Determinants of Corporate Cash Holdings and Dividend Payout Decision : Evidence from Japan

Group 5:

WEIJIA WANG (1A172G39-7)

YIJU CHEN (1A162G07-9)

YIFAN CAI (1A173G01-1)

2020/12/21

1. Introduction : Important Implications of Cash Holding and Dividend Payout

1.1 Cash Holding

According to Besley(2005), cash holding of a company is referred to as the cash in hand or readily available for investment in physical assets and to distribute to investors.

The benefits of cash holding include the following : it helps reduce the likelihood of financial distress; it allows the pursuance of investment policy when financial constraints are met; it contributes to minimization of the costs of raising external funds or liquidating existing assets. Additionally, cash holding reduces the pressure to perform well and allows managers to invest in projects that they are mostly interested in but may not represent the best interest of shareholders (Ferreira & Vilela, 2004).

Generally speaking, there are three motives of companies' cash holding, i.e transaction motive, precautionary motive and future requirements. Firstly, cash serves a buffer between a company's daily cash inflows and outflows that are not synchronized. In the sense of precautionary motive, companies hold cash to mitigate the volatility of operating earnings, the fluctuations in the cash balances, or for the increasing uncertainties in the market. Last but not the least, cash is held to meet future payment obligations such as payment of dividend and repayment of loans.

The decisions of a company's cash holding level can therefore suggest a company's financial status, the market environment, and the company's incoming payment obligations. And such information is critical for investors when they are making investment decisions. For example, an increase in a company's cash holding level may indicate that the company is facing escalating uncertainties, which could lead to a decrease in the company's future share price. The investors can then decide not to invest in this share to avoid potential loss. By applying statistical learning models, we are able to predict the future cash holding levels of a company, which provides useful information for investors to make better investment strategies.

Besides being beneficial for investors, applying statistical learning methods to the prediction of cash hold levels is also helpful for companies. Through statistical learning, companies are able to see which predictors are more significant in predicting the future cash holding levels of their competitors. That is to say, executives of a company can identify what motivates their competing companies' cash levels, which can help them make more informed decisions for their own company.

1.2 Dividend Payout Decision

A dividend payment is the distribution of a company's profits to its shareholders and is determined by the company's board of directors. Periodic dividend payments provide evidence of a company's stability, confidence in earnings growth, and sufficient profitability to fund future expansion. Therefore, a company's dividend decision serves as another piece of useful information for investors to make investment decisions. If the company is predicted to offer dividend payments in the next year, investors will be more inclined to buy this company's share. And such predictions can be effectively made by statistical learning models. Through classification models such as KNN, logistic regression and LDA, we are

able to predict whether a company is going to make a dividend payment decision based on specific predictors, which will be good references for investors to make wiser decisions.

Moreover, dividend payments typically lead to money going out of a company's books forever, so they can impact the company's share price. The share price may rise on the announcement roughly by the amount of the dividend declared and then decline by a similar amount at the opening session of the ex-dividend date. Therefore, if there exists credible analysis predicting that a company will make a dividend payment in the future, investors can buy the company's share in advance to take advantage of the increase in the share's price when the dividends are paid. Again, statistical learning methods are effective in providing such credible analysis.

2. Literature Review

2.1.1 Main Determinants of Cash Holding

Researchers have made great efforts so far to explore the potential factors correlated with the cash holding levels in corporations using different sample data. The paper attempts to review the previous literature and provide an extensive account of the principle variables that have been found to determine the level of cash holding.

Firm Size (Total Asset)

The firm size has been widely found as one of the most essential determinants of cash holding in the majority of studies. Based on the assumption that larger firms are better diversified in the sense of their asset composition, which entitles them a higher possibility and ability to liquidate part of assets to obtain cash to encounter financial distress instead of being forced to keep cash holding (Bates, Kahle and Stulz, 2009). Trade-off theory also stands for this negative relationship in the sense that large firms have a lower cost of raising money in the capital market than those small ones, which means that they are more likely to

hold less cash. Amarjit Gill found firm size is negatively correlated with cash holding using the sample of 166 Canadian firms listed on Toronto Stock Exchange from 2008 to 2010.

Similar findings have been verified in the majority of studies related using different samples of corporation data. (Afza, T., & Adnan, S.M.(2007), Ferreira, M.A., & Vilela, A.S. (2004), Nguyen, P. (2005), Drobetz, W., & Grüninger, M.C. (2007), Kim, J., Kim, H., & Woods, D. (2011)) In addition, some scholars pointed out that the effect of determinants of cash holding also differs for corporations in different development stages. For growth companies, there is a negative relationship between cash holding and firm's size, level of liquid assets and short term debt. While for mature companies, the level of cash holding shows a positive relationship with their size, investment, and the amount of dividend payout and stock repurchases. (Saddour, K., 2006).

Capital Expenditure

Kim, J., Kim, H., and Woods, D. (2011) found that firms with higher capital expenditures were shown to hold less cash in the research of 125 publicly traded US restaurant firms between 1997 and 2008. Scholars also suggest diversified relationships between expenditure of specific genres — expenditure to R&D has been found positively correlated with cash holding level (Yifan Ma, 2012 ; Saddour, K. ,2006); while expenditures on acquisitions are shown to be negatively correlated.

Leverage

The ratio of long-term debt over the total asset, namely the leverage ratio, also came to be acknowledged as a determinant of cash holding of corporations.

In the study of 400 firms in 12 Economic and Monetary Union (EMU) countries for the period of 1987-2000 done by Ferreira, M.A., and Vilela, A.S. (2004), leverage is found to be

negatively related with cash holding. Drobetz and Grüninger (2007) also found the non-linear relationship between the leverage ratio and liquidity holding. This negative relationship has also been found in other literature. (Afza, T., & Adnan, S.M., 2007; Hardin III, W.G., Highfield, M.J., Hill, M.D., & Kelly, G.W., 2009; Rizwan, M.F., & Javed, T., 2011; Yifan MA, 2012; Khaoula SADDOUR, 2006)

Company Age

Though the age of a company is not directly related with its degree of growth, it is correlated in directly. And in terms of growth opportunity, younger corporations with better growth space and opportunities have been found to positively correlate with cash holding level. Opler, T., Pinkowitz, L., Stulz, R., and Williamson, R.(1999) observed that firms with strong growth opportunities and riskier cash flows hold a relatively high ratio of cash to total non-cash assets.

Sales & Earnings

Yet another factor that influenced the corporate cash holdings is its benefitablity, which can be reflected from the data of earnings before interest, taxes, depreciation and amortization. Sales, as another variable indirectly related with profitability, is not covered by previous literature as crucial explanatory variables though, we included it in our model to observe its effect and expect a similar effect as earning's. Opler et al. found that firms do well tend to accumulate more cash in their hand in the study of 1048 publicly traded US firms from 1971 to1994. Megginson, W.L., and Wei, Z. (2010) looked into a group of China's share-issue privatized firms from 1993-2007 and found that smaller, more profitable and high growth firms hold more cash.

Dividend Payout

Last but not least, besides the ones mentioned above, studies have also indicated that another independent we are going to discuss — dividend payout decision (or its ratio) is closely correlated with cash holding. Drobetz and Grüninger (2007) found that dividend payments and operating cash flows are positively related to cash holding using the sample of 156 Swiss non-financial firms between 1995 and 2004. In the study of 125 publicly traded US restaurant firms between 1997 and 2008, Kim et al. (2011) argued that firms that decided to pay dividends are supposed to have a smaller level of cash holding. Another comparative approach of corporations in Brazil, Russia, India and China and firms from the US and the UK showed that dividend policy does play a significant role in determining cash holding.

2.1.2 Other Determinants of Cash Holding

Though the variables we included in our models are unavoidably limited to few main variables, we also take an account of other variables found to be significant determinants in other studies and summarise in this section.

TABLE1 : Other Explanatory Variables and Their Effect on Cash Holding in Existing Literature

Explanatory Variables	Effect	Literature
Market-to-book ratio	-	Amarjit Gill (2012) ; Afza, T., & Adnan, S.M. (2007).
	+	Rizwan, M.F., & Javed, T. (2011)
Cash flow ratio	+	Afza, T., & Adnan, S.M. (2007); Ferreira, M.A., & Vilela, A.S. (2004) ; Amarjit Gill (2012); Drobetz, W., & Grüninger, M.C. (2007); Rizwan, M.F., & Javed, T. (2011); Tariq Alzoubi (2013); Yifan Ma (2012); Saddour, K. (2006).
Working capital	-	Afza, T., & Adnan, S.M. (2007); Megginson, W.L., & Wei, Z. (2010); Rizwan, M.F., & Javed, T. (2011); Yifan Ma (2012);
	+	Amarjit Gill (2012) ;

CEO duality	+	Amarjit Gill (2012) ; Drobetz, W., & Grüninger, M.C. (2007).
Risk of cash flow	+	Opler, T., Pinkowitz, L., Stulz, R., & Williamson, R. (1999); Nguyen, P. (2005); Saddour, K. (2006).
Asset Liquidity	-	Ferreira, M.A., & Vilela, A.S. (2004); Kim, J., Kim, H., & Woods, D. (2011)
	+	Tariq Alzoubi (2013)
Debt Ratio	-	Nguyen, P. (2005); Megginson, W.L., & Wei, Z. (2010)
Operation Funds	-	Hardin III, W.G., Highfield, M.J., Hill, M.D., & Kelly, G.W. (2009)
Board Size	Non-significant	Drobetz, W., & Grüninger, M.C. (2007).
	+	Amarjit Gill (2012) ;
Asset Tangibility	-	Drobetz, W., & Grüninger, M.C. (2007).
Trade Credit	-	Saddour, K. (2006).

2.2.1 Determinants of Dividend Payout

There has been amountful of discussions about dividend payout decisions and it's still a pretty controversial topic in the area of corporate finance. A large number of researchers developed diversified theories or hypotheses and countless empirical evidence of correlates of dividend payout decision but the heterogeneity of results shown in different groups of firms with different characteristics the fact that researchers haven't agreed on a single point suggest that though discussed comprehensively, the issue is still unresolved and open for further discussion just as how Black(1976) once argued — "the harder we look at the dividend picture, the more it seems like a puzzle".

Dividend payout decisions had been taken for granted for a long period of time based on the acknowledgement that a company's value is supposed to increase together with the increase of dividend payout before Miller and Mpdigliani's paper (1961), therefore, shares of high dividend companies should be sold at a high price (Frankfurter and Wood 1997) based on the idea that current value of the company is discounted form of company's future profit.

Miller and Modigliani (1961) firstly raised the question towards “common sense” and suggested that the value of the company is not affected by dividend payout. thereafter, scholars started to overthrew the previous assumptions and thrived to explain the reasons for dividend payout decisions under different circumstances.

Firm Size (Total Asset)

The size of the firm has been found closely related with its dividend payout decision. Denis and Osobov(2008) found that the majority of large and profitable firms pay the dividends in a wide range of countries though this effect is generated by the size or profitability of firms is not clear but it suggested that size does matter. This correlation has also been verified in the study of food sector companies done by Fama and French (2001) that larger size is correlated with a larger likelihood of dividend payout

Capital Expenditure

The literature including capital expenditure as one of the predictors or explanatory variables is quite limited and we include this variable to observe its effect on dividend payout decision.

Leverage

Another crucial factor of dividend payout is found to be the leverage ratio. The capital structure of the firm is closely connected with variance in dividends (Belo, Collin-Dufresne & Goldstein, 2015). The study done by Von Eije and Megginson in 2008 found the negative relationship between leverage ratio and likelihood of dividend payout decision. Similar conclusions can be found in other studies (Cooper & Lambertides, 2018). Besides long-term debt, a higher debt ratio is found to be correlated with the decision not to pay out dividends due to the need to keep cash in hand to meet creditors’ demands.

Company Maturity (Age)

The findings regarding the relationship between company’s ages are somewhat ambiguous and limited. Only a limited amount of research verify the positive relationship between the

maturity and its dividend payout decision (Al-Malkawi et al., 2010; Brawn and Ševic´, 2018). However, the study done by Agyei and Marfo-Yiadom in 2011 using the sample of firms in Ghana shows the contrary result. Therefore, we look forward to verifying the effect of the firm's age in the circumstance of Japanese corporations.

Sales and Earnings

As what we mentioned above, corporations believe that corporations choose to pay out dividends to increase their profitability and some of previous literature did find a significant correlation. It's been found that whether a firm chooses to pay out dividend is depending on earnings of the firm and the percentage they used from the retained earnings. (DeAngelo, DeAngelo, & Stulz, 2006; Fama and Babiak,1968;Benartzi, Michaely and Thaler, 1997). In addition, ROE and ROA are significant indicators of a firm's profitability and they are correlated positively with dividend payouts. However, contrary to the previous findings, while ROE is significantly and negatively related with the dividend payout ratio, ROA is positively related in the study done by Kazmierska-Jozwiak (2015) Generally speaking, the relationship between the two is expected to be positive.

3. Models and Data

3.1.1 Variables Measurement

In this paper, we have developed three models towards our two respondent variables —cash holding and dividend payout decision. Based on literature review, we found the majority of the main variables are significantly related though the effect of several variables shown to be ambiguous and limited, we still include them due to the limited number of variables in this circumstance.

Cash holding is measured by the ratio between cash and total assets : *casholding* = *cash/at*. Dividend payment decision is a binary indicator, and it's yes if total dividend is larger than zero and it's no if it's not : *divdecision* .

Firm size is measured using the variable total asset size, which takes the natural logarithm of total assets : *tlasset*. Capital expenditure is measured by the ratio between capital expenditure and total assets : *cerat*. Leverage is measured as the ratio between long-term debt and total assets : *lrat*. Company age is taken from the natural logarithm of company age : *comage* . Sales is taken from the natural logarithm of company sales : *SALES.* Earnings before interest, taxes, depreciation and amortization is measured as the ratio between it and total asset : EBITDA.

3.1.2 Descriptive Statistics of Variables

Through the processing of data, we got seven continuous variables except the binary indicator — dividend decision and the descriptive statistics are shown as the following:

TABLE2: Descriptive Statistics of All Continuous Variables

Statistics	N	Mean	Standard Deviation	Min	Percentage 25%	Percentage 75%	Max
casholding	24997	0.214	0.161	0.015	0.097	0.285	0.768
tlasset	24997	5.765	1.741	2.282	4.536	6.832	10.574
cerat	24997	0.032	0.034	0.000	0.008	0.044	0.177
lrat	24997	0.087	0.107	0.000	0.000	0.137	0.466
comage	24997	3.956	0.669	1.946	3.611	4.431	5.147
SALES	24997	5.754	1.757	1.853	4.525	6.892	10.331
EBITDA	24997	0.093	0.058	0.005	0.053	0.119	0.314

3.2 Models

As mentioned above we built up three models to establish our dependent variables with other crucial explanatory variables.

The linear regression model of cash holding with and without divdecision :

$$(1) : casholding_{it} = \beta_0 + \beta_1 tlasset_{it} + \beta_2 cerat_{it} + \beta_3 lrat_{it} + \beta_4 comage_{it} + \beta_5 SALES_{it} + \beta_6 EBITDA_{it} + \beta_7 divdecision_{it} + \varepsilon_{it}$$

$$(1) : casholding_{it} = \beta_0 + \beta_1 tlasset_{it} + \beta_2 cerat_{it} + \beta_3 lrat_{it} + \beta_4 comage_{it} + \beta_5 SALES_{it} + \beta_6 EBITDA_{it} + \varepsilon_{it}$$

Because the second variable dividend decision is a binary variable, we instead used two models , the first one is the logistic regression model and the second one is the LDA model :

$$(2) : divdecision_{it} = \beta_0 + \beta_1 tlasset_{it} + \beta_2 cerat_{it} + \beta_3 lrat_{it} + \beta_4 comage_{it} + \beta_5 SALES_{it} + \beta_6 EBITDA_{it} + \varepsilon_{it}$$

4. Linear Regression Models on Cash Holding Level

4.1 Linear regression models with main predictors

Prediction results and model fit

	Dependent variable:
	casholding
tlasset	0.0237***
	(0.0015)
cerat	-0.8219***
	(0.0246)
lrat	-0.2874***
	(0.0078)
comage	-0.0622***
	(0.0014)
SALES	-0.0505***
	(0.0014)
EBITDA	0.6263***
	(0.0147)
Constant	0.6072***
	(0.0056)
Observations	24,997
R^2	0.4260
Adjusted R^2	0.4258
Residual Std. Error	0.1221 (df = 24990)
F Statistic	3,090.7730*** (df = 6; 24990)
Note:	*p<0.1; **p<0.05; ***p<0.01

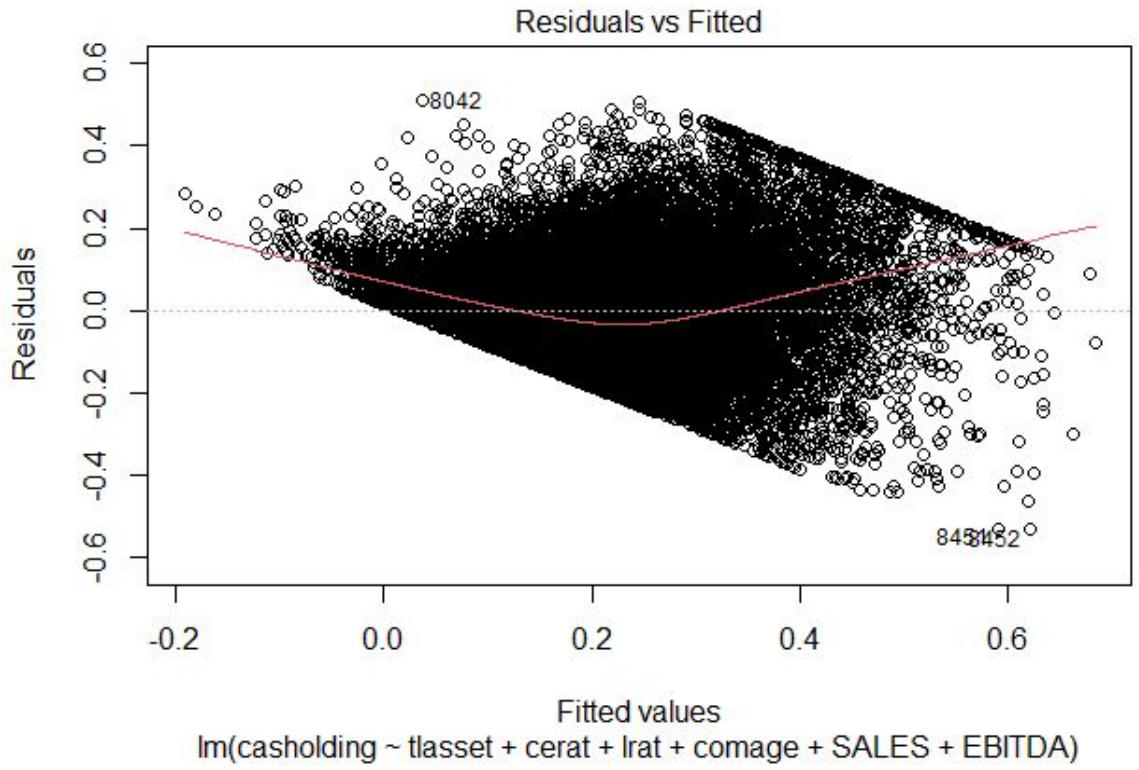
(table 4.1 summary results of linear regression)

From table 4.1, we can see empirical results of the linear regression, some observations we find are: first of all, all 6 of our predictors meet the standards of the 0.01 significance level, which implies that the results are statistically significant. The F-value also aligns with the results as well. The Adjusted R^2 is at 0.4258, which by rule of thumb suggested (Cohen.,1995), is considered high. With all these said, we can conclude that in terms of statistical significance, the empirical results are robust enough for us to discuss deeper on its economic significance. In this paper, we defined it as "whether a variable influence on the dependent variable is important enough to have any actual significance ".

From the results, we can see that the variables Capital expenditure ratio(*cerat*) , leverage ratio (*lrat*) and *EBITA* have the largest coefficients in terms of absolute value. The numbers of these three variables, symbolizing the percentage of growth of the dependent variable for each 1% increase of the variable, are -0.8219%, -0.2874% and 0.6263% respectively. This holds along the previous literature mentioned in the previous sections: capital spending and leverage ratios have a negative correlation with cash holdings, whereas *EBITA*, which can be viewed as profitability , has a positive correlation with holdings.

The remaining three variables, i.e. total assets, company age and sales, have relatively smaller impact on the dependent variable. This also to a sense confirms our findings in previous studies, of which rarely mentions total assets and sales as important determinants of cash holdings. Though the age of the company may be indirectly correlated with factors such as maturity of a company and growth, since mature and growing companies have different correlations with cash holding, without further segenmation, it is to no surprise that the total influence is relatively small.

The regression results suggest that all six predictors are statistically significant at the 0.05 significance level in predicting a company's level of cash holding. Total assets size is positively correlated with the level of corporate cash holding, whereas capital expenditure ratio, leverage ratio, company age, sales and the ratio of earnings/total assets are negatively correlated with the level of corporate cash holding.



(plot 4.1 Residuals vs Fitted)

plot 4.1 shows us that when we plot the residuals with fitted values, we can see an obvious linear decreasing trend, which means the relationship between cash holding levels and those predictors may not be linear. Therefore, we may need non-linear models to better predict the cash holding levels with those predictors. This also holds with the fact that the MSE is 0.014, which is relatively low.

4.2 Linear regression models with main predictors plus dividend payment decision

From the previous literature review, we found that the dividend payout decision (or its ratio) is closely correlated with cash holding. Therefore, we constructed another linear regression that combines the dividend payout decision as one of the predictors. The new regression results suggest that the dividend decision YES is a statistically significant predictor for dividend decision at the 0.05 significance level. And the coefficient for this

predictor is negative, meaning that paying a dividend is negatively correlated with the corporate cash holding levels. This result is in line with the findings of Kim et al. in 2011, but different from those of Drobetz and Grüninger in 2007, who found that dividend payments and operating cash flows are positively related to cash holding.

5. Logistic Regression and LDA Analysis on Dividend Decision

Results analysis for logistic regression

	Dependent variable:
	divdecision
tlasset	0.3523***
	(0.0364)
cerat	5.6053***
	(0.6282)
lrat	-3.6611***
	(0.1859)
comage	0.5904***
	(0.0338)
SALES	0.4698***
	(0.0337)
EBITDA	2.9233***
	(0.3371)
Constant	-4.7350***
	(0.1413)
Observations	24,997
Log Likelihood	-7,944.1330
Akaike Inf. Crit.	15,902.2700
Note:	*p<0.1; **p<0.05; ***p<0.01

(table 5.1 Logistic regression results)

According to table 5.1, all six predictors are statistically significant in predicting the dividend payout decision.

The variables Capital expenditure ratio(*cerat*) , leverage ratio (*lrat*) and *EBITA* have the largest coefficients in terms of absolute values. The coefficient of *cerat* is 5.6053, which means for every one unit increase in the capital expenditure ratio, the log odds of paying a

dividend increases by 5.6053. The coefficient of *lrat* is -3.6611, which means for a one unit increase in leverage ratio, the log odds of paying a dividend decreases by 3.6611. Other four coefficient values can be interpreted in the similar way.

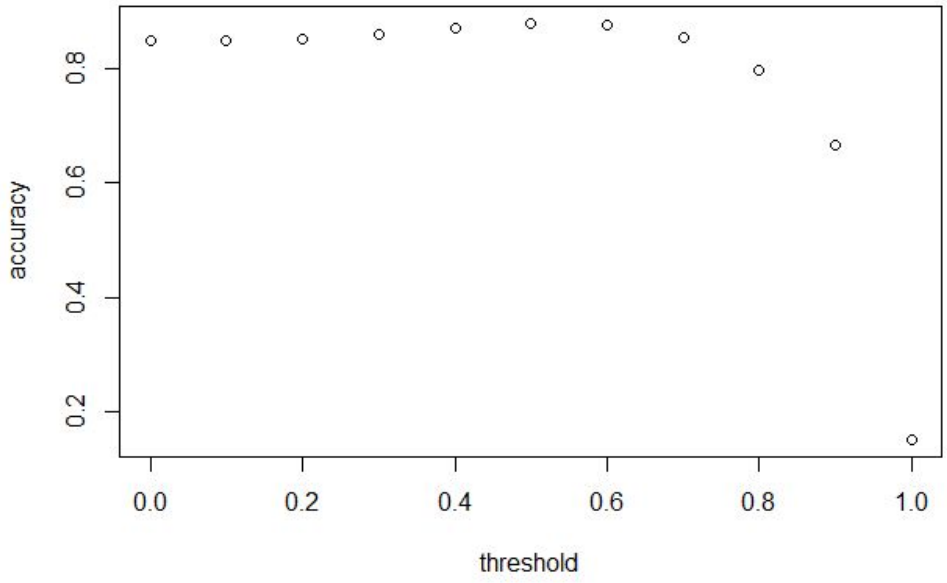
Output explanation of LDA model

The LDA output indicates that $\hat{\pi}1 = 0.1518$ and $\hat{\pi}2 = 0.8482$; in other words, 15.2 % of the training observations correspond to times when companies don't pay a dividend payment.

The group means are the average of each predictor within each class, and are used by LDA as estimates of μ_k. The group means results suggest that variable *tlasset* might have a greater influence on the decision of not paying a dividend (4.2272) than on the decision of paying a dividend (6.0399). This situation also happens with the variables *cerat*, *comage*, and *SALES*. On the contrary, for variable *lrat* and *EBITDA*, they may have a greater impact on the decision of paying a dividend than on the decision of not paying a dividend.

The coefficients of linear discriminants output provides the linear combination of all predictors that are used to form the LDA decision rule. If $0.060\times tlasset + 5.295\times cerat - 3.132\times lrat + 0.638\times comage + 0.417\times SALES + 1.272\times EBITDA$ is large, then the LDA classifier will predict the company to pay a dividend, and if it is small, then the LDA classifier will predict the company not to pay a dividend.

Prediction performances for the two models and how they vary with different thresholds



(plot 5.1 Prediction Accuracy of Logistics Regression vs Threshold levels)

Plot 5.1 illustrates how the prediction accuracy of the logistic regression model changes with different thresholds. We can see that the threshold of 0.5 generates the highest prediction accuracy among the 10 thresholds. Indeed, the R code provides the results that when the threshold is 0.5, the percentage of dividend decisions that are correctly predicted by the model is 87.86%. If we increase the threshold to 0.7 and 0.9, the percentages will decrease to 85.40% and 66.52%, respectively. On the contrary, if we decrease the threshold to 0.4 and 0.2, the percentages will decrease to 87.13% and 85.11%.

```
> table(lda_class,new_df$divdecision,dnn = c("Predicted", "Actual"))
         Actual
Predicted    No   Yes
      No   1255   519
      Yes  2539 20684
```

(table 5.1 Confusion Table for LDA model with the default threshold)

For the LDA model with the default threshold of 0.5, the percentage of dividend decisions that are correctly predicted by LDA is around (1255+20684)/24997=87.77%, that is to say,

the overall error rate is 12.23%. However, the percentage of those companies that do not make a dividend payment and are correctly predicted is only 33.08%: according to the confusion table, among 3794 companies which do not make a dividend payment, merely 1255 companies are correctly predicted by LDA. To improve LDA's prediction accuracy among these companies, we can lower the threshold.

```
> table(new.class.30, new_df$divdecision,dnn = c("Predicted", "Actual"))
         Actual
Predicted    No   Yes
      No   2079  1994
      Yes  1715 19209
```

(table 5.2 Confusion Table for LDA model with the threshold = 0.3)

Indeed, as shown in the table 5.2, if we lower the threshold to 30%, among the 3794 companies that do not make a dividend payment, now 2079 companies are correctly predicted. Though the overall error increases a bit to 14.84%, the error rate for the dividend decision prediction of those companies which do not make a dividend payment decreases from 66.92% to 54.80%, which is an obvious improvement.

6. Prediction Accuracy Evaluation of the Three Models

Splitting the data

Due to the fear of overfitting and accounting for randomness in using the entire data set, we now split the data set into 2 parts: training and testing data sets. Since this is a time series data, we view data before 2016(including) as training data and those after 2016 as testing data. The cleaned data set has 24,997 data points and the training and testing datasets have 16,044 and 8,953 data points respectively. This means the data is split in an approximately 65:35 ratio.

Prediction accuracy

On cash holding levels, we use linear regression and for Dividend payment decision, we use logistic regression and LDA. For the logistic regression, we chose the threshold of .5 according to the plot we constructed earlier. The results are as follows:

	model	dependent_variable	accuracy
1	linear regression	Cash holding	0.8688
2	logistic regression	Dividend payment descision	0.0496
3	LDA	Dividend payment descision	0.8864

(table6.1 prediction results summary of different models)

We can see that three logistic regression models performed extremely badly, with an accuracy of a mere 5 %. On the other hand both linear regression and LDA did moderately well, having 86.7% and 88.6% accuracy, respectively.

7. Validation and polynomials

Validation Methods

Lastly, to put in consideration of polynomials, we chose validation sets in two methods: validation sets and 10-fold-validation. Due to the size of our data, doing leave one out cross validation(LOOCV) proved to be computationally difficult, therefore we did not use the method. For the validation set, we chose to use the test data set as used above for the train-test models. The results are as shown below. Table 7.1 summaries all the errors of all the models using the two sorts of validation methods respectively. Plot 7.1~7.6 visualizes the change in error rates along the polynomials from power of 1 to power of 19.

Interpretation and results

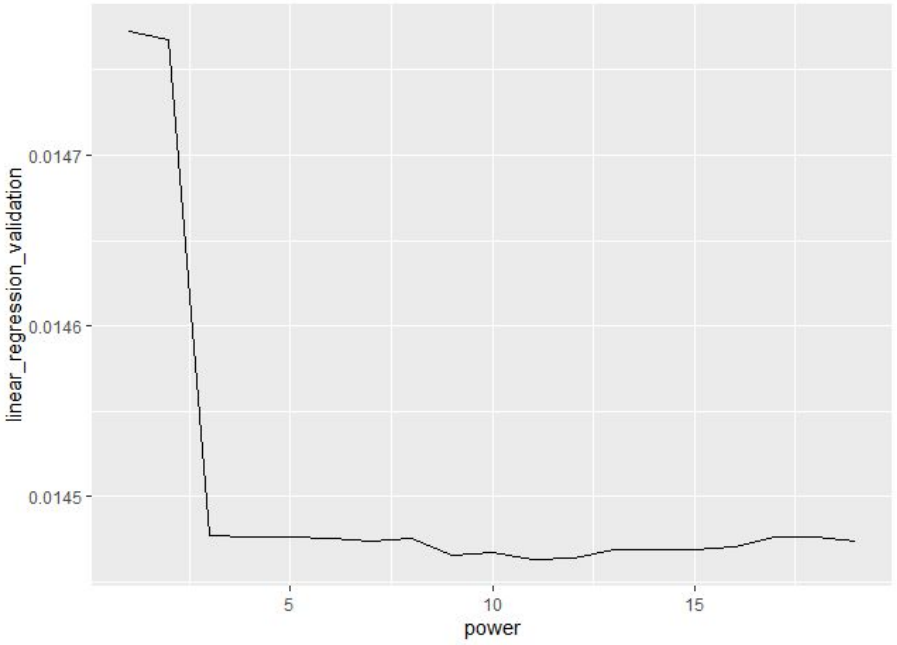
For linear regression we can see a clear drop both two kinds of validation at the 3 polynomial, the error rate falling by 3 %.

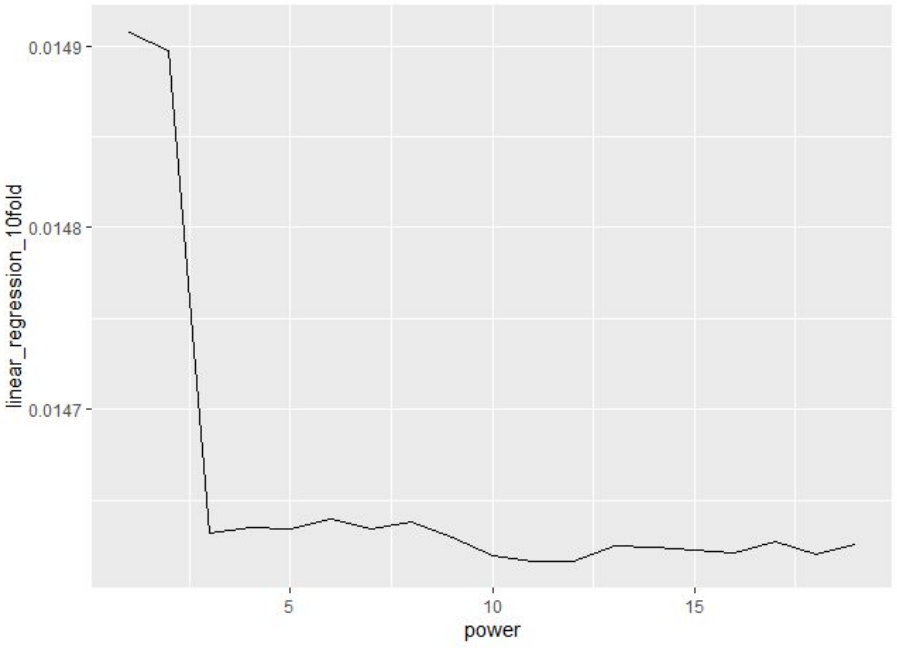
As for logistic regression, although as mentioned earlier, the overall error rate is still extremely high. However, we could also witness similar drop in both methods at around the power of 10. Finally, the LDA methods show different patterns by using different methods,

but overall we can still witness an upward trend with larger polynomials. Therefore we would chose to use the original variable, that is to use the power of 1.

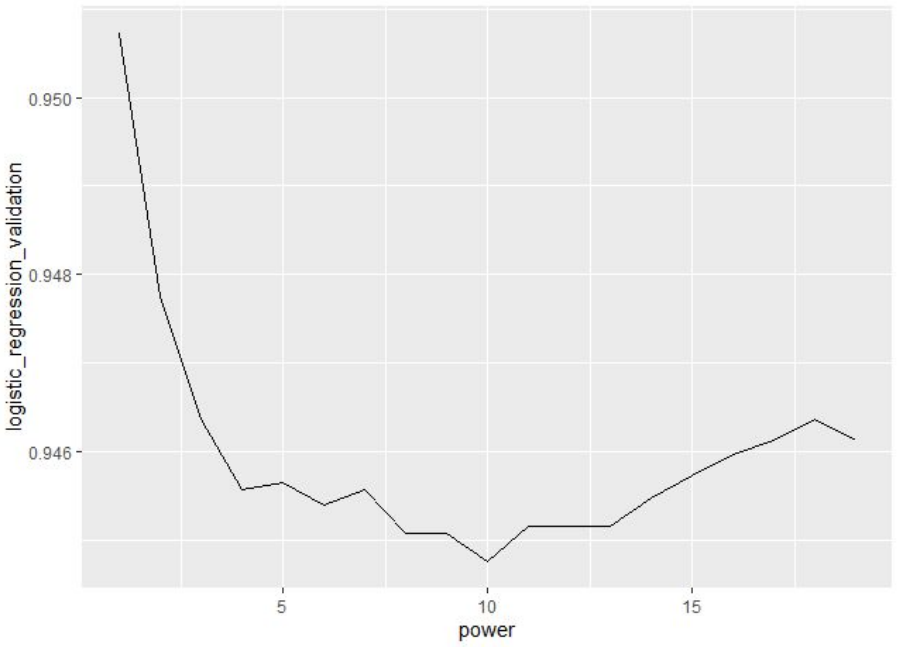
	power	linear_regression_validation	linear_regression_10fold	logistic_regression_validation	logistic_regression_10fold	LDA_validation	LDA_10fold
1	1	0.0148	0.0149	0.9507	0.0939	0.1246	0.1223
2	2	0.0148	0.0149	0.9477	0.0914	0.1272	0.1275
3	3	0.0145	0.0146	0.9464	0.0914	0.1245	0.1258
4	4	0.0145	0.0146	0.9456	0.0908	0.1242	0.1276
5	5	0.0145	0.0146	0.9456	0.0908	0.1256	0.1290
6	6	0.0145	0.0146	0.9454	0.0907	0.1254	0.1292
7	7	0.0145	0.0146	0.9456	0.0907	0.1263	0.1295
8	8	0.0145	0.0146	0.9451	0.0907	0.1255	0.1291
9	9	0.0145	0.0146	0.9451	0.0907	0.1265	0.1294
10	10	0.0145	0.0146	0.9448	0.0907	0.1258	0.1291
11	11	0.0145	0.0146	0.9452	0.0907	0.1261	0.1292
12	12	0.0145	0.0146	0.9452	0.0907	0.1252	0.1292
13	13	0.0145	0.0146	0.9452	0.0908	0.1252	0.1291
14	14	0.0145	0.0146	0.9455	0.0907	0.1254	0.1290
15	15	0.0145	0.0146	0.9457	0.0907	0.1255	0.1291
16	16	0.0145	0.0146	0.9460	0.0907	0.1259	0.1289
17	17	0.0145	0.0146	0.9461	0.0907	0.1253	0.1288
18	18	0.0145	0.0146	0.9464	0.0907	0.1259	0.1289
19	19	0.0145	0.0146	0.9461	0.0907	0.1257	0.1292

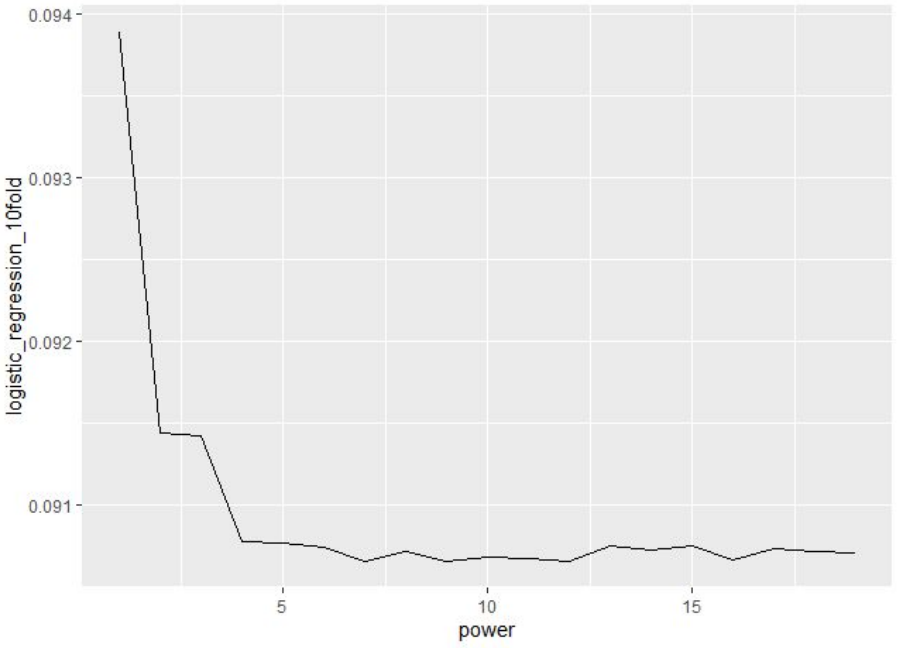
(Table 7.1 chart of polymnals error validation results)



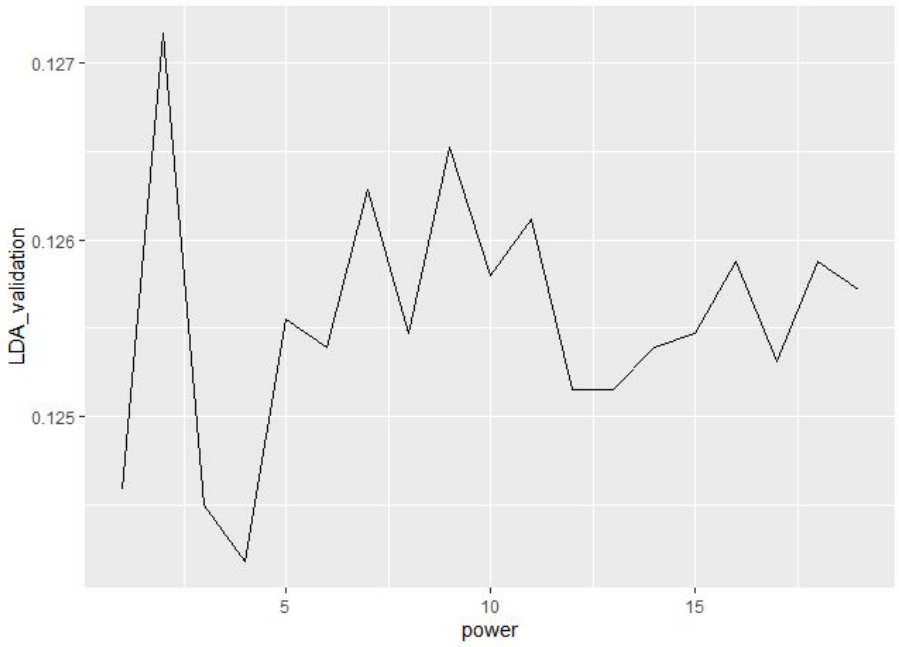


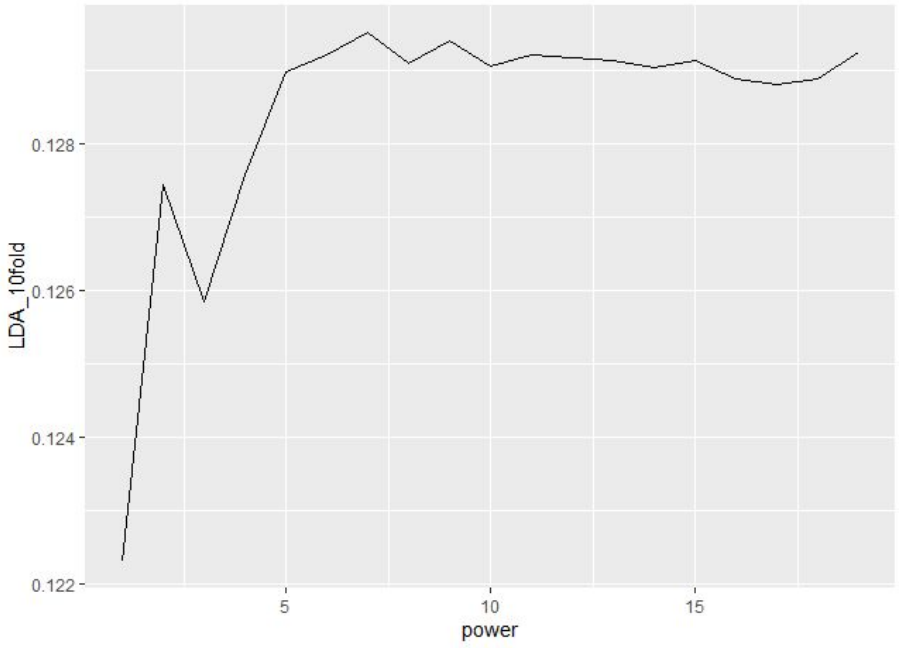
(Plot 7.1,7.2 plots of error rates of linear regression over polynomials using validation set and 10-fold)





(Plot 7.3,7.4 plots of error rates of logistic regression over polynomials using validation set and 10-fold)





(Plot 7.5,7.6 plots of error rates of LDA over polynomials using validation set and 10-fold)

Conclusion for empirical results

Mainly to answer the questions mentioned in part 3 and 4 we can conclude the following.

1. Linear Regression is a moderately suitable model for predicting cash holding levels at an accuracy of 86.7% on the test data

2. For determining Payout decision, LDA is more suitable for prediction providing an 88% accuracy on the test data. Logistic regression on the other hand is not suitable, at least for the given predictors and did poorly.
3. For polynomials on the age predictor, based on the validation results, we conclude that for linear regression the power of 3 is most appropriate. As for the LDA, despite the differences in pattern, we consider 10-fold to be more accurate and therefore concluded with retaining the power of 1.

Reference

Part 2:

A.P. Motives for Holding Cash Balances in a Firm: 5 Motives. (2016, December 12). Retrieved from https://www.accountingnotes.net/firm/motives-for-holding-cash-balances-in-a-firm-5-motives/11106

Hankins, K. W., & Petersen, M. (2020, January 17). Why Are Companies Sitting on So Much Cash? Retrieved from https://hbr.org/2020/01/why-are-companies-sitting-on-so-much-cash

Chen, J. (2020, November 19). Dividend Definition. Retrieved from https://www.investopedia.com/terms/d/dividend.asp

DeCecco, D. (n.d.). What is a Dividend Payment? Retrieved December 21, 2020, from https://www.wisegeek.com/what-is-a-dividend-payment.htm

Part 3 :

Afza, T., & Adnan, S.M. (2007). Determinants of corporate cash holdings: A case study of Pakistan. Proceedings of Singapore Economic Review Conference (SERC) 2007, August

01-04, Organized by Singapore Economics Review and The University of Manchester (Brooks World Poverty Institute), Singapore 164-165. [Online] Available: [https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=SERC2007&paper_id=166], http://ciit-isb.edu.pk/NEFiles/Images/Updated%20DrTalat%20Afza%20CV.pdf (Page 11) (August 3, 2011)

Dahya, J., & Travlos, N.G. (2000). Does the one man show pay? Theory and evidence on the dual CEO revisited. European Financial Management, 6(1), 461–483, http://dx.doi.org/10.1111/1468-036X.00113.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. Journal of Financial Economics, 40(2), 185–211, http://dx.doi.org/10.1016/0304-405X(95)00844-5.

Nadiri, M.I. (1969). The determinants of real cash balances in the U.S. total manufacturing sector. The Quarterly

Journal of Economics, 83(2), 173-196, http://dx.doi.org/10.2307/1883079.

Campbell, T., & Brendsel, L. (1977). The impact of compensating balance requirements on the cash balances of manufacturing corporations: An empirical study. The Journal of Finance, 32(1), 31-40, http://dx.doi.org/10.2307/2326900

Opler, T., Pinkowitz, L., Stulz, R., & Williamson, R. (1999). The determinants and implications of corporate cash holdings. Journal of financial economics, 52(1), 3-46, http://dx.doi.org/10.1016/S0304-405X(99)00003-3.

Ferreira, M.A., & Vilela, A.S. (2004). Why do firms hold cash? Evidence from EMU countries. European Financial Management, 10(2), 295-319, http://dx.doi.org/10.1111/j.1354-7798.2004.00251.x.

Nguyen, P. (2005). How sensitive are Japanese firms to earnings risk? Evidence from cash holdings. 1-41. [Online] Available:

http://www.musashi.jp/~togo/Seminar/Sensitivity_earnings_riskNguyen.pdf (August 3, 2011).

Saddour, K. (2006). The determinants and the value of cash holdings: Evidence from French firms. CEREG, 1-33. [Online] Available: http://www.dauphine.fr/cereg/cahiers_rech/cereg200606.pdf (August 3, 2011)

Drobetz, W., & Grüninger, M.C. (2007). Corporate cash holdings: Evidence from Switzerland. Financial markets Portfolio, 21, 293–324, http://dx.doi.org/10.1007/s11408-007-0052-8.

Hardin III, W.G., Highfield, M.J., Hill, M.D., & Kelly, G.W. (2009). The determinants of REIT cash holdings. Journal of Real Estate Finance and Economics, 39(1), 39-57, http://dx.doi.org/10.1007/s11146-007-9103-1.

Megginson, W.L., & Wei, Z. (2010). Determinants and value of cash holdings: Evidence from China's privatized firms. SSRN Working Paper Series, 1-37.

Kim, J., Kim, H., & Woods, D. (2011). Determinants of corporate cash-holding levels: An empirical examination of the restaurant industry. International Journal of Hospitality management, 30(3), 568-574, http://dx.doi.org/10.1016/j.ijhm.2010.10.004.

Rizwan, M.F., & Javed, T. (2011). Determinants of corporate cash holdings: Evidence from Pakistani public sector. Economics, Management and Financial Markets, 6(1), 344-358.

Bates, T., Kahle, K. and Stulz, R. (2009), „Why do U.S. firms hold so much more cash than they used to?", Journal of Finance, 64, (5).

Black, F. (1976) "The Dividend Puzzle" The Journal of Portfolio Management, Vol. 2: 5-8.

Amidu, M., & Abor, J. (2006). Determinants of dividend payout ratios in Ghana. The Journal of Risk Finance, 7(2), 136–145. doi:10.1108/15265940610648580

Allen, F. & Michaely, R. (2003) “Payout Policy, Handbook of the Economics of Finance, in: G.M. Constantinides & M. Harris & R. M. Stulz (ed.)” Handbook of the Economics of Finance, edition 1: chapter 7: 337-429.

Frankfurter, G. M., & Lane, W. R. (1992). The rationality of dividends. International Review of Financial Analysis, 1(2), 115–129. https://doi.org/10.1016/S1057-5219(15)30010-7 doi:10.1016/S1057-5219(15)30010-7

Denis, D. J., & Osobov, I. (2008). Why do firms pay dividends? International evidence on the determinants of dividend policy. Journal of Financial Economics, 89(1), 62–82. doi:10.1016/j. jfineco.2007.06.006

Belo, F., Collin-Dufresne, P., & Goldstein, R. S. (2015). Dividend dynamics and the term structure of dividend strips. Journal of Finance, 70(3), 1115–1160. doi:10.1111/jofi.12242

Cooper, I. A., & Lambertides, N. (2018). Large dividend increases and leverage. Journal of Corporate Finance, 48, 17–33. doi:10.1016/j.jcorpfin.2017.10.011

DeAngelo, H., DeAngelo, L., & Stulz, R. M. (2006). Dividend policy and the earned/contributed capital mix: A test of the life-cycle theory. Journal of Financial Economics,81(2),227–254.doi:10.1016/j.jfineco.2005.07.005

Kazmierska-Jozwiak, B. (2015). Determinants of dividend policy: Evidence from Polish Listed companies. Procedia Economics and Finance, 23, 473–477. https://doi.org/10.1016/S2212-5671(15)00490-6 doi:10.1016/S2212-5671(15)00490-6

Appendix : Group Members’ Contribution

CAI YIFAN :

- Data processing in part 1

- Provide some extra materials for Weijia in terms of theories of two dependent variables in part 2
- Found relevant literature of two dependent variables and wrote the whole literature review part of part 3
- Wrote R code of Part 3 and discussed with members

Weijia Wang:

- Found relevant literature and wrote most the part of Part 2 (combined with some theories found by Yifan Cai)
- Found and reviewed some literature on the determinants of corporate cash holdings
- Wrote R code for Part 3&4
- Explanation for part 3 b&c

Chen Yi Ju

- Created visualizations and tables for describing models and data
- Wrote and optimized R code for part 3&4
- Responsible for writing and interpreting the empirical results of the statistical methods and validation methods(part 3 a) and part 4)