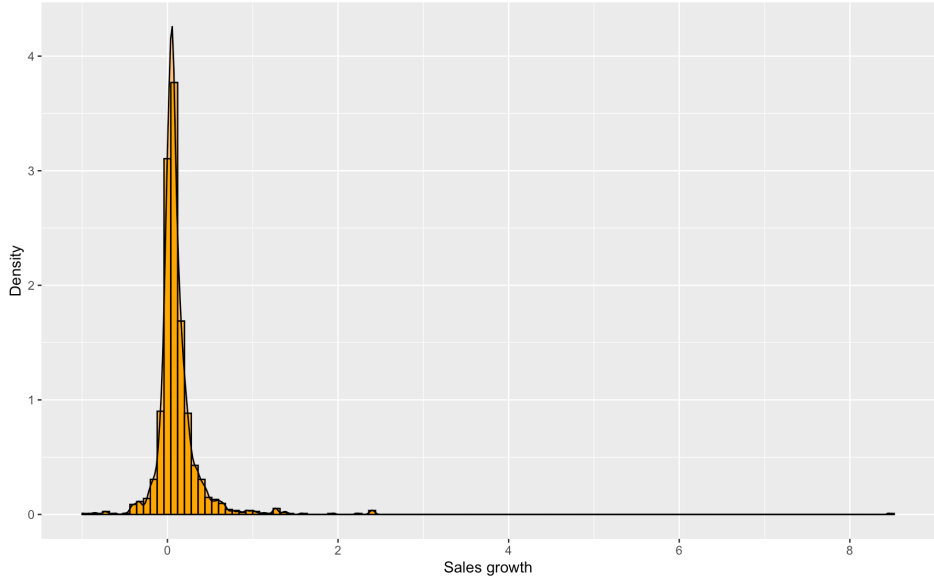


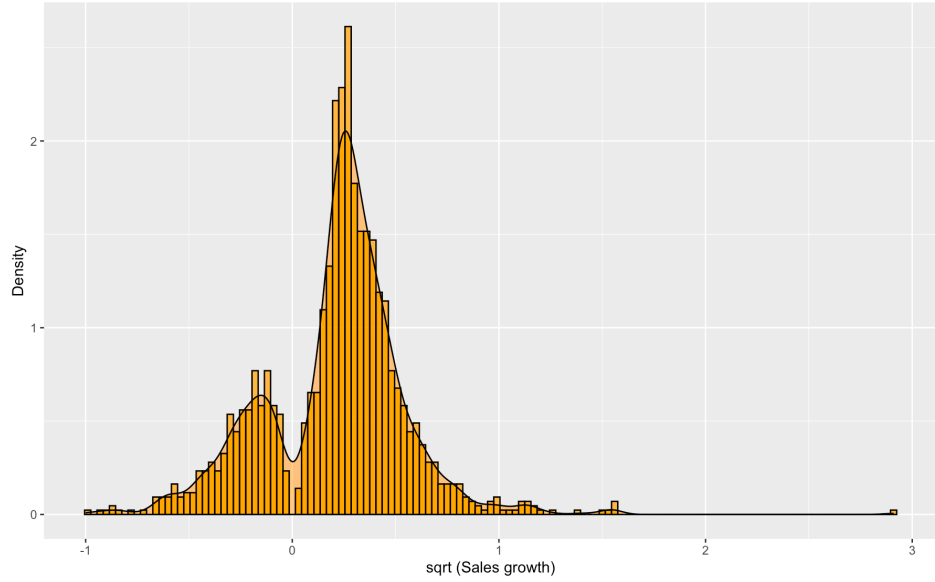


Square-root transformation (with sign)

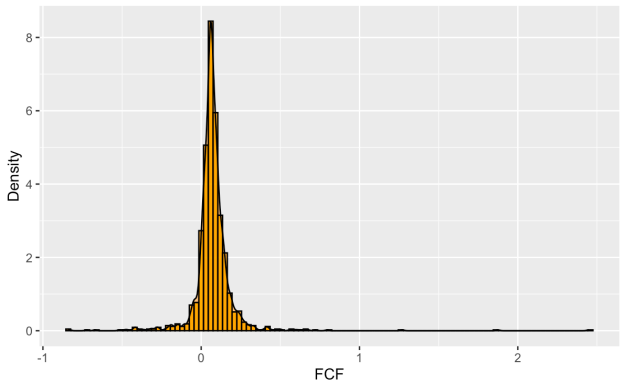


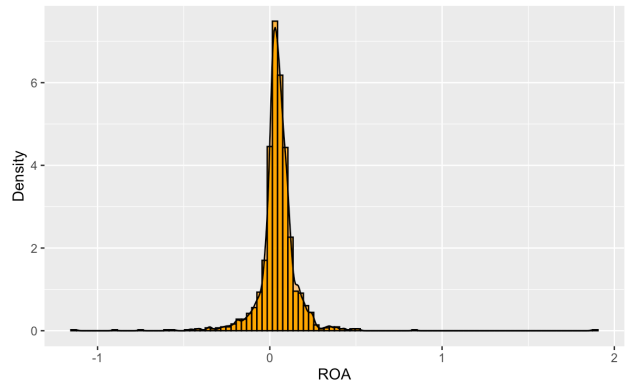


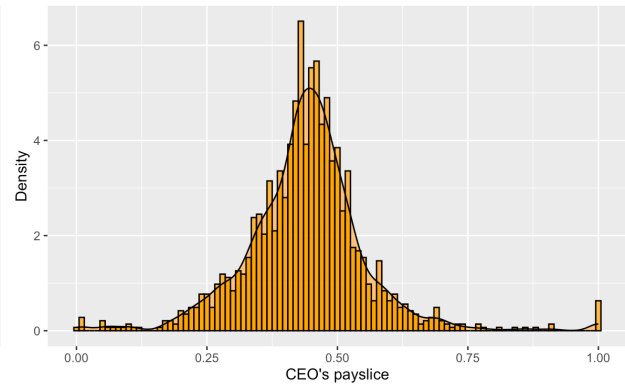


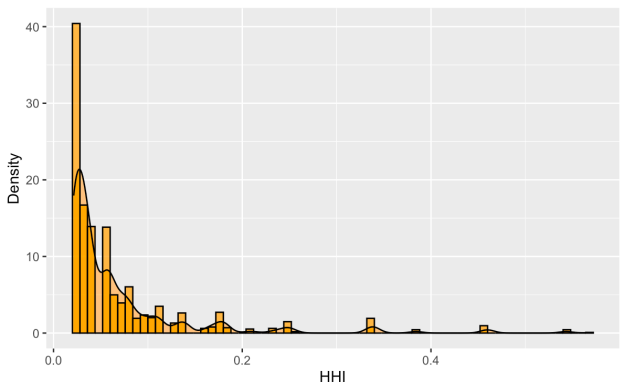


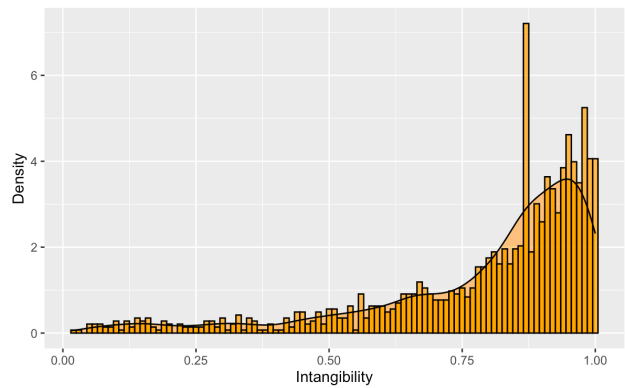
Not transformed

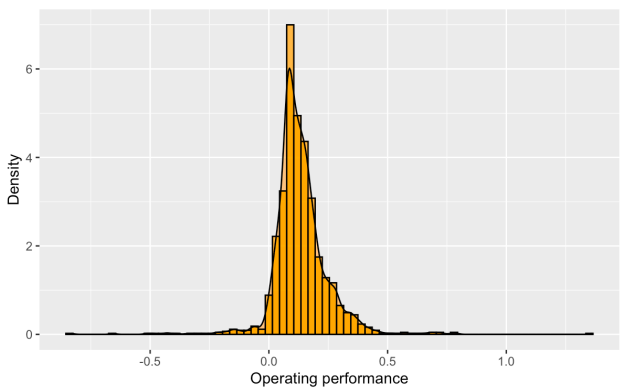












Outline

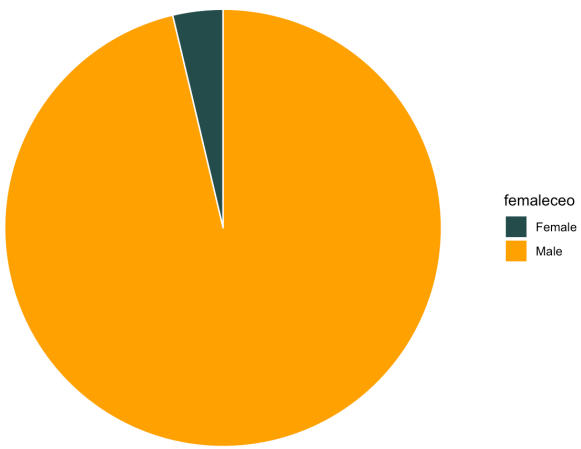
Summary statistics for continuous variables

Variable	Mean	St. Dev.	Q1	Q3	Max	Min	Kurtosis	Skewness
$\log$ SIZE	8.098	1.745	6.941	9.163	15.023	2.483	0.470	0.333
$\log$ LEVERAGE	-1.951	1.266	-2.319	-1.217	0.435	-9.648	8.524	-2.336
$\log$ TOBINSQ	0.613	0.544	0.212	0.914	2.973	-1.281	1.105	0.869
$\log$ CASHRATIO	-2.568	1.421	-3.418	-1.543	0.458	-11.645	3.454	-1.068
$\log$ INVEST	-1.649	0.798	-2.129	-1.171	3.530	-6.430	3.174	-0.060
$\sqrt{\text{BM}}$	0.569	0.643	0.435	0.765	9.096	-15.762	313.873	-10.471
$\sqrt{\text{SALESGROWTH}}$	0.213	0.339	0.068	0.391	2.910	-1.000	3.755	0.187
FCF	0.073	0.140	0.329	0.106	2.447	-0.844	85.110	4.820
HHI	0.068	0.077	0.026	0.073	0.566	0.021	12.951	3.296
OPPERF	0.133	0.120	0.080	0.178	1.363	-0.832	16.421	0.656
ROA	0.045	0.120	0.010	0.087	1.878	-1.128	50.972	1.169
INTAN	0.796	0.212	0.722	0.946	1.000	0.015	2.230	-1.631
RNDRATIO	0.039	0.069	0.000	0.050	0.809	0.000	25.756	4.033
CEOAGE	33.000	7.060	53.000	61.000	87.000	33.000	0.996	0.407
$\log$ CEOCOMP	8.489	1.003	7.994	9.135	11.596	-0.453	7.256	-1.380
$\log$ CEOEQUITY	-5.725	1.774	-6.858	-4.699	-0.377	-13.999	0.671	-0.061
$\log$ CEOTENURE	1.625	1.023	0.916	2.418	4.009	-2.643	-0.339	-0.328
CEOPAYSLICE	0.440	0.116	0.381	0.494	1.000	0.001	4.312	0.520

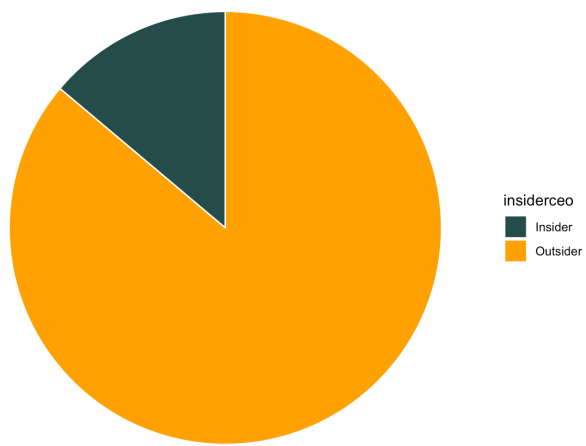
Table: Summary statistics

Pie chart of categorical variables

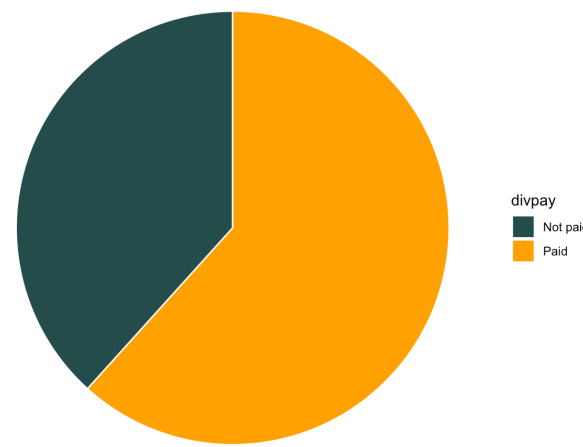
Female CEO



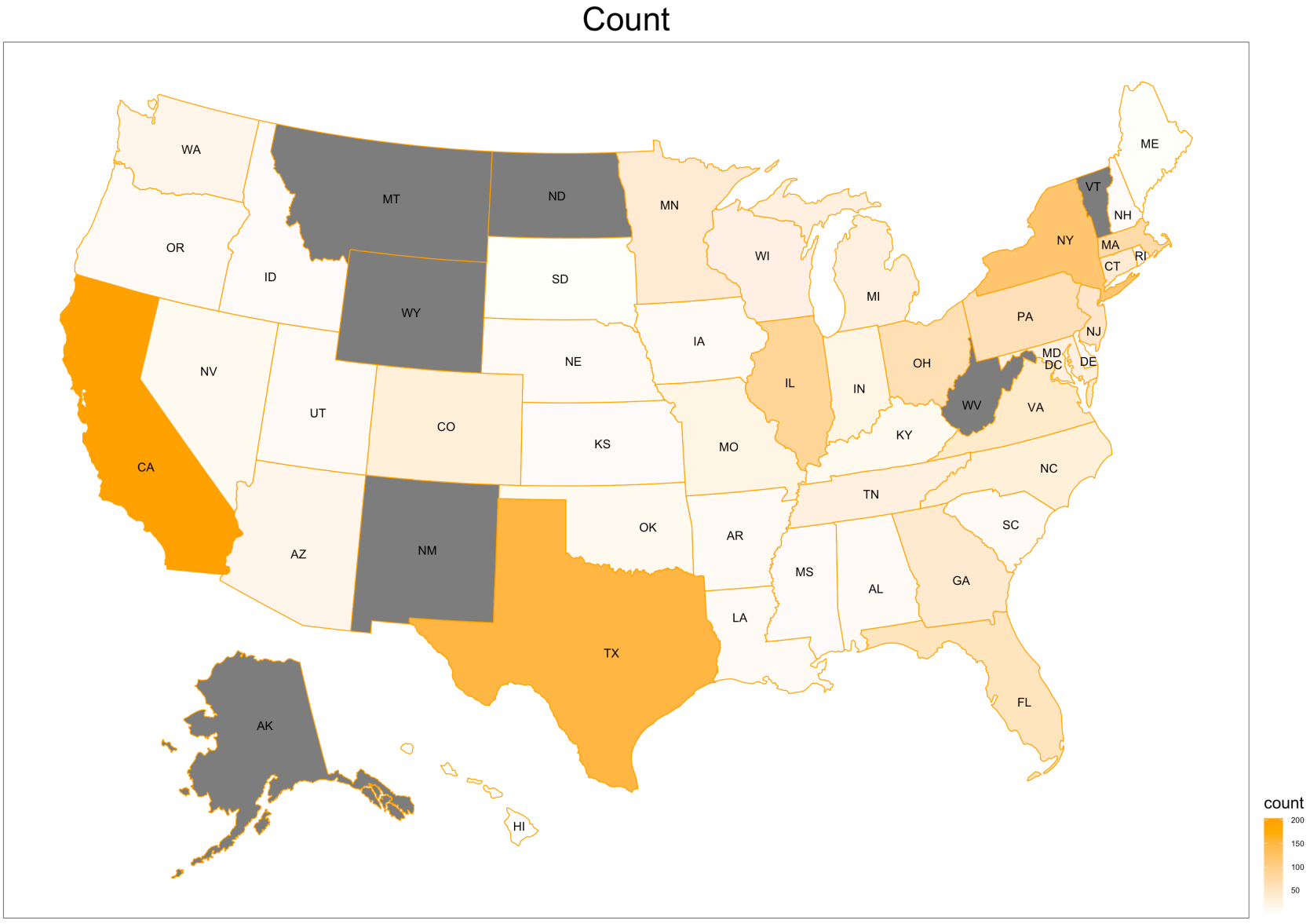
Insider CEO



Dividends paid



Count by state



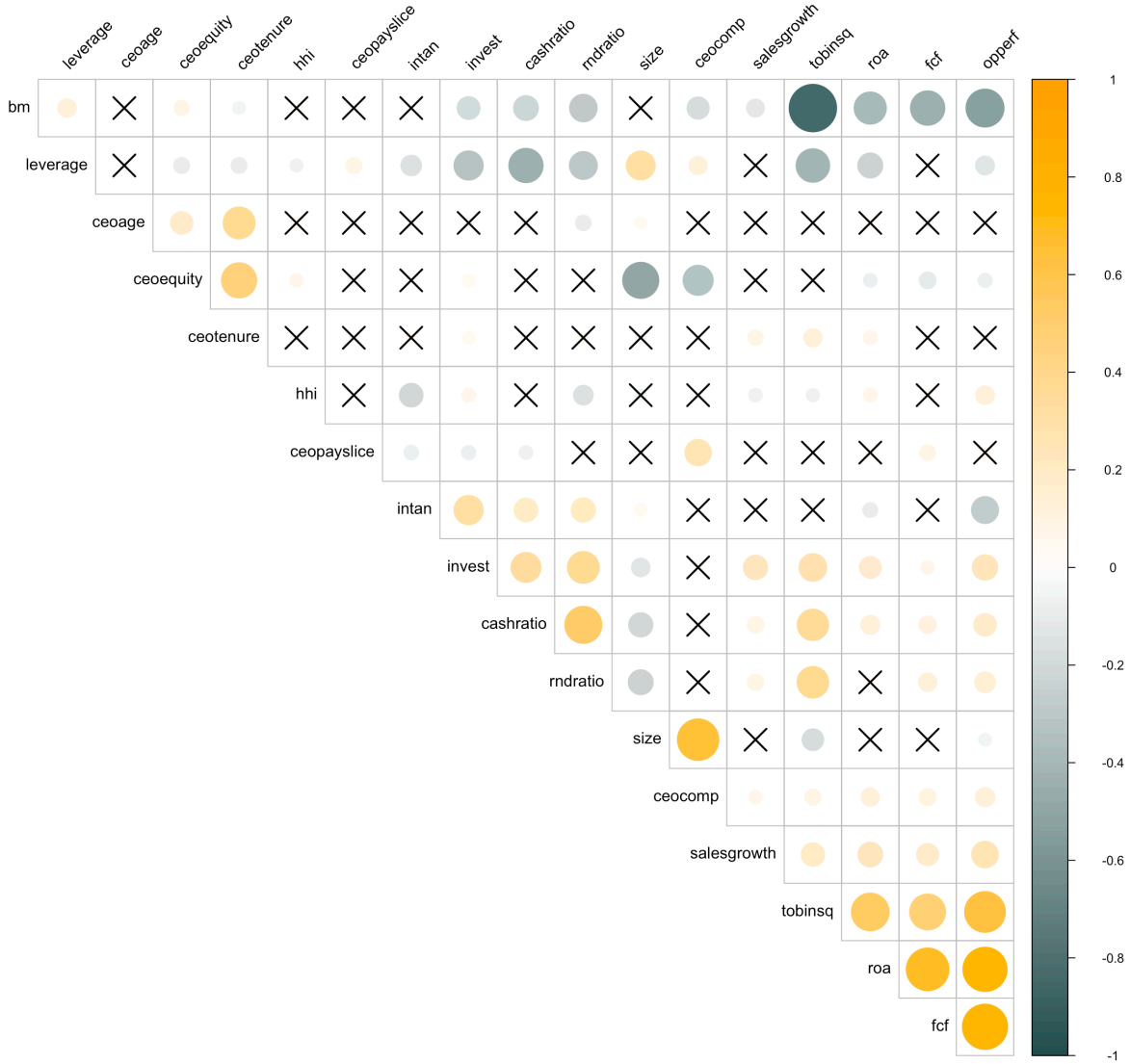
Outline

Kendall correlation test

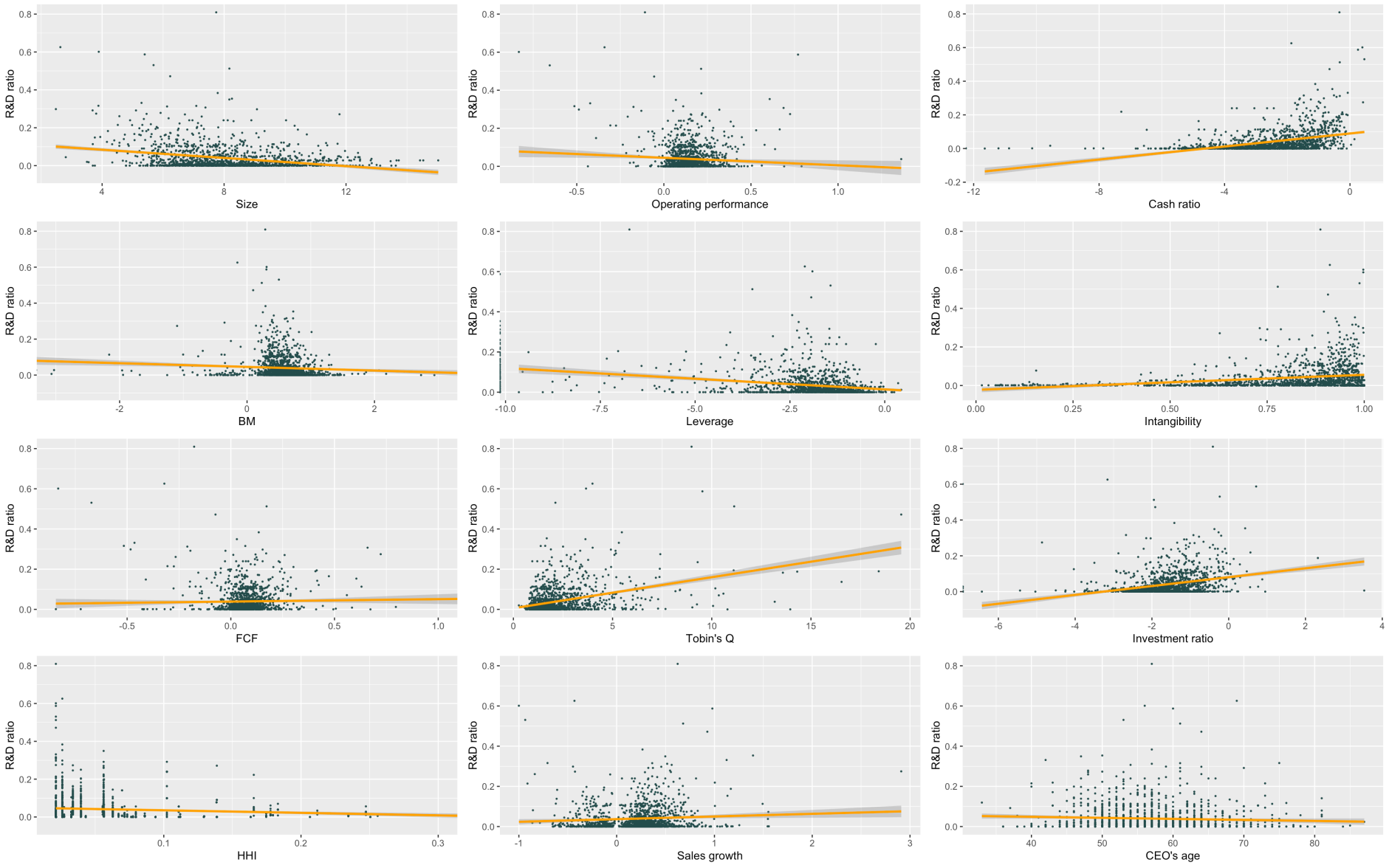
Variable	τ	p-value
ROA	0.025	0.174
$\log$ CEOCOMP	-0.004	0.842
$\log$ CEOEQUITY	0.012	0.493
$\log$ CEOTENURE	0.028	0.122
CEOPAYSLICE	-0.027	0.144

Table: Variables that can't reject the H_0

Combining Spearman correlogram with significance test



Scatter plot and regression line



Stargazer for simple linear regression (1)

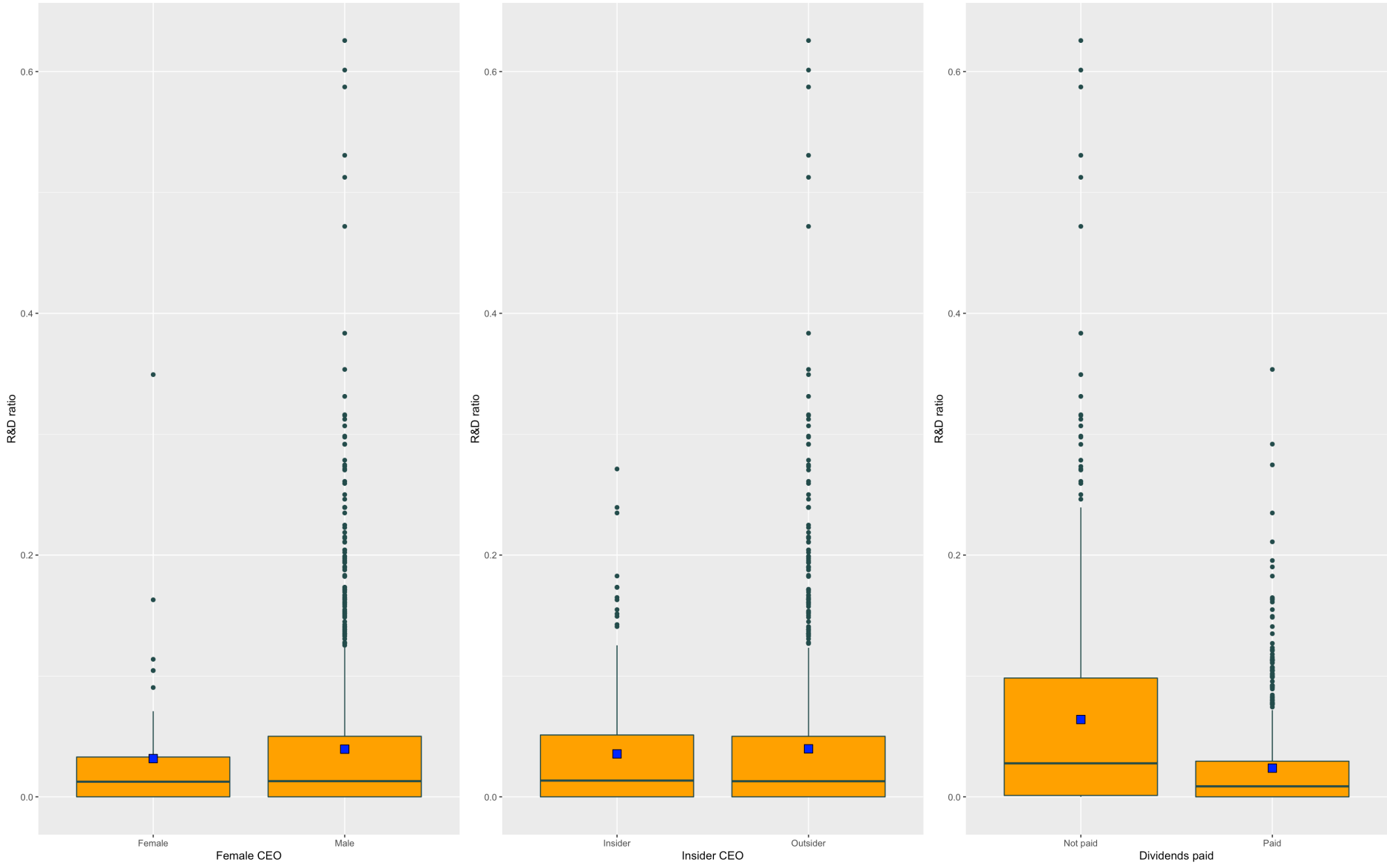
	Dependent variable:					
	R&D ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
log SIZE	-0.011^{***}					
	(0.001)					
log LEVERAGE		-0.011^{***}				
		(0.001)				
log TOBINSQ			0.015^{***}			
			(0.001)			
log CASHRATIO				0.019^{***}		
				(0.001)		
log INVEST					0.025^{***}	
					(0.002)	
CEOAGE						-0.001^{**}
						(0.0003)
Constant	0.127^{***}	0.014^{***}	0.006^{**}	0.089^{***}	0.081^{***}	0.069^{***}
	(0.008)	(0.003)	(0.003)	(0.003)	(0.004)	(0.015)
Observations	1,429	1,301	1,429	1,426	1,402	1,429
R^2	0.074	0.043	0.147	0.157	0.084	0.003
Adjusted R^2	0.074	0.043	0.147	0.157	0.083	0.002
Residual Std. Error	0.067 (df = 1427)	0.063 (df = 1299)	0.064 (df = 1427)	0.064 (df = 1424)	0.065 (df = 1400)	0.069 (df = 1427)
F Statistic	114.643^{***} (df = 1; 1427)	59.035^{***} (df = 1; 1299)	246.718^{***} (df = 1; 1427)	266.045^{***} (df = 1; 1424)	128.159^{***} (df = 1; 1400)	3.897^{**} (df = 1; 1427)
Note:						*p<0.1; **p<0.05; ***p<0.01

Stargazer for simple linear regression (2)

	Dependent variable:					
	R&D ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
$\sqrt{\text{BM}}$	-0.010^{***}					
	(0.003)					
$\sqrt{\text{SALESGROWTH}}$		0.013^{**}				
		(0.005)				
FCF			0.012			
			(0.013)			
HHI				-0.134^{***}		
				(0.024)		
OPPERF					-0.039^{***}	
					(0.015)	
INTAN						0.078^{***}
						(0.008)
Constant	0.045^{***}	0.037^{***}	0.039^{***}	0.049^{***}	0.045^{***}	-0.023^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.007)
Observations	1,429	1,429	1,429	1,429	1,429	1,429
R^2	0.009	0.004	0.001	0.022	0.005	0.057
Adjusted R^2	0.008	0.004	-0.0001	0.021	0.004	0.057
Residual Std. Error (df = 1427)	0.069	0.069	0.069	0.069	0.069	0.067
F Statistic (df = 1; 1427)	12.984^{***}	6.020^{**}	0.846	32.064^{***}	6.706^{***}	86.702^{***}
Note:						*p<0.1; **p<0.05; ***p<0.01

Outline

Box plot by femaleceo, insiderceo, divpay



Welch's T-test and Two sample F-test

Welch's T-test (in which a two-tailed test is applied)

$H_{T,0}$: The two population means are equal.

$H_{T,1}$: True difference in means is not equal to 0.

F-test of equality of variances[5]

$H_{F,0}$: Two normal populations have the same variance.

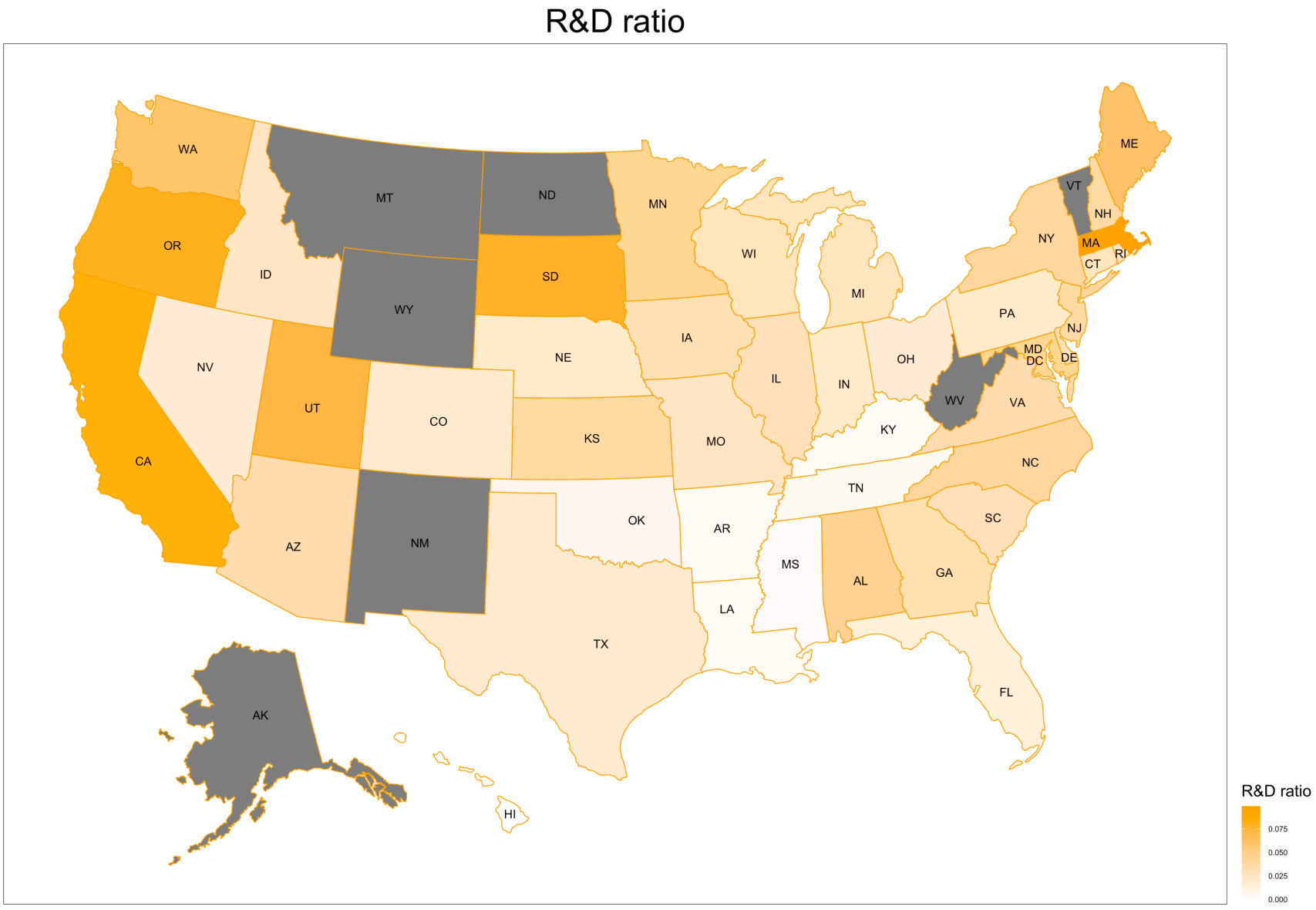
$H_{F,1}$: True ratio of variances is not equal to 1.

Variable	F	T	result
DIVPAY	$<$ 2.2e-16	0.3033	$H_{F,1}$, $H_{T,0}$
INSIDERCEO	3.995e-09	0.244	$H_{F,1}$, $H_{T,0}$
FEMALECEO	0.06457	0.3895	$H_{F,0}$, $H_{T,0}$

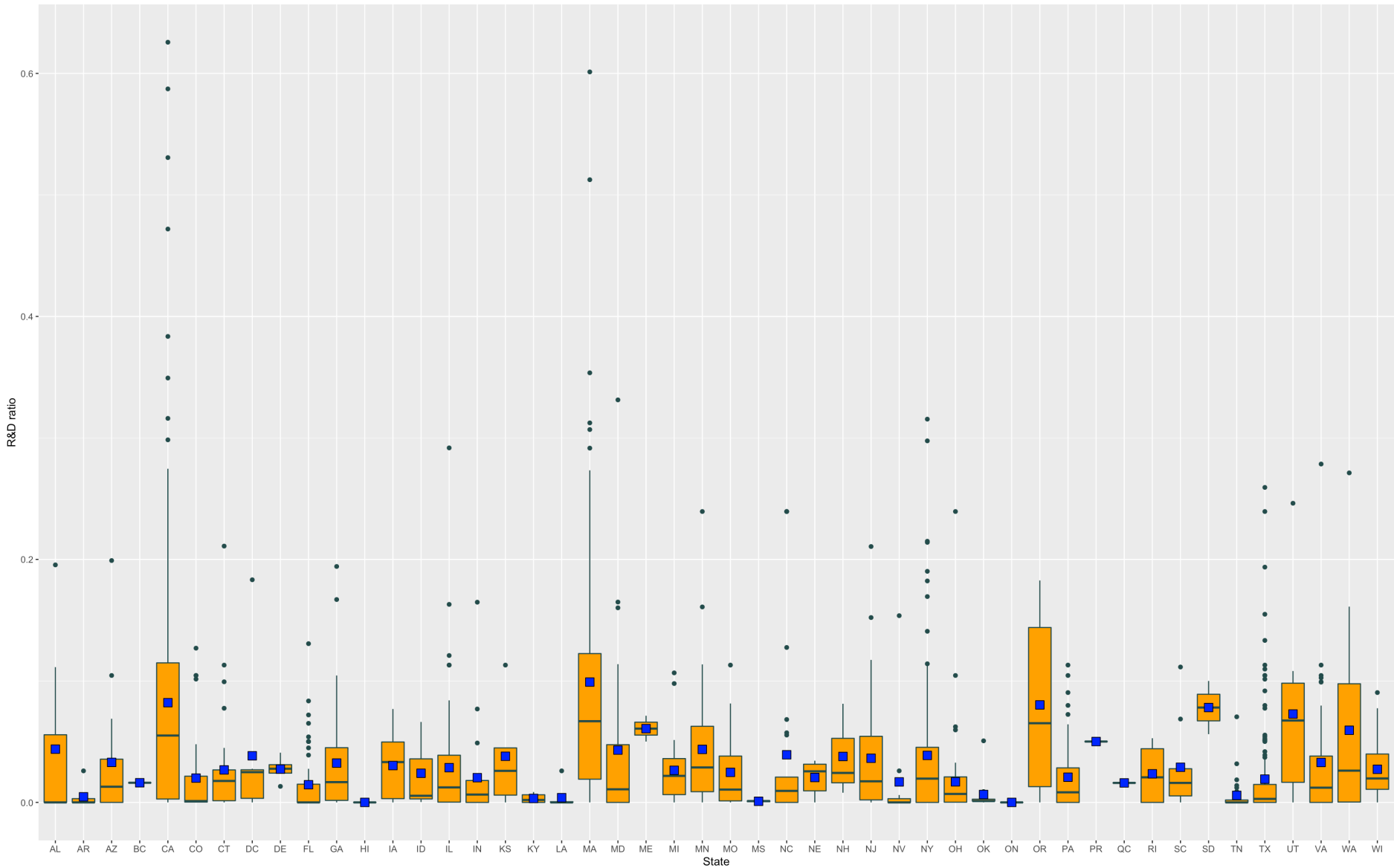
Table: p-values of two sample tests

[5] $H_{F,0}$: Homogeneity of variance, $H_{F,1}$: Heteroscedasticity of variance

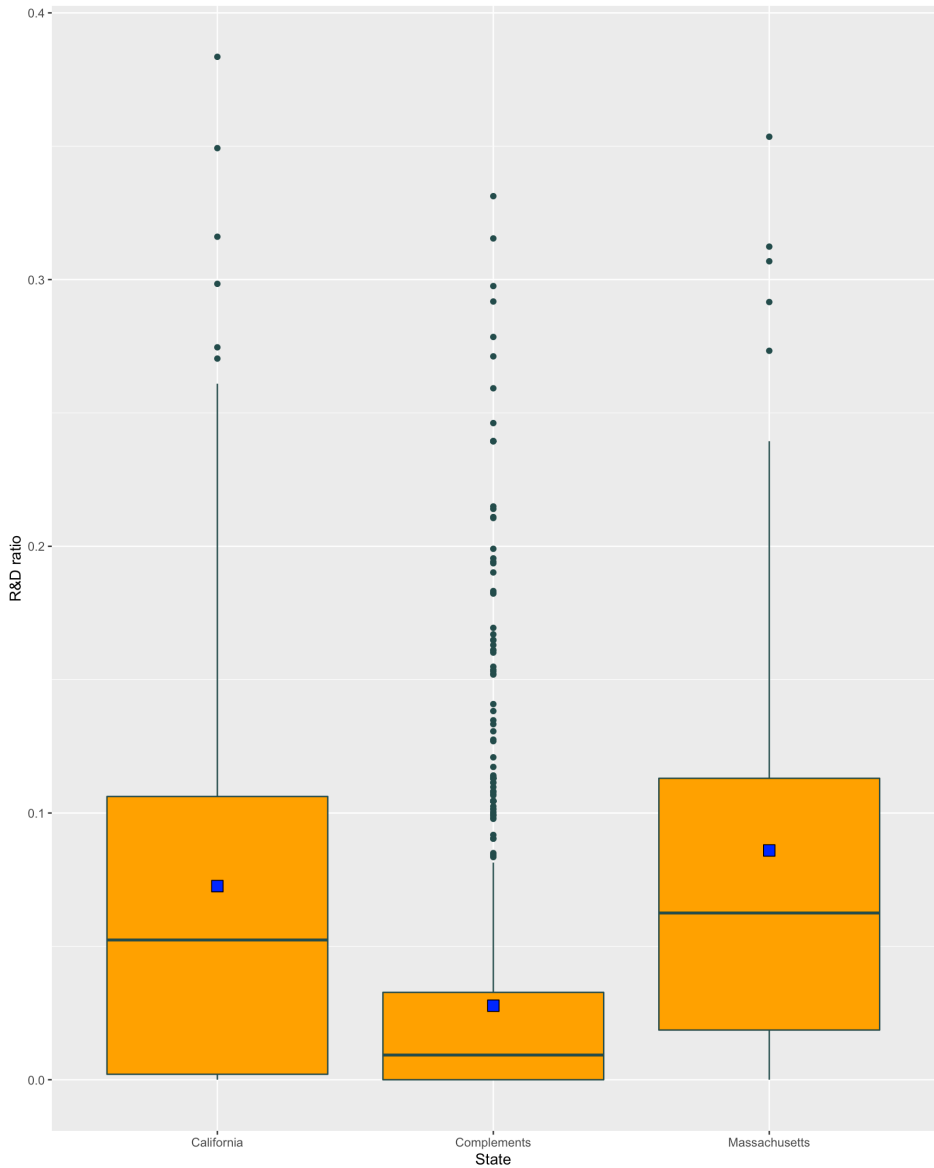
Mean of R&D ratio by state

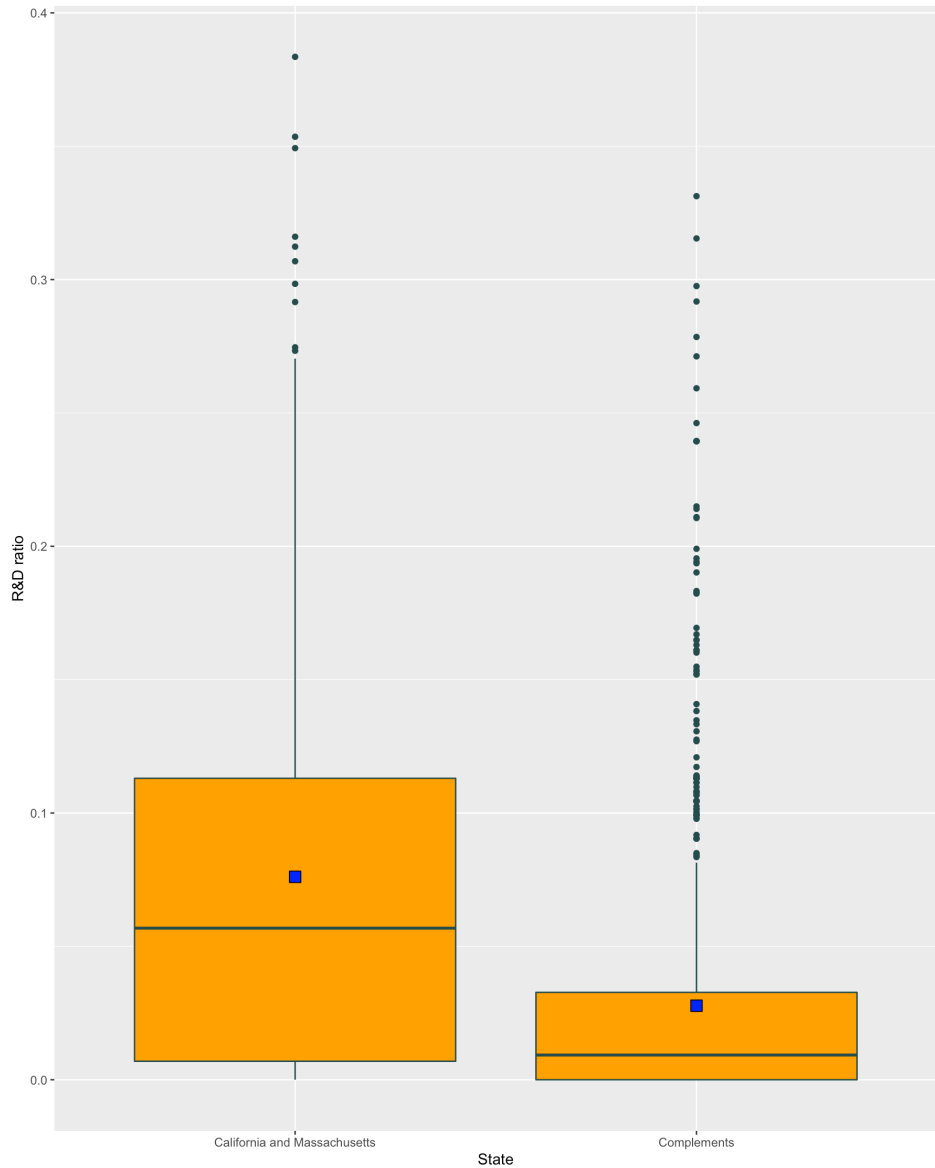


Box plot by state



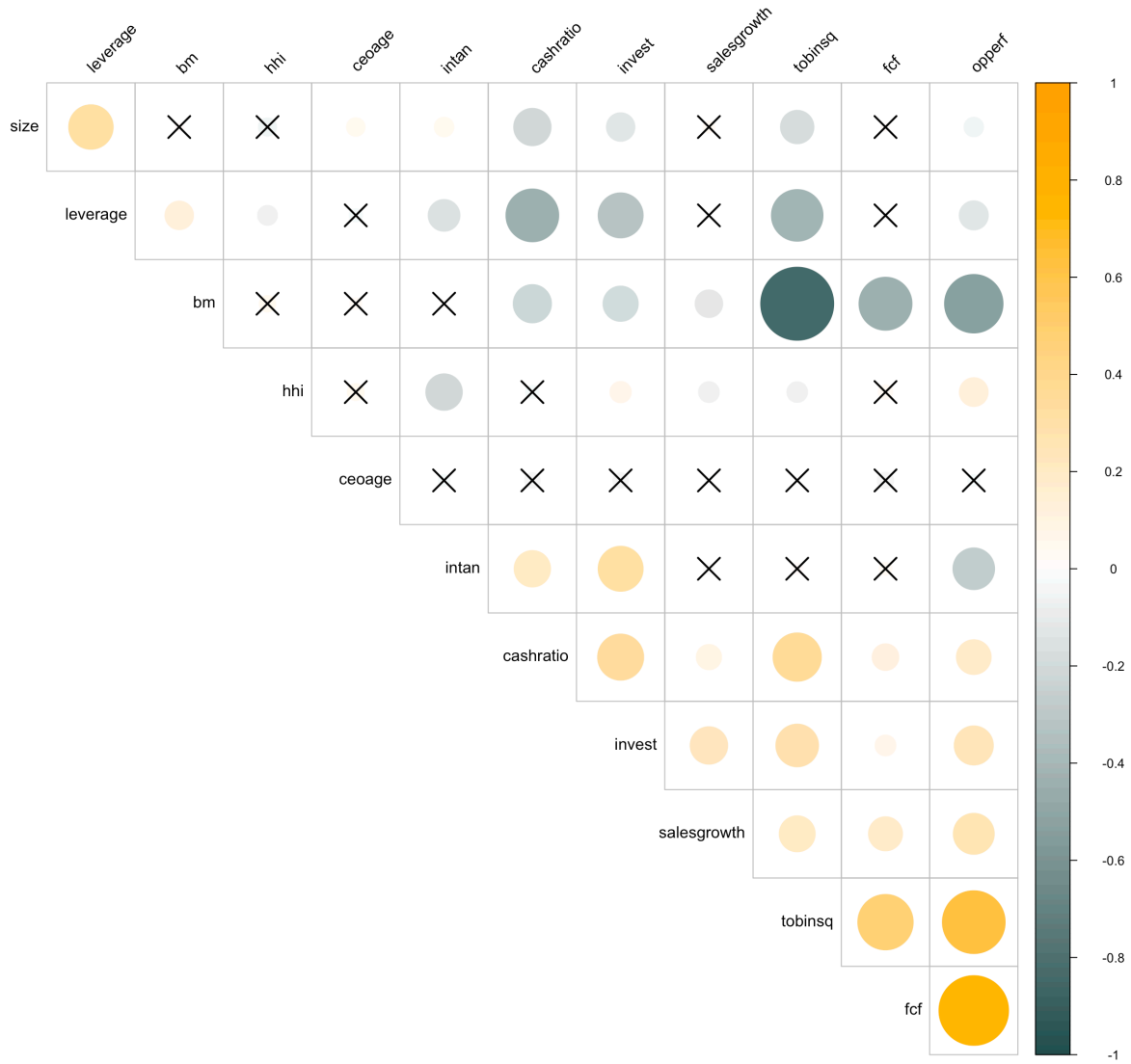
Regrouping state



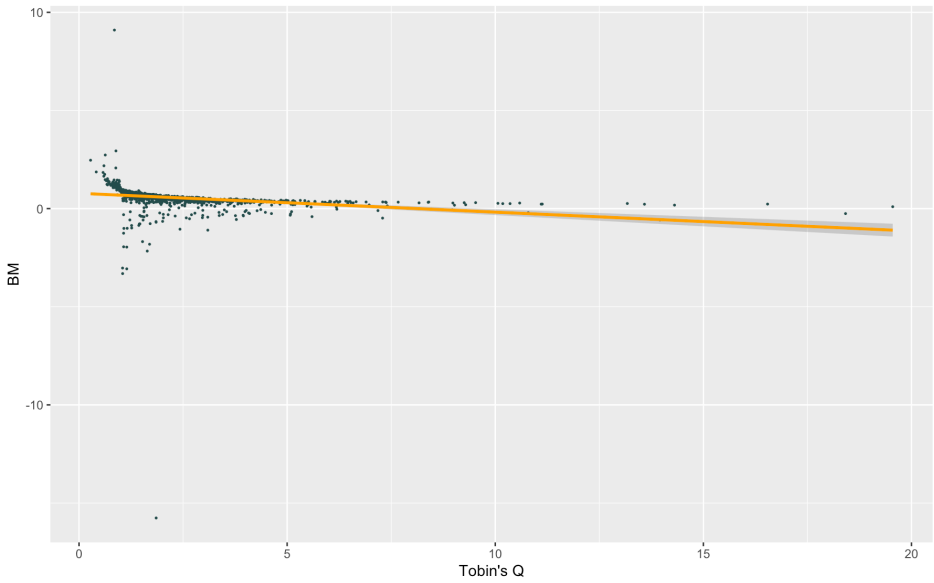


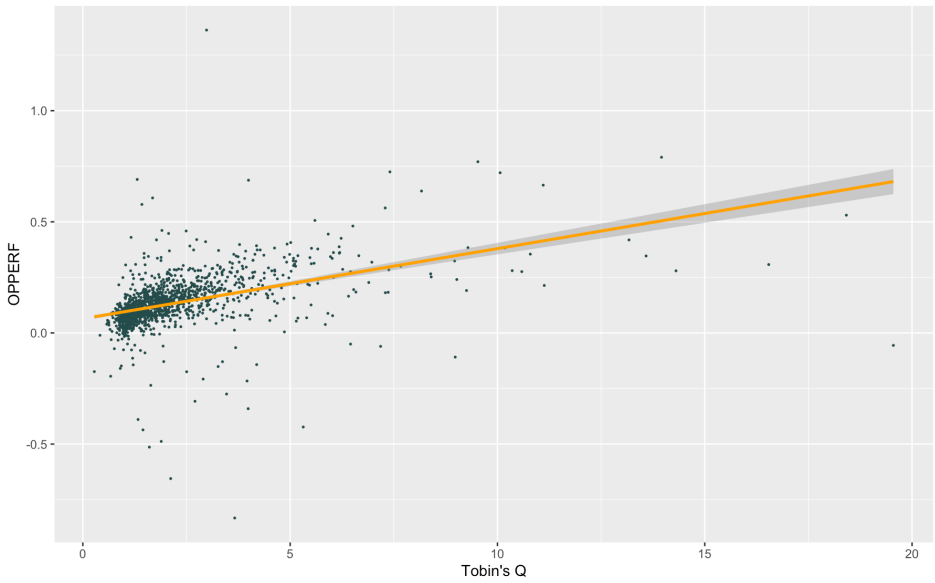
Outline

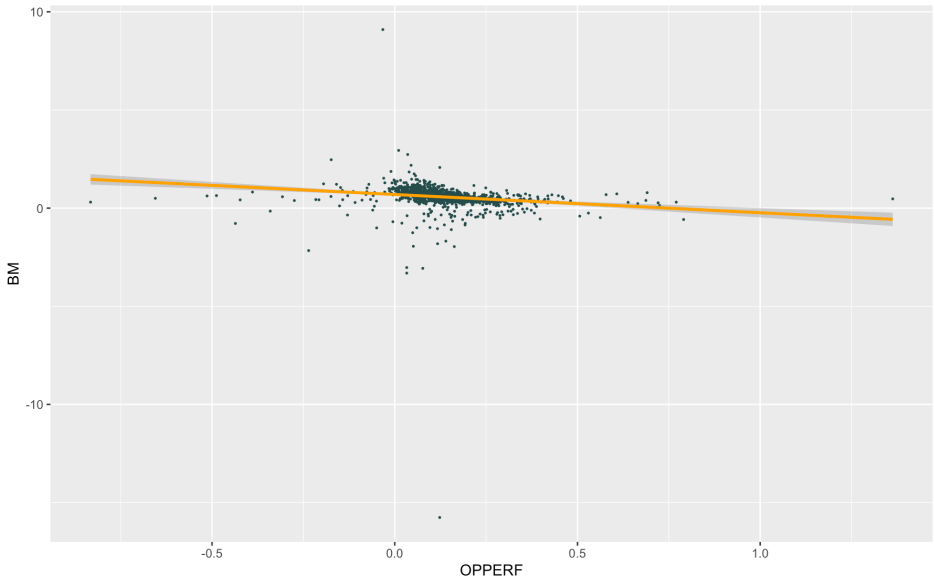
Combining Spearman correlogram with significance test

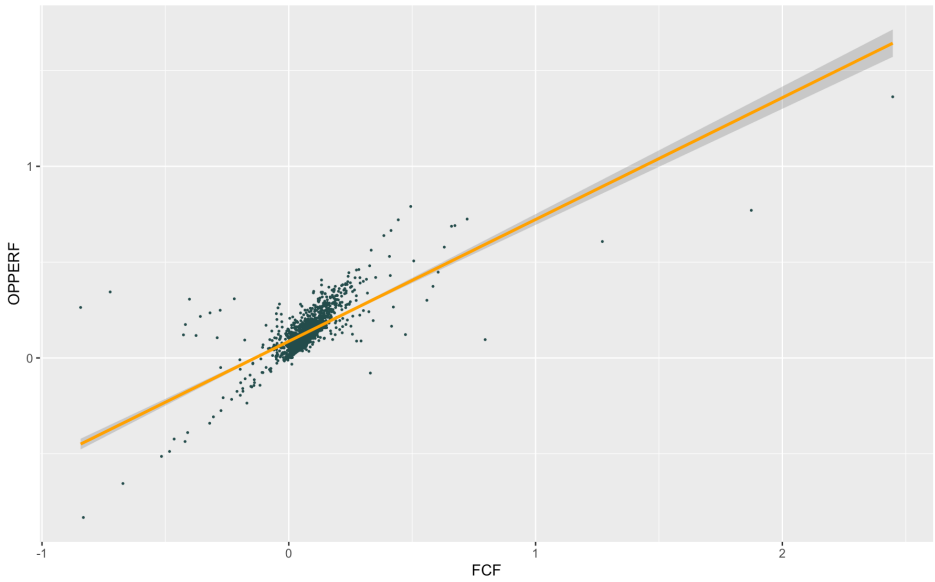


Cases with higher absolute value of correlation than 0.5





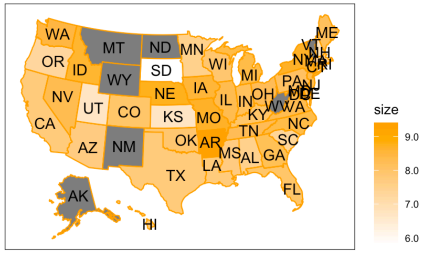




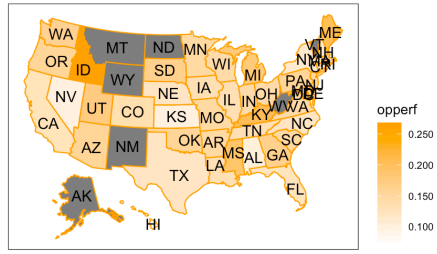
Outline

Surviving continuous variables by state (indistinguishable)

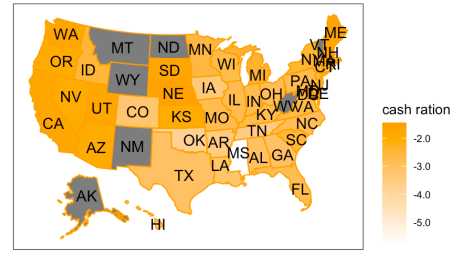
Size



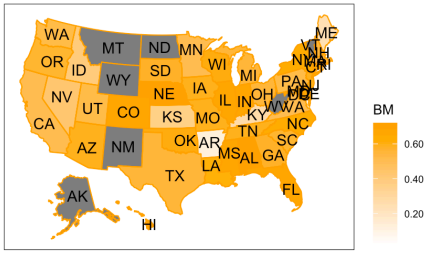
Operating performance



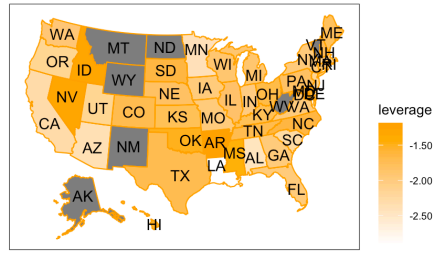
Cash ratio



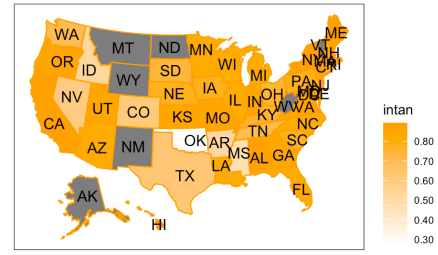
BM



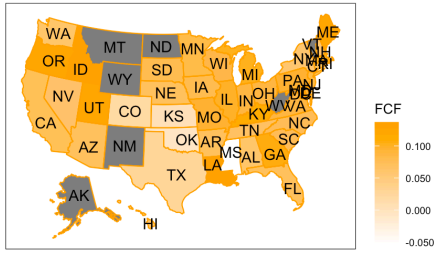
Leverage



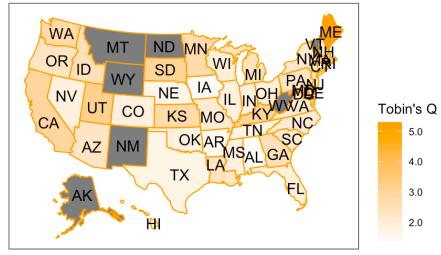
Intangibility



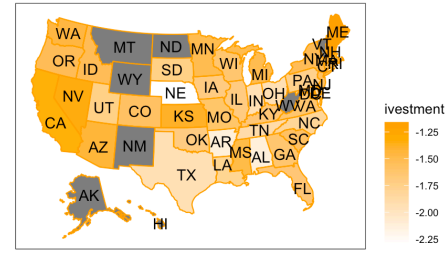
FCF



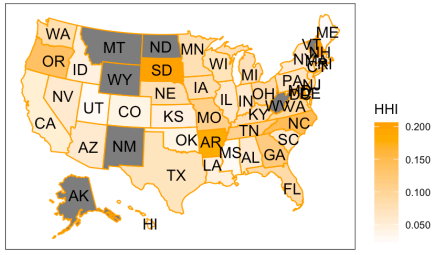
Tobin's Q



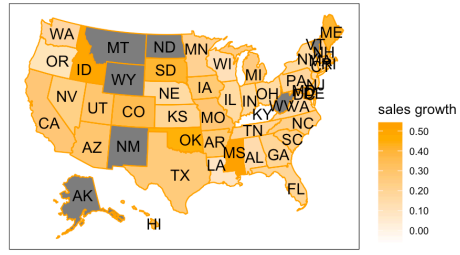
Investment



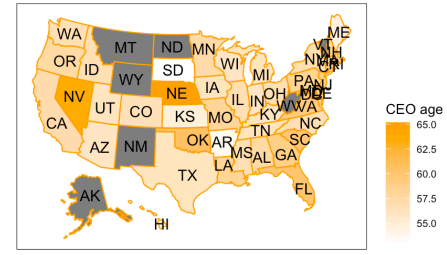
HHI



Sales growth



CEO age



Outline

Conclusion

What financial variables are relevant to the extent of the firm's R&D investment?

All financial variables used except ROA.

Does the variable that characterizes the CEO lack explanatory power for R&D investments?

Correlation analysis confirms that not all CEO characteristics are uncorrelated with R&D ratio.

If so, what characteristics correlate?

CEO age is most relevant to the degree of R&D investment.

Outline

Remaining analysis

Remaining analysis

- Check effects of normalization of R&D ratio (standard score)
- Eliminate multicollinearity based on VIF
- Select variables: FS, BS, SM
- Residual Analysis: Regression Assumptions, Influence and Outliers

References

- Chatterjee, Benmelech, E., & Frydman, C. (2015). Military ceos. *Journal of Financial Economics, 117*(1), 43-59.
- Barker III, V. L., & Mueller, G. C. (2002). CEO characteristics and firm R&D spending. *Management Science, 48*(6), 782-801.
- Serfling, M. A. (2014). CEO age and the riskiness of corporate policies. *Journal of Corporate Finance, 25*, 251-273.
- Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*. John Wiley & Sons.
- Kronmal, R. A. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 156*(3), 379-392.
- Kang, H. (2013). The prevention and handling of the missing data. *Korean journal of anesthesiology, 64*(5), 402.