VIET NAM NATIONAL UNIVERSITY HO CHI MINH CITY

**UNIVERSITY OF ECONOMICS AND LAW**

**FINAL-TERM REPORT**

**PROGRAM PACKAGE IN FINANCE 2**

**THE FACTORS THAT AFFECT THE DEBT MATURITY STRUCTURE OF DEVELOPMENT INVESTMENT CONSTRUCTION JOINT STOCK COMPANY (DIG)**

|  |  |
| --- | --- |
| Study program | : K20414C\_ Fintech |
| Course | : Program package in finance2 |
| Full Name | : Phạm Thị Kim Phượng |
| Class | : K20414C |
| ID student | : K204141927 |
| Email | : phuongptk20414c@st.uel.edu.vn |

Ho Chi Minh city, May 31, 2023

VIET NAM NATIONAL UNIVERSITY HO CHI MINH CITY

**UNIVERSITY OF ECONOMICS AND LAW**

**FINAL-TERM REPORT**

**PROGRAM PACKAGE IN FINANCE 2**

**TOPIC:**

**THE FACTORS THAT AFFECT THE DEBT MATURITY STRUCTURE OF DEVELOPMENT INVESTMENT CONSTRUCTION JOINT STOCK COMPANY (DIG)**

# DECLARATION

I hereby declare that the report "The factors that affect the debt maturity structure of Development Investment Construction Joint Stock Company (DIG)" is the result of my work under the guidance of Dr. Nguyen Thanh Liem within the framework of Program package in finance 2 module.

**TABLE OF CONTENTS**

[**DECLARATION 1**](#_Toc136784220)

[**LIST OF TABLES 3**](#_Toc136784221)

[**CONTENTS 4**](#_Toc136784222)

[**Abstract 4**](#_Toc136784223)

[**1. Introduction 4**](#_Toc136784224)

[**2. Literature review 4**](#_Toc136784225)

[**3. Create Dataset 7**](#_Toc136784226)

[**4. Descriptive statistics 9**](#_Toc136784227)

[**5. Data visualization 12**](#_Toc136784228)

[**6. Multiple regression 17**](#_Toc136784229)

[**7. ARIMA model 23**](#_Toc136784230)

[**8. How Decision Tree algorithm can be used to make prediction whether the firm will increase/decrease the debt maturity structure? 32**](#_Toc136784231)

[**REFERENCES 33**](#_Toc136784232)

[**APPENDIX 34**](#_Toc136784233)

# LIST OF TABLE

[Table 1: Variable, definition, measure, and sources of the variables 8](#_Toc136814610)

[Table 2: Head of data 9](#_Toc136814611)

[Table 3: Descriptive statistics of all the variables for the entire period 10](#_Toc136814612)

[Table 4: Descriptive statistics of all the variables from quarter 1 of 2010 to quarter 4 of 2019 12](#_Toc136814613)

[Table 5: Descriptive statistics of all the variables from quarter 1 of 2020 to quarter 4 of 2021 13](#_Toc136814614)

[Table 6: Result of predicting the response value for all the quarters of the sample 26](#_Toc136814615)

[Table 7: Predict the debt maturity structure for the four quarters in 2022 34](#_Toc136814616)

[Table 8: Comparison of actual DMS and forecasted DMS 35](#_Toc136814617)

# LIST OF FIGURE

[Figure 1: Box plot of DIG’s debt maturity structure for the entire period 13](#_Toc136814511)

[Figure 2: Box plot of DIG’s debt maturity structure before COVID-19 14](#_Toc136814512)

[Figure 3: Box plot of DIG’s debt maturity structure during COVID-19 14](#_Toc136814513)

[Figure 4: Histogram of DIG's debt maturity structure for the entire period 15](#_Toc136814514)

[Figure 5: Histogram of DIG's debt maturity structure before COVID-19 16](#_Toc136814515)

[Figure 6: Histogram of DIG's debt maturity structure during COVID-19 17](#_Toc136814516)

[Figure 7: The relationship between the independent and dependent variables 19](#_Toc136814517)

[Figure 8: Result of regression model 1 19](#_Toc136814518)

[Figure 9: Result of checking important assumptions for linear regression 20](#_Toc136814519)

[Figure 10: Result of regression model 2 22](#_Toc136814520)

[Figure 11: Result of checking important assumptions for linear regression 24](#_Toc136814521)

[Figure 12: The debt maturity structure of DIG from 2010 to 2021 26](#_Toc136814522)

[Figure 13: The debt maturity structure series at the first hierarchical error 27](#_Toc136814523)

[Figure 14: The debt maturity structure series at the second hierarchical error 28](#_Toc136814524)

[Figure 15: ACF and PACF schemes of the debt maturity structure series at the second hierarchical error 29](#_Toc136814525)

[Figure 16: ARIMA(3,2,2) model Residuals 31](#_Toc136814526)

[Figure 17: Residuals from ARIMA(3,2,2) 32](#_Toc136814527)

# CONTENTS

## Abstract

The maturity structure of corporate debt is one of the significant financing choices that a firm must make simultaneously while deciding how to finance its operational and investment decisions. The choice of a suitable debt maturity structure is exceptionally relevant for firms because it can enable them to avoid mismatch by aligning assets in line with liabilities, addressing agency-related problems, sidestepping the ill effects of the cost of capital, and signaling the firms’ earning quality and value. The study investigates the firm-specific determinants and COVID-19 determinants significant for the debt maturity structure of Development Investment Construction Joint Stock Company with data from the company's financial statements collected on finance.vietstock.com from the first quarter of 2010 to the fourth quarter of 2021. The results of the regression model indicate that liquidity, leverage, asset intensity, and the dividend policy significantly determine the debt maturity structure. This study also explores the impact of COVID-19 on firm-specific determinants of debt maturity structure.

## Introduction

Debt is an important aspect of finance and a tool to control disadvantages and increase advantages when choosing between short-term and long-term debt. Debt maturity is the maturity of short-term and long-term debt. Deciding to choose an optimal debt term will help businesses have many opportunities, reduce risks from funding sources, increase transparency, and exploit opportunities from tax deductions due to debt. Short-term debt plays an important role in reducing the costs of enterprises. However, borrowing in the short term causes enterprises to face many risks in terms of liquidity, refinancing, and reinvesting. This leads businesses, especially construction enterprises such as Development Investment Construction Joint Stock Corporation (DIG Corp.), to face a dilemma in choosing between short-term debt and long-term debt to bring the greatest benefit.

This study focuses on factors affecting the debt maturity structure of Development Investment Construction Joint Stock Corporation, analyzing the impact of internal factors of the company and the impact of COVID-19 on firm-specific determinants of debt maturity structure.

## Literature review

There are research papers on debt maturity structure in both Vietnam and abroad, such as:

Mohammed (2020) researched the factors affecting debt maturity structure among the listed non-financial firms in Nigeria. The results indicate that the non-debt tax-shield, liquidity, asset intensity, diversification, growth opportunity, firm size, and dividend policy significantly determine the debt maturity structure among the listed non-financial firms in Nigeria. However, the evidence is not enough to conclude that profitability and investors’ confidence determine the debt maturity structure among the non-financial firms in Nigeria. Firm diversification and liquidity appeared to have the most profound negative effect on the debt maturity structure, in line with predictions of the special use of debt hypothesis and the pecking order theory. Overall, it is concluded that firm-specific factors determine the choice of debt maturity structure among Nigerian listed non-financial firms.

Lemma and Negash (2012) show that internal factors impact a firm’s debt maturity structure. Their findings confirm that asset maturity, income volatility, and debt ratio positively affect the debt maturity structure. Besides examining company-specific characteristics, they explore the effect of the industry sector and characteristics of the economy on the decisions about the debt maturity structure of firms in African countries. Their studies have shown that the size of the economy has a positive effect on debt maturity structure; thus, firms in low-income countries tend to use less long-term debt, whereas taxes and economic growth rate (GDP) have the opposite effect.

Acording Méndez (2013), the empirical determinants of a firm's debt maturity structure are examined for a sample of 38,993 non-financial Spanish firms over the period 1995-2006. The results show the relevance of growth opportunities, size, asymmetric information, and asset maturity in explaining debt maturity.

Correia et al. (2014) explore the factors affecting the debt maturity structure in European countries. Research results show that internal factors, including firm size, asset maturity, and leverage ratio, positively correlate with long-term debt, whereas profitability is negatively correlated with long-term debt. Their findings are in line with Lemma and Negash (2012); specifically, the larger the size of the banking system, the more these firms use short-term debt.

In Vietnam, Nguyen (2018) shows that the debt maturity structure of companies in Vietnam is dynamic. The author examines the internal and external factors that influence the debt maturity structure of Vietnamese enterprises. Internal characteristics, such as earnings volatility, liquidity, tangible assets, and firm size, positively affect debt maturity structure. In Vietnam, physical assets are the most important intrinsic factor affecting long-term debt. External circumstances impact a firm’s debt term structure. In contrast, institutional quality and economic growth had no effect, and interest rate term structure, inflation, and the level of financial development, which included the intermediary financial system and financial markets, were all positively connected.

Pham (2018) studies the debt maturity structure of real estate companies listed on the Vietnamese stock market. The research results show that financial leverage, company size, asset structure, solvency, profit fluctuations are the factors affecting the debt maturity structure of enterprises, other factors such as growth opportunities, corporate income tax is not statistically significant.

Pham (2020) conducts another study regarding the capital structure and debt maturity structures of Vietnamese real estate investment and business firms. The findings of the study suggest that institutions hurt debt term structure decisions. Liquidity, business risk, firm size, financial development, and inflation affect debt maturity structure choice.

The Ngo and Le (2021) study investigates the firm-specific and macroeconomic determinants significant for the debt maturity structure of Vietnamese corporate firms. A sample of 722 non-financial firms listed on the Ho Chi Minh and Hanoi Stock Exchange in Vietnam from 2007 to 2018 was taken to test the hypothesis. The study’s methods of fixed effects panel data analysis provide empirical evidence that firm size, firms’ quality, liquidity, leverage, asset maturity, tax impact, and macro variables are significantly related to the debt maturity structure. Their study measures debt maturity based on the balance sheet approach and define debt maturity as the proportion of long-term debt to total debt.

Nguyen (2022) used the GMM method to find the factors that are significant predictors of debt maturity structure in Vietnamese listed firms from 2010 to 2019, namely the lagged debt maturity structure, leverage ratio, profitability, firm size, growth opportunities, GDP, and inflation. Research shows that debt maturity structure is positively correlated with lagged debt maturity structure, firm size, growth opportunities, and GDP under agency cost theory, while signaling theory creates the ground for the negative effects of profitability and leverage on debt maturity structure.

Based on the literature review, the way to measure the variables and the expectation of the sign of the estimator coefficients βi in the model are presented and explained in Table 1.

Table 1: Variable, definition, measure, and sources of the variables

|  |  |  |  |
| --- | --- | --- | --- |
|  | Variable | Measure | Sources |
| Dependent variable | Debt maturity structure | Long term debt/ Total debt | Ngo and Le (2021), Nguyen (2022) |
| Independent  variable | Liquidity | Current assets/current liabilities | Kalsie and Nagpal (2018), Hussain et al. (2018), Mohammed (2020), Pham (2018), Pham (2020) |
|  | Leverage | Total debt/ Total assets | Pham (2020), Nguyen (2022) |
|  | Profitability | Earnings before interest and tax/ Total assets | Correia et al. (2014), Pham (2018), Nguyen (2022) |
|  | Asset intensity | Fixed asset/Total assets | Lemma and Negash (2012), Mohammed (2020) |
|  | Covid | 0: Before Covid-19 period,  1: Covid-19 period |  |

## Create Dataset

First, read the data, check for missing values, and handle it to make sure the dataset is clear. Then calculate the necessary variables, including debt maturity structure, liquidity, leverage, profitability, and asset intensity. In addition, the data included a dummy variable, COVID-19.

**CODE:**

|  |
| --- |
| #required packages  library(tidyverse)  library(readxl)  library(ggplot2)  library(lmtest)  library(forecast)  library(xts)  library(tseries)  #load data  data <- read\_excel("D://học tập//NĂM 3//hk 6//gói ứng dụng trong tài chính//cuối kỳ//K204141927.xlsx", sheet = 1)  #all variable names  colnames(data) |

**RESULT:**

|  |
| --- |
| [1] "Time" "long\_term\_debt" "short\_term\_debt"  [4] "total\_current\_assets" "fixed\_assets" "total\_assets"  [7] "interest\_expense" "EBT" "covid" |

**CODE:**

|  |
| --- |
| #checking Na values  sum(is.na(data)) |

**RESULT:**

|  |
| --- |
| [1] 0 |

The data has no missing values. So we will proceed to calculate the variables.

**CODE:**

|  |
| --- |
| #create variables  data <- data %>%  mutate(DMS= round((long\_term\_debt/(long\_term\_debt + short\_term\_debt)),4),  liquidity = round((total\_current\_assets/short\_term\_debt),4),  profitability = round((EBT + interest\_expense)/total\_assets,4), # + interest\_expense  leverage = round(((long\_term\_debt+short\_term\_debt)/total\_assets),4),  asset\_intensity = round((fixed\_assets/total\_assets),4)  )  df <- data %>%  select(Time, DMS, liquidity, profitability, leverage, asset\_intensity)  head(df)  #checking Na values  sum(is.na(data)) |

**RESULT:**

Table 2: Head of data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Time** | **DMS** | **liquidity** | **profitability** | **leverage** | **asset\_intensity** |
| 31/03/2010 | 0.335 | 2.54 | 0.0202 | 0.376 | 0.0767 |
| 30/06/2010 | 0.346 | 2.76 | 0.0265 | 0.358 | 0.0748 |
| 30/09/2010 | 0.343 | 2.75 | 0.0824 | 0.364 | 0.0751 |
| 31/12/2010 | 0.257 | 2.38 | 0.0481 | 0.369 | 0.058 |
| 31/03/2011 | 0.368 | 2.43 | 0.0076 | 0.412 | 0.0742 |
| 30/06/2011 | 0.466 | 2.82 | 0.0069 | 0.424 | 0.0726 |

**CODE:**

|  |
| --- |
| #checking Na values  sum(is.na(df)) |

**RESULT:**

|  |
| --- |
| [1] 0 |

## Descriptive statistics

**CODE:**

|  |
| --- |
| #entire period  entire\_period <- df %>%  summarise(variables = c('DMS','liquidity','profitability','leverage','asset\_intensity'),  obs = nrow(df),  min = c(min(DMS), min(liquidity), min(profitability), min(leverage), min(asset\_intensity)),  mean = c(mean(DMS), mean(liquidity), mean(profitability), mean(leverage), mean(asset\_intensity)),  median = c(median(DMS), median(liquidity), median(profitability), median(leverage), median(asset\_intensity)),  std = c(sd(DMS), sd(liquidity), sd(profitability), sd(leverage), sd(asset\_intensity)),  max = c(max(DMS), max(liquidity), max(profitability), max(leverage), max(asset\_intensity))  )  entire\_period = data.frame(entire\_period)  entire\_period |

**RESULT:**

Table 3: Descriptive statistics of all the variables for the entire period

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variables** | **Obs** | **Min** | **Mean** | **Median** | **Std** | **Max** |
| DMS | 48 | 0.1467 | 0.43899375 | 0.47920 | 0.16144022 | 0.6798 |
| liquidity | 48 | 1.2172 | 2.70494792 | 2.68900 | 0.68744383 | 4.0971 |
| profitability | 48 | -0.0044 | 0.01278125 | 0.00660 | 0.01846407 | 0.0824 |
| leverage | 48 | 0.3575 | 0.50190000 | 0.50715 | 0.05709864 | 0.6340 |
| asset\_intensity | 48 | 0.0361 | 0.06423542 | 0.05555 | 0.02677784 | 0.1182 |

From Table 3, it can be seen that there were 48 observations for the entire period from quarter 1 of 2010 to quarter 4 of 2021. The mean value of the debt maturity structure is 0.439, which means that DIC Joint Stock Company used average long-term debts of 0.439 during the period 2010–2021. Compared to the average long-term debt ratio of enterprises in the real estate industry of 0.308 (Pham, 2020), DIC Joint Stock Company uses more long-term debt. However, the company still uses short-term debt more than long-term debt in its debt structure. The financial market in Vietnam is still underdeveloped, the funding sources of companies are limited, most of these sources depend mainly on bank loans. Although banks have many types of loans from short term to long term. However, short-term loans are always preferred, and the conditions for making loans are also lighter than proving that the company has enough requirements for medium and long-term loan procedures. The minimum and maximum values of the debt maturity structure are 0.1467 and 0.6798, respectively. The minimum values of the liquidity and profitability variables are 1.2172 and -0.0044, respectively. However, their maximum values are 4.0971 and 0.0824. Leverage ratio has a mean value of 0.5019, and its minimum and maximum values are 0.3575 and 0.6340, respectively. The mean value of asset intensity is 0.0642. Simultaneously, the minimum and maximum values of asset intensity are 0.0361 and 0.1182, respectively.

**CODE:**

|  |
| --- |
| #before Covid-19 pandemic  before\_covid <- df[1:40, ]  before\_covid\_period <- before\_covid %>%  summarise(variables = c('DMS','liquidity', 'profitability','leverage', 'asset\_intensity'),  obs = nrow(before\_covid),  min = c(min(DMS), min(liquidity), min(profitability), min(leverage), min(asset\_intensity)),  mean = c(mean(DMS), mean(liquidity), mean(profitability), mean(leverage), mean(asset\_intensity)),  median = c(median(DMS), median(liquidity), median(profitability), median(leverage), median(asset\_intensity)),  std = c(sd(DMS), sd(liquidity), sd(profitability), sd(leverage), sd(asset\_intensity)),  max = c(max(DMS), max(liquidity), max(profitability), max(leverage), max(asset\_intensity))  )  before\_covid\_period = data.frame(before\_covid\_period)  before\_covid\_period |

**RESULT:**

Table 4: Descriptive statistics of all the variables from quarter 1 of 2010 to quarter 4 of 2019

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variables** | **Obs** | **Min** | **Mean** | **Median** | **Std** | **Max** |
| DMS | 40 | 0.1653 | 0.4757550 | 0.50500 | 0.14375338 | 0.6798 |
| liquidity | 40 | 1.8681 | 2.8874075 | 2.75480 | 0.56334075 | 4.0971 |
| profitability | 40 | -0.0044 | 0.0112800 | 0.00615 | 0.01687874 | 0.0824 |
| leverage | 40 | 0.3575 | 0.4903425 | 0.50170 | 0.05070914 | 0.5455 |
| asset\_intensity | 40 | 0.0361 | 0.0642650 | 0.05395 | 0.02794183 | 0.1182 |

Table 4 shows that there are 40 observations for the period from quarter 1 of 2010 to quarter 4 of 2019. The mean value of the debt maturity structure is 0.4758, which means that DIC Joint Stock Company used average long-term debts of 0.4758 during the period 2010–2019. The minimum and maximum values of the debt maturity structure are 0.1653 and 0.6798, respectively. The minimum values of the liquidity and profitability variables are 1.8681 and -0.0044, respectively. However, their maximum values are 4.0971 and 0.0824. Leverage ratio has a mean value of 0.5019, and its minimum and maximum values are 0.3575 and 0.5455, respectively. The mean value of asset intensity is 0.0643. Simultaneously, the minimum and maximum values of asset intensity are 0.0361 and 0.1182, respectively.

**CODE:**

|  |
| --- |
| #during Covid-19 pandemic  on\_covid <- df[41:48, ]  on\_covid\_period <- on\_covid %>%  summarise(variables = c('DMS','liquidity', 'profitability','leverage', 'asset\_intensity'),  obs = nrow(on\_covid),  min = c(min(DMS), min(liquidity), min(profitability), min(leverage), min(asset\_intensity)),  mean = c(mean(DMS), mean(liquidity), mean(profitability), mean(leverage), mean(asset\_intensity)),  median = c(median(DMS), median(liquidity), median(profitability), median(leverage), median(asset\_intensity)),  std = c(sd(DMS), sd(liquidity), sd(profitability), sd(leverage), sd(asses\_intensity)),  max = c(max(DMS), max(liquidity), max(profitability), max(leverage), max(asset\_intensity))  )  on\_covid\_period = data.frame(on\_covid\_period)  on\_covid\_period |

**RESULT:**

Table 5: Descriptive statistics of all the variables from quarter 1 of 2020 to quarter 4 of 2021

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **variables** | **obs** | **min** | **mean** | **median** | **std** | **max** |
| DMS | 8 | 0.1467 | 0.2551875 | 0.23705 | 0.11630539 | 0.4913 |
| liquidity | 8 | 1.2172 | 1.7926500 | 1.84140 | 0.51332173 | 2.4033 |
| profitability | 8 | 0.0023 | 0.0202875 | 0.00835 | 0.02499042 | 0.0648 |
| leverage | 8 | 0.4869 | 0.5596875 | 0.56920 | 0.05462667 | 0.6340 |
| asset\_intensity | 8 | 0.0407 | 0.0640875 | 0.06025 | 0.02155448 | 0.0963 |

Table 5 shows that there are 8 observations for the period from quarter 1 of 2020 to quarter 4 of 2021. The mean value of the debt maturity structure is 0.2552, which means that DIC Joint Stock Company used average long-term debts of 0.2552 during the period 2020–2021. That is down more than 22% the debt maturity structure during the period 2010–2019. The minimum and maximum values of the debt maturity structure are 0.1467 and 0.4913, respectively. The minimum values of the liquidity and profitability variables are 1.2172 and 0.0023, respectively. However, their maximum values are 2.4033 and 0.0648. The mean value of liquidity decreased from 2.8874 to 1.7927. However, the mean value of profitability increased from 0.0113 to 0.0203. Leverage ratio has a mean value increased from 0.5019 to 0.560, and its minimum and maximum values are 0.4869 and 0.6340, respectively. The mean value of asset intensity is 0.0641, unchanged so much from the previous period. Simultaneously, the minimum and maximum values of asset intensity are 0.0407 and 0.0963, respectively. Thus, we can conclude that the COVID-19 pandemic has had an effect on the debt maturity structure of the company. The company's debt maturity structure has declined since the COVID-19 pandemic emerged. This shows that the majority of enterprises' loans are short-term loans and are mainly from financial sources such as bank financing.

## Data visualization

**CODE:**

|  |
| --- |
| #box plot  boxplot(df$DMS,  main = "Box plot of DIG's debt maturity structure",  col = "lightblue",  xlab= "Debt maturity structure",  horizontal = TRUE) |

**RESULT:**

Figure 1: Box plot of DIG’s debt maturity structure for the entire period

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| #before covid  boxplot(before\_covid$DMS,  main = "Box plot of DIG's debt maturity structure before Covid19",  col = "lightblue",  xlab= "Debt maturity structure",  horizontal = TRUE) |

**RESULT:**

Figure 2: Box plot of DIG’s debt maturity structure before COVID-19

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| #during covid  boxplot(on\_covid$DMS,  main = "Box plot of DIG's debt maturity structure during Covid19",  col = "lightblue",  xlab= "Debt maturity structure",  horizontal = TRUE) |

**RESULT:**

Figure 3: Box plot of DIG’s debt maturity structure during COVID-19

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| # histogram  ggplot(df, aes(x = DMS)) +  geom\_histogram(binwidth = 0.08, fill = "lightblue", color = "black") +  labs(title = "Histogram of DIG's debt maturity structure for the entire period",  x = "Debt maturity structure",  y = "Frequency") |

**RESULT:**

Figure 4: Histogram of DIG's debt maturity structure for the entire period

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| # before covid  ggplot(before\_covid, aes(x = DMS)) +  geom\_histogram(binwidth = 0.07, fill = "lightblue", color = "black") +  labs(title = "Histogram of DIG's debt maturity structure before Covid19",  x = "Debt maturity structure",  y = "Frequency") |

**RESULT:**

Figure 5: Histogram of DIG's debt maturity structure before COVID-19

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| #during covid  ggplot(on\_covid, aes(x = DMS)) +  geom\_histogram(binwidth = 0.12, fill = "lightblue", color = "black") +  labs(title = "Histogram of DIG's debt maturity structure during Covid19",  x = "Debt maturity structure",  y = "Frequency") |

**RESULT:**

Figure 6: Histogram of DIG's debt maturity structure during COVID-19

|  |
| --- |
|  |

Both the histogram plot and the box plot of the debt maturity structure of the DIC company for the entire period show that the debt maturity structure distribution of DIC is right skewed and mostly ranges from the lowest of 0.15 to the highest of around 0.68. However, the DIC company regularly maintains a debt maturity structure ranging from 0.3 to 0.58, of which the most frequent rate is 0.47.

During the period before Q1 2020, the debt term structure of the DIC company is right skewed and mostly ranges from the lowest of 0.15 to the highest of around 0.68. However, the DIC company regularly maintains a debt maturity structure ranging from 0.36 to 0.59, of which the most frequent rate is 0.51.

During the period from Q1 2020 to Q4 2021, the debt term structure of the DIC company is left skewed and mostly ranges from the lowest of 0.15 to the highest of around 0.49. However, the DIC company regularly maintains a debt maturity structure ranging from 0.17 to 0.31, of which the most frequent rate is 0.23.

## Multiple regression

* 1. **With all the individual variables (model 1)**

The relationship between the independent and dependent variables should be linear. We can check this visually through scatterplot.

**CODE:**

|  |
| --- |
| #Check linearity & multicollinearity  par(mfrow=c(1,4))  plot(liquidity ~ DMS, data = df)  plot(profitability ~ DMS, data = df)  plot(leverage ~ DMS, data = df)  plot(asset\_intensity ~ DMS, data = df) |

**RESULT:**

Figure 7: The relationship between the independent and dependent variables

|  |
| --- |
|  |

The relationship between DMS and liquidity seems clear. But the relationships between DMS and profitability, DMS and leverage, DMS and asset intensity are a bit less clear.

**CODE:**

|  |
| --- |
| #multiple regression  dms.lm<-lm(DMS ~ liquidity + profitability + leverage + asset\_intensity, data = df)  summary(dms.lm) |

**RESULT:**

Figure 8: Result of regression model 1

|  |
| --- |
|  |

The research results show that except for profitability, there is no impact on the debt maturity of the companies; the remaining variables all have a statistically significant impact on debt maturity. Specifically, as follows: liquidity, combined leverage, and asset intensity are positively correlated with debt maturity. All other things being unchanged, a company's liquidity increases by 1%, the DIC's debt maturity structure increases by 0,237%, the DIC's leverage increases by 1%, the DIC's debt maturity structure increases by 0,722%, the DIC's asset intensity increases by 1%, and the debt maturity structure increases by 1.279%. In this model, independent variables have explanatory significance at the significance level of 10%, and these independent variables explained 89.47% of the variation of DIC's debt maturity structure through the adjusted R-squared index.

**CODE:**

|  |
| --- |
| #Check multicollinearity  car::vif(dms.lm)  #check important assumptions for linear regression  par(mfrow=c(2,2))  plot(dms.lm) |

**RESULT:**

Figure 9: Result of checking important assumptions for linear regression

|  |
| --- |
| > car::vif(dms.lm)  liquidity profitability leverage asset\_intensity  1.435991 1.355786 1.464662 1.278439 |

Test results for multicollinearity (VIF) among independent variables in the model show that all values are less than 10, so there is no multicollinearity between the variables.

**CODE:**

|  |
| --- |
| #check normality  shapiro.test(resid(dms.lm)) # Null hypothesis is normality  # Check homoscedasticity  bptest(dms.lm) #Null hypothesis is homoscedasticity |

**RESULT:**

|  |
| --- |
| Shapiro-Wilk normality test  data: resid(dms.lm)  W = 0.93117, p-value = 0.007535  studentized Breusch-Pagan test  data: dms.lm  BP = 4.1371, df = 4, p-value = 0.3878 |

A Shapiro-Wilk test of the normality of the model residual indicated that the residuals are normal with p-value = 0.6909 > α = 0.1 (accepting H0). At the same time, this data set does not have homoscedasticity with p-value = 0.0005093 < α = 0.1 (reject H0) in the Breusch-Pagan test. That indicates that the OLS is a good estimator of the sampled data.

* 1. **With the usual individual variables and the interaction between Covid-19 dummy variable and the independent variables (model 2)**

Since the profitability variable is not statistically significant, we will leave it out of the model.

**CODE:**

|  |
| --- |
| #convert data  data <- data %>%  mutate(covid = ifelse(covid == 'covid\_period',1,0)  )  data$covid = factor(data$covid)  df2 <- data %>%  select(Time, DMS, liquidity, profitability, leverage, asset\_intensity, covid)  #multiple regression  dms2.lm<-lm(DMS ~ liquidity + leverage + asset\_intensity + liquidity\*covid + leverage\*covid + asset\_intensity\*covid, data = df2) #liquidity\*covid + + leverage\*covid  summary(dms2.lm) |

**RESULT:**

Figure 10: Result of regression model 2

|  |
| --- |
|  |

The research results show that except for asset\_intensity\*covid, there is no impact on the debt maturity of the companies; the remaining variables all have a statistically significant impact on debt maturity. Specifically, as follows: liquidity, leverage, asset intensity, liquidity\*covid1, and leverage\*covid1 are positively correlated with debt maturity while covid1 is negatively correlated with debt maturity. All other things being unchanged, a company's liquidity increases by 1%, the DIC's debt maturity structure increases by 0,24202%, the DIC's leverage increases by 1%, the DIC's debt maturity structure increases by 0,63225%, the DIC's asset intensity increases by 1%, and the debt maturity structure increases by 1.55624%, the DIC's liquidity\*covid1 increases by 1%, and the debt maturity structure increases by 0.13015%, the DIC's leverage\*covid1 increases by 1%, and the debt maturity structure increases by 1.91954%. The impact of COVID-19 causes the debt maturity structure to decrease by 1.24016%. In this model, independent variables have explanatory significance at the significance level of 10%, and these independent variables explained 92.64% of the variation of DIC's debt maturity structure through the adjusted R-squared index.

**CODE:**

|  |
| --- |
| #Check multicollinearity  car::vif(dms2.lm)  #check important assumptions for linear regression  par(mfrow=c(2,2))  plot(dms2.lm) |

**RESULT:**

Figure 11: Result of checking important assumptions for linear regression

|  |
| --- |
| liquidity leverage asset\_intensity  1.991930 1.685618 1.287007  covid liquidity:covid leverage:covid  1023.068306 54.197181 537.598126  asset\_intensity:covid  18.613400 |

Test results for multicollinearity (VIF) among independent variables in the model show that covid, liquidity\*covid, leverage\*covid, and asset\_intensity\*covid are higher than 10, so there is multicollinearity between the variables.

**CODE:**

|  |
| --- |
| # Check normality  shapiro.test(resid(dms2.lm)) # Null hypothesis is normality  # Check homoscedasticity  bptest(dms2.lm) #Null hypothesis is homoscedasticity |

**RESULT:**

|  |
| --- |
| Shapiro-Wilk normality test  data: resid(dms2.lm)  W = 0.95576, p-value = 0.06805  studentized Breusch-Pagan test  data: dms2.lm  BP = 8.0146, df = 7, p-value = 0.3313 |

A Shapiro-Wilk test of the normality of the model residual indicated that the residuals are not normal with p-value = 0.06805 < α = 0.1 (reject H0). At the same time, this data set has homoscedasticity with p-value = 0.3313 > α = 0.1 (accepting H0) in the Breusch-Pagan test. This make the estimated regression coefficients ineffective.

To obtain reliable estimates in the presence of heteroscedasticity, there are several corrective measures that can be taken, including:

- Transform the variables in the model to reduce the impact of heteroscedasticity. For example, taking the logarithm or square root of a variable may help to reduce heteroscedasticity.

- Use a different modeling technique that can handle heteroscedasticity, such as weighted least squares. This type of regression assigns a weight to each data point based on the variance of its fitted value. Essentially, this gives small weights to data points that have higher variances, which shrinks their squared residuals. When the proper weights are used, this can eliminate the problem of heteroscedasticity.

* 1. **Predict the value of the variable of assigned topic for all the quarters of the sample using Model 1**

**CODE:**

|  |
| --- |
| #use model 1 to predict the response value for all the quarters of the sample  pred\_model = predict(dms.lm, data = df[,3:6])  pred\_table =data.frame(Time = df$Time,  Actual\_DMS = df$DMS,  Predicted\_DMS = round(pred\_model,4))  pred\_table$Difference = abs(pred\_table$Actual\_DMS - pred\_table$Predicted\_DMS)  pred\_table  summary(pred\_table$Difference) |

**RESULT:**

Table 6: Result of predicting the response value for all the quarters of the sample

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Actual\_DMS** | **Dredicted\_DMS** | **Difference** |
| 31/03/2010 | 0.3348 | 0.3233 | 0.0115 |
| 30/06/2010 | 0.3464 | 0.3586 | 0.0122 |
| 30/09/2010 | 0.3432 | 0.3505 | 0.0073 |
| 31/12/2010 | 0.2568 | 0.2518 | 0.0050 |
| 31/03/2011 | 0.3678 | 0.3228 | 0.0450 |
| 30/06/2011 | 0.4657 | 0.4217 | 0.0440 |
| ... | … | … | … |
| 30/09/2020 | 0.2522 | 0.3549 | 0.1027 |
| 31/12/2020 | 0.1475 | 0.1462 | 0.0013 |
| 31/03/2021 | 0.1765 | 0.1652 | 0.0113 |
| 30/06/2021 | 0.1467 | 0.1556 | 0.0089 |
| 30/09/2021 | 0.3460 | 0.2678 | 0.0782 |
| 31/12/2021 | 0.4913 | 0.3620 | 0.1293 |

|  |
| --- |
| > summary(pred\_table$Difference)  Min. 1st Qu. Median Mean 3rd Qu. Max.  0.00040 0.01123 0.02630 0.03750 0.05903 0.12930 |

The regression model predicts quite well the debt maturity structure for all the quarters of the sample. The difference between the actual debt maturity structure and the predicted debt maturity structure is not large, in the range of 0.00040 to 0.12930.

## ARIMA model

**CODE:**

|  |
| --- |
| class(df$Time)  date.time = seq(as.Date("2010/03/31"), by = "quarter", length.out = nrow(df))  date.time  # Create a xts Dataframe  dms = xts(df[,2],date.time)  class(dms)  #plot the data - look at the debt maturity structure of DIG  par(mfrow=c(1,1))  plot(dms) |

**RESULT:**

Figure 12: The debt maturity structure of DIG from 2010 to 2021

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| #stationary check  par(mfrow=c(1,2))  print(adf.test(dms)) |

**RESULT:**

|  |
| --- |
| Augmented Dickey-Fuller Test  data: dms  Dickey-Fuller = -2.2378, Lag order = 3, p-value = 0.4792  alternative hypothesis: stationary |

P value is 0.4792 > 0.05. So, the debt maturity structure series is not stationary. Therefore, we take the first hierarchical error in the series and check for stationarity.

**CODE:**

|  |
| --- |
| dmsdiff1 <- diff(dms, differences = 1)  par(mfrow=c(1,1))  plot(dmsdiff1)  dmsdiff1 = na.omit(dmsdiff1)  print(adf.test(dmsdiff1)) |

**RESULT:**

Figure 13: The debt maturity structure series at the first hierarchical error

|  |
| --- |
|  |

|  |
| --- |
| Augmented Dickey-Fuller Test  data: dmsdiff1  Dickey-Fuller = -2.5208, Lag order = 3, p-value = 0.3665  alternative hypothesis: stationary |

P value is 0.3665 > 0.05. So, the 1st hierarchy of the debt maturity structure series is not stationary. Therefore, we take the second hierarchical error in the series and check for stationarity.

**CODE:**

|  |
| --- |
| dmsdiff2 <- diff(dms, differences = 2)  par(mfrow=c(1,1))  plot(dmsdiff2)  dmsdiff2 = na.omit(dmsdiff2)  print(adf.test(dmsdiff2)) |

**RESULT:**

Figure 14: The debt maturity structure series at the second hierarchical error

|  |
| --- |
|  |

|  |
| --- |
| Augmented Dickey-Fuller Test  data: dmsdiff2  Dickey-Fuller = -5.6808, Lag order = 3, p-value = 0.01  alternative hypothesis: stationar |

P value is 0.01 < 0.05. The ADF test results also show that the 2nd hierarchy of the debt maturity structure series is stationary, so d = 2.

**CODE:**

|  |
| --- |
| #graph the ACF and PACF looking for identifiable lags PACF -> p ACF -> q  par(mfrow=c(1,2))  acf(dmsdiff2, main = 'ACF for differenced Series', lag.max = 40)  pacf(dmsdiff2, main = 'PACF for differenced Series', lag.max = 40)  acf(dmsdiff2, lag.max = 40, plot = FALSE)  pacf(dmsdiff2, lag.max = 40, plot=FALSE) |

**RESULT:**

Figure 15: ACF and PACF schemes of the debt maturity structure series at the second hierarchical error

|  |
| --- |
|  |

Based on the PACF and ACF schemes in Figure 15, we have:

PACF: The order of AR(p) can be selected as: 1, 2, 3

ACF: The order of MA(q) can be selected as: 1, 2, 18

Therefore, we consider the following 9 ARIMA models: ARIMA(1,2,1), ARIMA(1,2,2), ARIMA(1,1,18), ARIMA(2,2,1), ARIMA(2,2,2), ARIMA(2,2,18), ARIMA(3,2,1), ARIMA(3,2,2), ARIMA(3,2,18). If the model has the smallest AIC, SigmaSQ, and maximum likelihood, it will be the chosen model.

**CODE:**

|  |
| --- |
| #building ARIMA model  a = auto.arima(dms,seasonal=F) #ARIMA(0,1,0)  a  b = arima(dms, order = c(1,2,1)) #sigma^2 estimated as 0.007535: log likelihood = 45.2, aic = -84.4  b  arima(dms, order = c(1,2,2)) #sigma^2 estimated as 0.007544: log likelihood = 45.2, aic = -82.4  arima(dms, order = c(1,2,18)) #sigma^2 estimated as 0.003476: log likelihood = 54.93, aic = -69.85  c = arima(dms, order = c(2,2,1)) #sigma^2 estimated as 0.007131: log likelihood = 46.2, aic = -84.4  c  arima(dms, order = c(2,2,2)) #sigma^2 estimated as 0.007129: log likelihood = 46.21, aic = -82.41  arima(dms, order = c(2,2,18)) #sigma^2 estimated as 0.002872: log likelihood = 56.69, aic = -71.38  arima(dms, order = c(3,2,1)) #sigma^2 estimated as 0.007434: log likelihood = 46.43, aic = -82.86  d = arima(dms, order = c(3,2,2)) #sigma^2 estimated as 0.006222: log likelihood = 48.29, aic = -84.59  d  tsdisplay(residuals(d), lag.max = 40 ,main = '(3,2,2) Model Residuals')  arima(dms, order = c(3,2,18)) #sigma^2 estimated as 0.00287: log likelihood = 56.71, aic = -69.41 |

**RESULT:**

Figure 16: ARIMA(3,2,2) model Residuals

|  |
| --- |
|  |

Choose model ARIMA(3,2,2) with sigma^2 estimated as 0.006222, log likelihood = 48.29, aic = -84.59. To see if the model is effective in predicting, we need to test whether the residual series et is white noise or not. To test the hypothesis that et is a white noise, we need to check the stationarity of the residual series et and check whether et follows the normal distribution.

**CODE:**

|  |
| --- |
| #perform the portmanteau test  checkresiduals(d, lag = 40)  summary(d$residuals) |

**RESULT:**

Figure 17: Residuals from ARIMA(3,2,2)

|  |
| --- |
|  |

|  |
| --- |
| Ljung-Box test  data: Residuals from ARIMA(3,2,2)  Q\* = 27.302, df = 35, p-value = 0.8203  Model df: 5. Total lags used: 40  > summary(d$residuals)  Min. 1st Qu. Median Mean 3rd Qu. Max.  -0.2082700 -0.0339960 -0.0003034 0.0051902 0.0449494 0.2797717 |

Since the ACF histogram shows that no autocorrelation in the first 16 lags exceeds the limit of statistical significance and the p-value of the Ljung-Box test is 0.8203. Therefore, we can conclude that there is little evidence for non-zero autocorrelation in the forecast errors at the first 16 lags. The forecast errors are normally distributed with zero mean and constant variance.

**CODE:**

|  |
| --- |
| #ARMA roots table  plot(d) |

**RESULT:**

|  |
| --- |
|  |

Both AR and MA roots are inside the circle. The estimated model is covariance stationary and the estimated process is invertible.

**CODE:**

|  |
| --- |
| #predict the DMS for the 4 quarters in 2022  dmsforecast = forecast(b,h = 4, level = c(95))  dmsforecast  par(mfrow=c(1,1))  plot(dmsforecast) |

**RESULT:**

Table 7: Predict the debt maturity structure for the four quarters in 2022

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Point Forecast** | **Lo 95** | **Hi 95** |
| 49 | 0.4879297 | 0.3300959 | 0.6457636 |
| 50 | 0.4908107 | 0.2819825 | 0.6996390 |
| 51 | 0.5390058 | 0.3042845 | 0.7737270 |
| 52 | 0.6164303 | 0.3529888 | 0.8798718 |

|  |
| --- |
|  |

**CODE:**

|  |
| --- |
| #comparison of actual DMS and forecasted DMS  actual\_dms2022 <- read\_excel("D://học tập//NĂM 3//hk 6//gói ứng dụng trong tài chính//cuối kỳ//K204141927.xlsx",  sheet = 2)  actual\_dms <- actual\_dms2022 %>% select(Time, DMS)  dmsforecast\_ = data.frame(forecast(b,h = 4, level = c(95)))  forecast\_table = data.frame(Time = actual\_dms$Time,  Actual\_DMS = round((actual\_dms$DMS),4),  Predicted\_DMS = dmsforecast\_$Point.Forecast)  forecast\_table$Difference = abs(forecast\_table$Actual\_DMS - forecast\_table$Predicted\_DMS)  forecast\_table |

**RESULT:**

Table 8: Comparison of actual DMS and forecasted DMS

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Actual\_DMS** | **Predicted\_DMS** | **Difference** |
| 31/03/2022 | 0.5304 | 0.4879297 | 0.04247027 |
| 30/06/2022 | 0.5567 | 0.4908107 | 0.06588926 |
| 30/09/2022 | 0.5673 | 0.5390058 | 0.02829424 |
| 31/12/2022 | 0.4366 | 0.6164303 | 0.17983030 |

The forecasted debt maturity structure is not much different from reality, about 0.03% to 0.18%.

## How Decision Tree algorithm can be used to make prediction whether the firm will increase/decrease the debt maturity structure?

The Decision Tree algorithm can be used to make predictions about whether a firm will increase or decrease its debt maturity structure. The Decision Tree model can be used as follows:

First, prepare the data: obtain the data from the company's financial statements, clean the data, handle missing values, and encode categorical variables if necessary. Divide the dataset into two parts: a training set and a test set.

Second, identify the features from the dataset that are most likely to influence the change in debt maturity structure.

Label encoding: Assign labels to the target variable that represents the change in debt maturity structure. For example, you can encode "increase" as 1 and "decrease" as 0.

Decision Tree training: Use the training set to train a Decision Tree classifier. The decision tree algorithm will use historical data to create a tree-like model that captures the relationships between the input features and the target.

Model evaluation: Using the test set, evaluate the trained decision tree model. Use appropriate evaluation metrics such as accuracy, precision, recall, or F1 score to assess its performance.

Prediction: Once the model has been trained and evaluated, it can be used to make predictions on new, previously unseen data. Provide the trained decision tree with the relevant features of a firm, and it will predict whether the firm's debt maturity structure is likely to increase or decrease.

# REFERENCES

Correia, S., Brito, P., & Brandão, E. (2014). *Corporate Debt Maturity-An international comparison of firm debt maturity choices* (No. 544). Universidade do Porto, Faculdade de Economia do Porto.

Hussain, H. I., Shamsudin, M. F., Salehuddin, S., & Jabarullah, N. H. (2018). Debt maturity and Shari’ah compliance: Evidence from Malaysian panel data.

Kalsie, A. & Nagpal, A. (2018). The determinants of corporate debt maturity for nse-listed corporates. *FIIB Business Review*, 7(1), 43-56. DOI: 10.1177/2319714518766117

Lemma, T. T., & Negash, M. (2012). Capital and debt maturity structures of a firm: Evidence from selected African countries. *Journal of Business and Policy Research*, 7(2), 60–92.

Méndez, V. M. G. (2013). Determinants of debt maturity structure across firm size*. Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad*, 42(158), 187-209.

Mohammed, L. (2020). Firm-specific determinants of debt maturity structure of listed non-financial firms in Nigeria. *Malaysian management journal*.

Nguyen, T. N. (2018). Factors affecting debt maturity structure of the companies in Vietnam. *Ph.D thesis*. University of Economics Ho Chi Minh City.

NGO, V. T., & LE, T. L. (2021). Factors influencing corporate debt maturity: An empirical study of listed companies in Vietnam. *The Journal of Asian Finance, Economics and Business*, 8(5), 551-559.

Nguyen, K. Q. T. (2022). Determinants of debt maturity structure: Evidence in Vietnam. *Cogent Business & Management*, 9(1), 2094588.

Pham, T. V. (2020). Capital Structure and Debt Term Structure of Real Estate Investment and Construction Enterprises in Vietnam. *Ph.D Thesis*. Banking University of Ho Chi Minh City.

Van Trinh, PT (2018). Debt term structure of real estate companies listed on Vietnam stock market. *HO CHI MINH CITY OPEN UNIVERSITY SCIENCE-ECONOMIC AND BUSINESS ADMINISTRATION*, *13* (1), 38-50.

# APPENDIX

#1/ required packages

library(tidyverse)

library(readxl)

library(ggplot2)

library(lmtest)

library(forecast)

library(xts)

library(tseries)

#2/ create dataset

#load data

data <- read\_excel("D://học tập//NĂM 3//hk 6//gói ứng dụng trong tài chính//cuối kỳ//K204141927.xlsx", sheet = 1)

#all variable names

colnames(data)

#checking Na values

sum(is.na(data))

#create variables

data <- data %>%

mutate(DMS= round((long\_term\_debt/(long\_term\_debt + short\_term\_debt)),4),

liquidity = round((total\_current\_assets/short\_term\_debt),4),

profitability = round((EBT + interest\_expense)/total\_assets,4), # + interest\_expense

leverage = round(((long\_term\_debt+short\_term\_debt)/total\_assets),4),

asset\_intensity = round((fixed\_assets/total\_assets),4)

)

df <- data %>%

select(Time, DMS, liquidity, profitability, leverage, asset\_intensity)

head(df)

#checking Na values

sum(is.na(df))

#3/descriptive statistics

#entire period

entire\_period <- df %>%

summarise(variables = c('DMS','liquidity', 'profitability','leverage', 'asset\_intensity'),

obs = nrow(df),

min = c(min(DMS), min(liquidity), min(profitability), min(leverage), min(asset\_intensity)),

mean = c(mean(DMS), mean(liquidity), mean(profitability), mean(leverage), mean(asset\_intensity)),

median = c(median(DMS), median(liquidity), median(profitability), median(leverage), median(asset\_intensity)),

std = c(sd(DMS), sd(liquidity), sd(profitability), sd(leverage), sd(asset\_intensity)),

max = c(max(DMS), max(liquidity), max(profitability), max(leverage), max(asset\_intensity))

)

entire\_period = data.frame(entire\_period)

entire\_period

#before Covid-19 pandemic

before\_covid <- df[1:40, ]

before\_covid\_period <- before\_covid %>%

summarise(variables = c('DMS','liquidity', 'profitability','leverage', 'asset\_intensity'),

obs = nrow(before\_covid),

min = c(min(DMS), min(liquidity), min(profitability), min(leverage), min(asset\_intensity)),

mean = c(mean(DMS), mean(liquidity), mean(profitability), mean(leverage), mean(asset\_intensity)),

median = c(median(DMS), median(liquidity), median(profitability), median(leverage), median(asset\_intensity)),

std = c(sd(DMS), sd(liquidity), sd(profitability), sd(leverage), sd(asset\_intensity)),

max = c(max(DMS), max(liquidity), max(profitability), max(leverage), max(asset\_intensity))

)

before\_covid\_period = data.frame(before\_covid\_period)

before\_covid\_period

#during Covid-19 pandemic

during\_covid <- df[41:48, ]

during\_covid\_period <- during\_covid %>%

summarise(variables = c('DMS','liquidity', 'profitability','leverage', 'asset\_intensity'),

obs = nrow(during\_covid),

min = c(min(DMS), min(liquidity), min(profitability), min(leverage), min(asset\_intensity)),

mean = c(mean(DMS), mean(liquidity), mean(profitability), mean(leverage), mean(asset\_intensity)),

median = c(median(DMS), median(liquidity), median(profitability), median(leverage), median(asset\_intensity)),

std = c(sd(DMS), sd(liquidity), sd(profitability), sd(leverage), sd(asset\_intensity)),

max = c(max(DMS), max(liquidity), max(profitability), max(leverage), max(asset\_intensity))

)

during\_covid\_period = data.frame(during\_covid\_period)

during\_covid\_period

#4/box & whisker plot and histogram

#box plot

boxplot(df$DMS,

main = "Box plot of DIG's debt maturity structure for the entire period",

col = "lightblue",

xlab= "Debt maturity structure",

horizontal = TRUE)

#before covid

boxplot(before\_covid$DMS,

main = "Box plot of DIG's debt maturity structure before Covid19",

col = "lightblue",

xlab= "Debt maturity structure",

horizontal = TRUE)

#during covid

boxplot(during\_covid$DMS,

main = "Box plot of DIG's debt maturity structure during Covid19",

col = "lightblue",

xlab= "Debt maturity structure",

horizontal = TRUE)

#histogram

ggplot(df, aes(x = DMS)) +

geom\_histogram(binwidth = 0.08, fill = "lightblue", color = "black") +

labs(title = "Histogram of DIG's debt maturity structure for the entire period",

x = "Debt maturity structure",

y = "Frequency")

#before covid

ggplot(before\_covid, aes(x = DMS)) +

geom\_histogram(binwidth = 0.07, fill = "lightblue", color = "black") +

labs(title = "Histogram of DIG's debt maturity structure before Covid19",

x = "Debt maturity structure",

y = "Frequency")

#during covid

ggplot(during\_covid, aes(x = DMS)) +

geom\_histogram(binwidth = 0.12, fill = "lightblue", color = "black") +

labs(title = "Histogram of DIG's debt maturity structure during Covid19",

x = "Debt maturity structure",

y = "Frequency")

#5/ multiple regression

#5.1 regression with all the individual variables

#Check linearity & multicollinearity

par(mfrow=c(1,4))

plot(liquidity ~ DMS, data = df)

plot(profitability ~ DMS, data = df)

plot(leverage ~ DMS, data = df)

plot(asset\_intensity ~ DMS, data = df)

#Multiple regression

dms.lm<-lm(DMS ~ liquidity + profitability + leverage + asset\_intensity, data = df)

summary(dms.lm)

#Check multicollinearity

car::vif(dms.lm) #VIF >10 thì đa cộng tuyến

#check important assumptions for linear regression

par(mfrow=c(2,2))

plot(dms.lm)

#check normality

shapiro.test(resid(dms.lm))# Null hypothesis is normality

# Check homoscedasticity

bptest(dms.lm) #Null hypothesis is homoscedasticity #p.value >5%, lúc đó phương sai đồng nhất, phương sai không đổi

#5.2 regression with the usual individual variables

#and the interaction between Covid-19 dummy variable and the independent variables

#convert data

data <- data %>%

mutate(covid = ifelse(covid == 'covid\_period',1,0)

)

data$covid = factor(data$covid)

df2 <- data %>%

select(Time, DMS, liquidity, profitability, leverage, asset\_intensity, covid)

dms2.lm<-lm(DMS ~ liquidity + leverage + asset\_intensity + liquidity\*covid + leverage\*covid + asset\_intensity\*covid, data = df2) #liquidity\*covid + + leverage\*covid

summary(dms2.lm)

#Check multicollinearity

car::vif(dms2.lm)

#check important assumptions for linear regression

par(mfrow=c(2,2))

plot(dms2.lm)

# Check normality

shapiro.test(resid(dms2.lm))# Null hypothesis is normality

# Check homoscedasticity

bptest(dms2.lm) #Null hypothesis is homoscedasticity

#5.3 Prediction

#use model 1 to predict the response value for all the quarters of the sample

pred\_model = predict(dms.lm, data = df[,3:6])

pred\_table =data.frame(Time = df$Time,

Actual\_DMS = df$DMS,

Predicted\_DMS = round(pred\_model,4))

pred\_table$Difference = abs(pred\_table$Actual\_DMS - pred\_table$Predicted\_DMS)

pred\_table

summary(pred\_table$Difference)

#6/ ARIMA model

class(df$Time)

date.time = seq(as.Date("2010/03/31"), by = "quarter", length.out = nrow(df))

date.time

# Create a xts Dataframe

dms = xts(df[,2],date.time)

class(dms)

#plot the data - look at the debt maturity structure of DIG

par(mfrow=c(1,1))

plot(dms)

#stationary check

par(mfrow=c(1,2))

print(adf.test(dms))

dmsdiff1 <- diff(dms, differences = 1)

par(mfrow=c(1,1))

plot(dmsdiff1)

dmsdiff1 = na.omit(dmsdiff1)

print(adf.test(dmsdiff1))

dmsdiff2 <- diff(dms, differences = 2)

par(mfrow=c(1,1))

plot(dmsdiff2)

dmsdiff2 = na.omit(dmsdiff2)

print(adf.test(dmsdiff2))

#graph the ACF and PACF looking for identifiable lags PACF -> p ACF -> q

par(mfrow=c(1,2))

acf(dmsdiff2, main = 'ACF for differenced Series', lag.max = 40)

pacf(dmsdiff2, main = 'PACF for differenced Series', lag.max = 40)

acf(dmsdiff2, lag.max = 40, plot = FALSE)

pacf(dmsdiff2, lag.max = 40, plot=FALSE)

# p = 1,2,3 ; q = 1,2,18

#building ARIMA model

a = auto.arima(dms,seasonal=F) #ARIMA(0,1,0)

a

b = arima(dms, order = c(1,2,1)) #sigma^2 estimated as 0.007535: log likelihood = 45.2, aic = -84.4

b

arima(dms, order = c(1,2,2))#sigma^2 estimated as 0.007544: log likelihood = 45.2, aic = -82.4

arima(dms, order = c(1,2,18)) #sigma^2 estimated as 0.003476: log likelihood = 54.93, aic = -69.85

c = arima(dms, order = c(2,2,1)) #sigma^2 estimated as 0.007131: log likelihood = 46.2, aic = -84.4

c

arima(dms, order = c(2,2,2)) #sigma^2 estimated as 0.007129: log likelihood = 46.21, aic = -82.41

arima(dms, order = c(2,2,18)) #sigma^2 estimated as 0.002872: log likelihood = 56.69, aic = -71.38

arima(dms, order = c(3,2,1)) #sigma^2 estimated as 0.007434: log likelihood = 46.43, aic = -82.86

d = arima(dms, order = c(3,2,2)) #sigma^2 estimated as 0.006222: log likelihood = 48.29, aic = -84.59

d

tsdisplay(residuals(d), lag.max = 40 ,main = '(3,2,2) Model Residuals')

arima(dms, order = c(3,2,18)) #sigma^2 estimated as 0.00287: log likelihood = 56.71, aic = -69.41

#perform the portmanteau test

checkresiduals(d, lag = 40)

summary(d$residuals)

#ARMA roots table

plot(d)

#predict the DMS for the 4 quarters in 2022

dmsforecast = forecast(d,h = 4, level = c(95))

dmsforecast

par(mfrow=c(1,1))

plot(dmsforecast)

#comparison of actual DMS and forecasted DMS

actual\_dms2022 <- read\_excel("D://học tập//NĂM 3//hk 6//gói ứng dụng trong tài chính//cuối kỳ//K204141927.xlsx",

sheet = 2)

actual\_dms <- actual\_dms2022 %>% select(Time, DMS)

dmsforecast\_ = data.frame(forecast(d,h = 4, level = c(95)))

forecast\_table = data.frame(Time = actual\_dms$Time,

Actual\_DMS = round((actual\_dms$DMS),4),

Predicted\_DMS = dmsforecast\_$Point.Forecast)

forecast\_table$Difference = abs(forecast\_table$Actual\_DMS - forecast\_table$Predicted\_DMS)

forecast\_table