Finance 701

Data Analysis

Connor Robert McDowall

April 22, 2021

Contents

List of Figures

List of Tables

1 Quantitative Analysis Exercise

This assignment was completed using a combination of technologies. An installation of Oracle Virtual Box enables remote access to SAS 9.4v. The data analysis investigation takes place in Jupyter Notebooks, implemented in Python. The remote SAS session is sent procedures using the Python API SASPy.

1.1 Merge Datasets

The pandas Python module provides merging functionalities for datasets. Subsequently, this library merges the fundamentals and markets datasets, matching 'FirmID' and 'Year' for both sets. Subsection 3.2 shows the implementation of merging datasets in Python.

1.2 Variable Creation

Additionally, the pandas Python module creates new variables by using arithmetic expressions to manipulate existing variables. Subsection 3.2 shows the creation of the following six variables:

1. CashRatio
2. DARatio
3. ROA
4. DivPayer
5. LnAssets
6. LnAge

Table 1 displays the output generate from FirmID: 4115, Year: 2005.

Variable	Value
CashRatio	0.189053
DARatio	0.053987
ROA	0.177876
DivPayer	1.0
LnAssets	25.490257
LnAge	4.110874

Table 1: FirmID: 4115, Year: 2005

1.3 Summary Statistics

Table 2 displays summary statistics as required.

Variable	N	Mean	Std Dev	Minimum	25th %	Median	75th %	Maximum
Age	156823	15.022854	15.175470	1.000000	4.000000	10.000000	20.000000	94.000000
Assets	156823	1768.782628	9956.925993	0.042000	29.775000	115.984000	565.156000	551669
Cash	156823	171.730476	1584.892751	-1.086000	1.866000	9.693000	52.101000	133819
EBITDA	156823	244.714791	1550.325207	-21913	1.317000	11.582000	72.333000	81730
Equity	156823	691.163110	4206.022395	-11476	13.208000	54.865000	253.472000	201934
Liabilities	156823	1077.619518	6232.627431	0	11.216000	47.772000	281.729000	349735
LTDebt	156823	436.124551	2542.605050	0	0.408000	9.178000	99.278000	173113
MarketCap	156823	1984.577322	13783	-1974.213333	13.228135	91.573208	547.917896	1011987
NetIncome	156823	86.575108	890.436267	-98696	-0.750000	2.832000	22.176000	59531
Return	156823	0.130102	0.728716	-0.996004	-0.242077	0.034090	0.329531	37.079998
Revenue	156823	1539.436559	8994.352077	-1964.999000	30.551000	127.136000	590.890000	511729
CashRatio	156823	0.162607	0.201545	-0.022170	0.027133	0.079384	0.216397	1.000000
DARatio	156823	0.180858	0.199670	0	0.012416	0.138444	0.280910	7.558009
DivPayer	156823	0.416693	0.493013	0	0	0	1.000000	1.000000
ROA	156823	0.065083	0.496958	-134.238095	0.051005	0.118900	0.177914	3.253335

Table 2: Summary Statistics

1.4 Correlation Matrix

Table 3 displays correlation matrix as required.

	Age	MarketCap	Return	CashRatio	DARatio	ROA
Age	1.00000	0.21826 <.0001	0.00490 0.0521	-0.16356 <.0001	0.04874 <.0001	0.07925 <.0001
MarketCap	0.21826 <.0001	1.00000	0.00580 0.0217	-0.00788 0.0018	0.01902 <.0001	0.02983 <.0001
Return	0.00490 0.0521	0.00580 0.0217	1.00000	0.05470 <.0001	-0.03201 <.0001	0.08121 <.0001
CashRatio	-0.16356 <.0001	-0.00788 0.0018	0.05470 <.0001	1.00000	-0.28855 <.0001	-0.19027 <.0001
DARatio	0.04874 <.0001	0.01902 <.0001	-0.03201 <.0001	-0.28855 <.0001	1.00000	0.01129 <.0001
ROA	0.07925 <.0001	0.02983 <.0001	0.08121 <.0001	-0.19027 <.0001	0.01129 <.0001	1.00000

Table 3: Correlation Matrix

1.5 T-Test Output

The T-Test analysis is split into three parts:

1. Hypotheses
2. Results
3. Summary

1.5.1 T-Test Hypotheses

The purpose of the t-test was to explore the following hypotheses, assuming unequal variances between the two groups:

H_0 : There is no difference in the cash ratio between dividend paying & non paying firms
H_1 : There is a difference in the cash ratio between dividend paying & non paying firms

1.5.2 T-Test Results

Tables 4, 5, 6, 7 display the results from this t-test procedure.

DivPayer	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		91476	0.2048	0.2347	0.000776	-0.0222	1.0000
1		65347	0.1035	0.1199	0.000469	-0.00732	0.9956
Diff (1-2)	Pooled		0.1013	0.1953	0.00100		
Diff (1-2)	Satterthwaite		0.1013		0.000907		

Table 4: T-Test Procedure: Part I - Variable Statistics I

DivPayer	Method	Mean	95% CL	Mean	Std Dev	95% CL	Std Dev
0		0.2048	0.2033	0.2063	0.2347	0.2336	0.2358
1		0.1035	0.1026	0.1044	0.1199	0.1192	0.1205
Diff (1-2)	Pooled	0.1013	0.0993	0.1033	0.1953	0.1946	0.1959
Diff (1-2)	Satterthwaite	0.1013	0.0995	0.1031			

Table 5: T-Test Procedure: Part II - Variable Statistics II

Method	Variances	DF	t Value	Pr > —t—
Pooled	Equal	156821	101.29	<.0001
Satterthwaite	Unequal	143669	111.72	<.0001

Table 6: T-Test Procedure: Part III - Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	91475	65346	3.83	<.0001

Table 7: T-Test Procedure: Part IV - Equality of Variances

1.5.3 T-Test Summary

A pvalue of <. 0.0001 indicates a very strong statistically significant difference in the mean of CashRatios (0.1013) for firms who pay dividends compared to those who do not pay dividends, suggesting we can reject the null hypothesis of no difference in favour of the alternative hypothesis ,assuming unequal variance using the Satterthwaite method.

1.6 Annual Cash Ratio Chart

Figure 1 displays fluctuations in the annual cash ratio across all firms from 1950 to 2019. Fluctuations over time are displayed from two perspectives:

1. Equally weighted average cash ratio across all firms by year.
2. Value weighted average cash ratio across firms according to firms' market capitalisation each year.

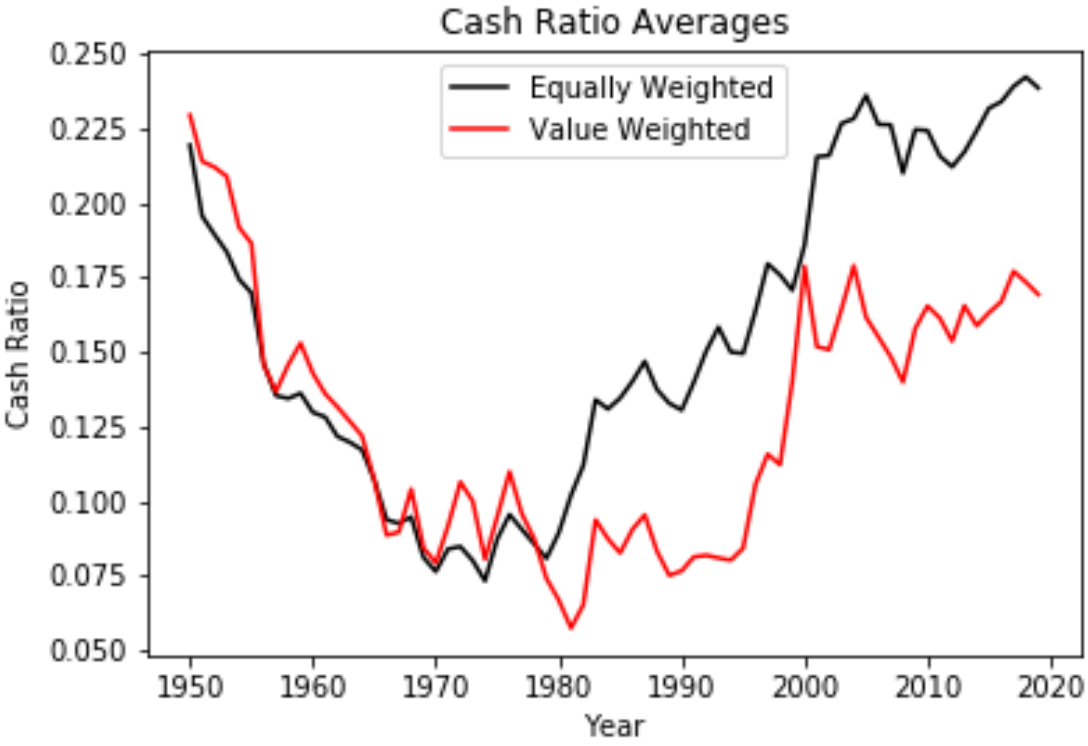


Figure 1: Average Cash Ratios: Equal & Value Weighted

1.7 Lag Return Calculation

The Lag Return variables are calculated by implementing the methodology described in subsection 3.2. The implementation correctly identifies 15,964 missing values, 14007 from initial listings, 1957 from entries related to delisting/re-listing events or missing observations. The target firm (FirmID: 5540, Year:2002) has a LagReturn of 0.044689 (6.d.p) as required.

1.8 Regression Analysis

Table 8 displays the co-efficient estimates, associated t-statistics, statistical significance indicators, R^2 values and the number of observations used in the sample to measure the co-efficients.

1.8.1 OLS Regression Notes

1. IGFE = SIG Industry Group Fixed Effects
2. YFE = Year Fixed Effects
3. Adjusted R^2 reported in Models I-II
4. Normal co-efficient & values reported in Models I-II, not values & co-efficients under the Heteroscedasticity Consistent section of the SASS Output.

1.8.2 OLS Regression Results

Model	Intercept (β_0)	LnAssets (β_1)	LnAge (β_2)	DARatio (β_3)	DivPayer (β_4)	LagReturn (β_5)	IGFE	YFE	R^2 (Adj)	Obs (# Used)
I										
β	0.44859	-0.01525					No	No	0.0264	156823
T-Stat	101.62	-65.21								
$Pr > \|t\|$	$< .0001$	$< .0001$								
II										
β	0.26766	0.00065555	-0.01601	-0.26894	-0.08487	0.01017	No	No	0.1424	140859
T-Stat	59.70	2.47	-26.60	-107.77	-79.60	15.40				
$Pr > \|t\|$	$< .0001$	0.0135	$< .0001$	$< .0001$	$< .0001$	$< .0001$				
III										
β	0.5253053793	-.0031356789	-.0066426725	-.1971642737	-.0293596941	0.0081145210	Yes	No	0.369559	140859
T-Stat	89.13	-11.96	-11.77	-83.87	-27.29	14.21				
$Pr > \|t\|$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$				
IV										
β	0.6539506648	-.0142617342	-.0235531355	-.2352176297	-.0327617500	0.0112630754	No	Yes	0.216192	140859
T-Stat	101.34	-49.24	-40.28	-97.29	-28.00	17.17				
$Pr > \|t\|$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	$< .0001$				
V										
β	0.7228296126	-.0131141677	-.0159296700	-.1813118405	-.0021391321	0.0088674634	Yes	Yes	0.401624	140859
T-Stat	105.26	-45.55	-28.06	-78.55	-1.89	15.34				
$Pr > \|t\|$	$< .0001$	$< .0001$	$< .0001$	$< .0001$	0.0588	$< .0001$				

Table 8: OLS Regression Results

1.8.3 OLS Regression Summary

β_3 in Model III is -.1971642737 with a t-stat of -83.87 and p-value $<$ .0001. This co-efficient relates to the effect the proportion of debt to long term assets (DARatio) has on the cash ratio, the firms proportion of cash holdings to total assets. The co-efficient reports a 1 unit change in the DARatio will result in a -0.1971.. unit change in the level of cash holdings to total assets. This is both statistically and economically significant. Firstly, the analysis reports a large t-statistics (-83.87) and low p-value ($<$ .0001), giving strong evidence to reject the null hypothesis of DARatio having no effect on CashRatio in favour of the alternative hypothesis. Secondly, this analysis aligns with financial concepts. Greater proportions of debt (DARatio) would require larger debt obligations in the form of principal and interest payments, leading to increasing cash holding diminishing cash holding.

β_4 in Model III is -.0293596941 with a t-stat of -27.29 and p-value of $<$ .0001. This co-efficient reports the presence of a dividend payer (dummy variable) results in a -0.0293596941 unit change in the level of cash holdings to total assets. This is both statistically and economically significant. Firstly, the analysis reports a large t-statistics (-27.29) and low p-value ($<$ .0001), giving strong evidence to reject the null hypothesis of DivPayer having no effect on CashRatio in favour of the alternative hypothesis. Secondly, this analysis aligns with financial concepts. The presence of a dividend payer will make distributions to equity holders, likely in the form of cash, leading to a reduction of cash holdings as a proportion of total assets.

The purpose of including industry and fixed effects is to prevent confounding analysis by control for observed and unobserved time-invariant variables. The most attractive feature of fixed effect models is controlling for all stable characteristics of the entity belonging to a particular industry or year. In summary, each entity becomes their own control and enables the ability to make causal inferences from non-experimental data, mainly the other time-varying variables (Allison, 2006).

2 Corporate Cash Holdings Research

2.1 Introduction

Mark Leary and John Graham conducted time-series, cross-sectional analysis on the fluctuations of corporate cash holdings by organisations. The authors found two main drivers of average and aggregate cash are current-period profitability realisations and fund current investment. The contributing causes to cash holding fluctuations vary from different reasons ways: Productivity and real GDP growth. Changes to cash management efficiencies. The entrapment of cash offshore from repatriation taxes. Initial public offering effects. Macroeconomic effects and Regulation (Graham and Leary, 2018). This paper continues to explore the impact of corporate cash holdings have on firm characteristics.

2.2 Research Question

The subsequent research question will be explored using quantitative methods:

Do Corporate Cash Holdings Affect Asset Profitability?

A succinct literature review informs the question, hypotheses and methodology.

2.3 Literature Review

Michael and Rong examined the cross-sectional variation in the marginal value of corporate cash holdings that arises from differences in corporate financial policy. Their analysis found cash is most highly valued by shareholders of firms with low levels of cash holdings, low leverage and constraints in accessing financial markets (Faulkender and Wang, 2006). Shareholders demand profitability. Therefore, we see evidence low cash holdings, potentially contributed to investing in assets, improves shareholder returns and satisfaction.

Laurent expanded on this research by finding firm holdings strategically influence the product market outcomes as follows:

1. Relative-to-rival cash reserves lead to systematic future market share gains obtained at the expense of industry rivals.
2. Competitive effect of cash contributes to increased firm value and operating performance.
3. Cash-rich firms partly gain shares in their product market by drawing down reserves to invest in fixed capital, research and development.

The analysis contributes three key insights. Firstly, the strategic value of cash is substantial. Secondly, the interconnectedness of finance and product market. Lastly, the interactions between firms' financial and real decisions are more complex than a simple association between debt financing and competitive strategies (Fresard, 2010).

Michael and Rong highlight low levels of cash holdings and their association with shareholder satisfaction. Laurent highlights the deployment of cash in investing in fixed assets to strategically influence product market outcomes to benefit shareholders. Theoretically, low cash holdings should partially correspond to greater capital expenditure to improve

the profitability of assets and satisfy shareholders. Subsequently, this quantitative analysis will explore the relationship between corporate cash holdings and return of assets, using the following hypotheses, to answer the aforementioned research question.

2.4 Hypotheses

2.4.1 H_0: Null Hypothesis

There is no difference in the cash ratio between firms with low ROA and high ROA.

2.4.2 H_1: Alternative Hypothesis

There is a difference in the cash ratio between firms with low ROA and high ROA.

2.5 Data

CRSP and Compustat provide the data source for this assignment. The granularity of the data is annual, covering the time period 1950 to 2019. The dataset is the combination of the following sets:

1. **Fundamentals**: Firm fundamentals and financial reporting data
2. **Market**: Stock return and other trading information

The merged dataset contains 156,823 firm-year observations.

2.6 Methodology

The analysis is relatively rudimentary, exploring a simple relationship. The methodology is as follows:

1. Re-adjust ROA for Net Income instead of EBITDA
2. Provision of descriptive statistics
3. Correlation analysis on ROA/ROE related data points
4. Time-series analysis on CashRatio, asset composition and ROA.
5. Regression analysis on ROA to predict CashRatio, controlling for industry and year fixed effects.

2.7 Results & Discussion

2.7.1 ROA Readjustment

There is some rationale adjusting return on assets (ROA) to reflect net income over assets opposed to EBITDA over assets. ROA is a measure of profitability based on the utilization of assets. The ratio measures the efficiency of a manager, investor or analyst at utilising assets to generate returns. Earnings before interest, tax, depreciation and amortization (EBITDA) is not a complete reflection of profit. Depending on the depreciation and amortization models, expenses must be made to accurately represent the value of existing assets. Additionally, obligations to debt holders and taxation entities must be met. Net Income is a better proxy for profitability as these obligations are addressed, leaving a more

accurate representation of the profit available to distribute to shareholders. Subsequently, a new return on assets variable was created. $ROANI = \frac{NetIncome}{Assets}$. Additionally, $ROE = \frac{EBITDA}{Equity}$, $ROENI = \frac{NetIncome}{Equity}$ were created for completeness.

2.7.2 Descriptive Statistics

Table 11 in Appendix subsection 3.1 displays summary statistics as required as the table was too wide to display practically. The summary statistics reveal some interesting insights. Firstly, there is a large discrepancy between the calculated ROA and ROE using both EBITDA and Net Income. This finding further confirms the need to adjust ROA for Net Income. Secondly, ROENI and ROE have some gross outliers as both the minimum and maximum values are excessive. These are most likely transcription errors. These errors would normally be omitted. However, both ROE and ROENI are not explored further in this instance so data cleansing is not required. Our main variable of interest in ROANI.

2.7.3 Correlation Matrices

The following correlation matrix was generated to explored the relationships between CashRatio, ROA, ROE, ROANI and ROENI. Table 9 displays correlation matrix as required.

	CashRatio	ROA	ROE	ROENI	ROANI
CashRatio	1.00000	-0.19027 <.0001	-0.00362 0.1516	0.00039 0.8757	-0.11505 <.0001
ROA	-0.19027 <.0001	1.00000	0.00150 0.0217	-0.00071 0.7795	0.82293 <.0001
ROE	-0.00362 0.0521	0.00150 0.5518	1.00000	0.21759 <.0001	-0.00110 0.6619
ROENI	0.00039 0.8757	-0.00071 0.7795	0.21759 <.0001	1.00000	-0.00108 <.0001
ROANI	-0.11505 <.0001	0.82293 <.0001	-0.00110 0.6619	-0.00108 0.6685	1.00000

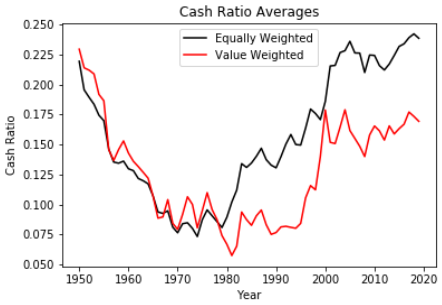
Table 9: Corporate Cash Holdings Correlation Matrix

The correlation matrix conveys two key concepts. Firstly, any variable correlated with ROE or ROENI is statistically insignificant, implying there are no linear correlations when either CashRatio, ROA of ROANI is paired with either ROE or ROENI. Secondly, the correlations between CashRatio and either ROA or ROANI are statistically significant at less than the 1% level, implying the null hypothesis of no linear correlations can be rejected in favour of the alternative hypothesis of a negative linear correlation existing (-0.19 and -0.12 for ROA and ROANI respectively). Lastly, both combinations of ROA/ROANI and ROE/ROENI are highly correlated but this is expected.

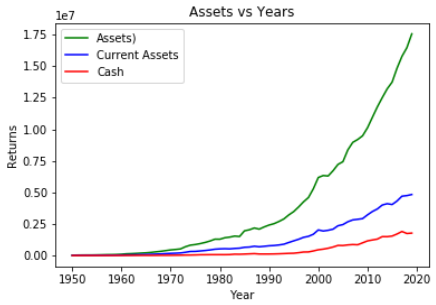
2.7.4 Time-series Analysis

Figure 2: Time-series Corporate Cash Holdings vs ROA (Net Income)

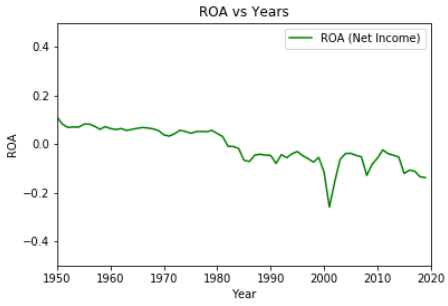
(a) CashRatio



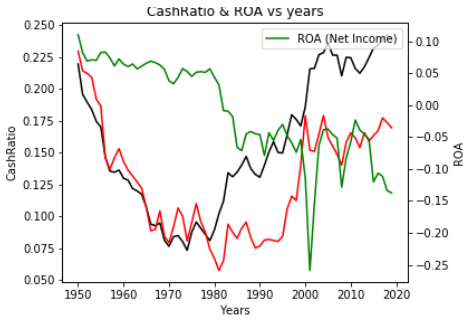
(b) Increase in Assets



(c) ROA vs Years



(d) CashRatio and ROA vs Year



The time-series analysis offers some interesting insights. Firstly, figure 2 (a) shows the upward trend in corporate cash holdings as the CashRatio increases over time from both equally-weighted and value-weighted perspectives. Secondly, figure 2 (b) shows a near exponential increase in aggregate assets over time with both current assets and cash increasing at a slower rate. The cause behind this observation could be the maturation of financial markets, increasing the number of entities and size of assets listed in the market. Thirdly, figure 2 (c) shows return on assets ($ROA = \frac{\text{Net Income}}{\text{Assets}}$) decreasing over time. Lastly, overlapping equally-weighted and value-weighted portfolio with ROA shows as corporate cash holdings begins to increase around 1980, ROA begins to substantially decline on a relative basis. Additionally, ROA took two large exogenous shocks around 2000 and 2008. These shocks were most likely the cause of the internet bubble bursting and global financial crisis respectively.

2.7.5 Regression Analysis

The following regression model was used to empirically test the aforementioned null hypothesis:

$$\text{Cash Ratio}_{i,t} = \beta_0 + \beta_1 \times \text{ROANI}_{i,t} + \text{SIG Industry Fixed Effects} + \text{Year Fixed Effects} + \epsilon_{i,t}$$

1. Net Income instead of EBITDA was used based on the results from the correlations and theoretical arguments.

2. SIG industry effects and year effects are incorporated to control for observed and unobserved time-invariant variables, preventing confounding analysis as explained in subsubsection 1.8.3.

The regression results are as follows:

Model	Intercept (β_0)	ROANI (β_1)	IGFE	YFE	R^2 (Adj)	Obs (# Used)
I						
β	0.4094039536	-0.0123209222	Yes	Yes	0.345139	156823
T-Stat	98.23	-20.34				
$Pr > \|t\|$	$< .0001$	$< .0001$				

Table 10: ROA vs CashRatio OLS Regression

The results deliver several key insights. Firstly, both the co-efficients are statistical significant at less than the 1% level with large t-statistics and small p-values. Subsequently, there is evidence to reject the null hypothesis of no difference in Cash Ratio between with firms with low and high ROA in favour of the alternative. The β_1 co-efficient of implies for each 1 unit change in ROANI, there is a -0.0123... decrease in CashRatio, a proxy for corporate cash holdings. The R^2, coefficient of determination, implies 34.51% of variance in CashRatio is explained by ROANI, SIG Industry and Year fixed effects. Subsequently, the empirical analysis informs the research question.

2.8 Research Design Issues

However, a few research design issues must be addressed.

1. Investigate the presence of endogeneity issues associated with potentially missing time-variant variables.
2. The methodology looks at the relationship between corporate cash holdings and ROA from a high level perspective. Investigate with greater granularity across industry for industry-specific insights.

2.9 Conclusion

This paper explored the impact corporate cash holdings have on firm characteristics, in particular return on assets. Theoretical frameworks supported the notion larger corporate cash holdings should reduce asset efficiency and profitability, agreeing with existing literature. Observations from time-series analysis agree with both the theoretical frameworks and literature. Empirical analysis in the form of regressions and correlations align with the aforementioned findings. There is sufficient evidence to reject the Null Hypothesis of no difference in CashRatio between firms with low and high ROA. It appears corporate cash holding do affect asset profitability, answering the research question. However, the research issues raised must be investigated in future research.

References

Allison, P. (2006). Fixed effects regression methods in sas.

Faulkender, M., & Wang, R. (2006). Corporate financial policy and the value of cash. *The Journal of Finance*, *61*(4), 1957–1990. doi:https://doi.org/10.1111/j.1540-6261.2006.00894.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2006.00894.x

Fresard, L. (2010). Financial strength and product market behavior: The real effects of corporate cash holdings. *The Journal of Finance*, *65*(3), 1097–1122. doi:https://doi.org/10.1111/j.1540-6261.2010.01562.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01562.x

Graham, J. R., & Leary, M. T. (2018). The Evolution of Corporate Cash. *The Review of Financial Studies*, *31*(11), 4288–4344. doi:10.1093/rfs/hhy075. eprint: https://academic.oup.com/rfs/article-pdf/31/11/4288/26132510/hhy075.pdf

3 Appendix

3.1 Corporate Cash Holding Summary Statistics

Variable	N	Mean	Std Dev	Minimum	25th Pctl	Median	75th Pctl	Maximum
Assets	156823	1768.782628	9956.925993	0.042000	29.775000	115.984000	565.156000	551669
Cash	156823	171.730476	1584.892751	-1.086000	1.866000	9.693000	52.101000	133819
CurAssets	156823	558.744661	3041.599403	0	16.025000	58.740000	240.584000	175552
EBITDA	156823	244.714791	1550.325207	-21913	1.317000	11.582000	72.333000	81730
Equity	156823	691.163110	4206.022395	-11476	13.208000	54.865000	253.472000	201934
NetIncome	156823	86.575108	890.436267	-98696	-0.750000	2.832000	22.176000	59531
CashRatio	156823	0.162607	0.201545	-0.022170	0.027133	0.079384	0.216397	1.000000
ROA	156823	0.065083	0.496958	-134.238095	0.051005	0.118900	0.177914	3.253335
ROE	156821	0.264617	59.648117	-16122	0.096059	0.250903	0.379391	15887
ROANI	156823	-0.040524	0.697813	-173.261905	-0.020275	0.039604	0.078563	21.789094
ROENI	156821	-0.312381	103.424944	-31837	-0.028610	0.088947	0.154666	7770.333333

Table 11: Corporate Cash Holding Summary Statistics

3.2 Jupyter Notebook

The assignment submission includes both the jupyter notebook and html describing the data analysis. Additionally, the file can be provided on request.