*Rhulane Mkhawane*

*MKHRHU002 | Word count: 1839*

*Time Series Assignment*

ECO3021S: Quantitative Methods in Economics

*University of Cape Town*

*Table of Contents*

Introduction 3

Literature Review 4

Data and Summary Statistics 5

**Model and Results 9**

**Conclusion 13**

**References 14**

**Appendix: R Code for Assignment 15**

Introduction

This article aims to produce and evaluate a forecast of gross domestic product (GDP) by making use of time series regression techniques. A study of the literature on the performance of yield curve’s ability to predict economic activity is conducted. The time series data provided (containing bond yield rates and GDP results) is then evaluated and used to form the regression models. This article later proves that the autoregressive model with lag 1 (AR (1)) is the best model for forecasting gross domestic product.

Literature Review

The yield curve is one of the most important financial tools in macroeconomics. It represents the relationship between the yield rates of a specific type of asset, including government bonds, and the different maturity terms (Kumar, Stauvermann and Vu, 2021). It is one of “the most closely watched financial indicator” (Haubrich and Dombrosky, 1996) because of its many advantages.

It is commonly used as an indicator of economic activity. There exists a direct relationship between the yield curve and economic activity: once the yield curve decreases, this is normally followed by a decrease in gross domestic product (Haubrich and Dombrosky, 1996). The yield curve can also be used to predict the real growth rate of gross domestic product: a direct relationship between the yield spread from the curve and the growth rate exists, with a decrease in the yield spread preceding a decrease in the real growth rate of the gross domestic product (Haubrich and Dombrosky, 1996).

A negative yield spread also normally indicates an incoming recession (Haubrich and Dombrosky, 1996). This is a result of the short term interest rate of government bonds being greater than their long term interest rates, indicating a tightening monetary policy adopted by central banks (Argyropoulos and Tzavalis, 2016). The slope and curvature of the yield curve hold significant “information about future economic activity” (Argyropoulos and Tzavalis, 2016).

Data and Summary Statistics

The data that was used for our models is a time series dataset of South Africa’s key macroeconomic indicators, spanning from 2001 to 2019. This dataset contains information about South Africa’s bond yields for different maturity dates and the gross domestic product of the country. Specifically, the following variables can be found in the dataset: the three-month bond yield named “three\_month”, the five-year bond yield named “five\_year”, the ten-year named “ten\_year” and the gross domestic product (GDP) named “gdp”. All four variables have a frequency of 64 observations.

A new variable was created to observe the year-on-year growth of the variable “gdp” called “gdp\_g”. This new variable represents the year-on-year growth of South African’s gross domestic product. Another variable, named “yield\_spread”, was created by taking the difference between the ten-year yield bond and the three-month yield bond for each observation. The data was analysed, and the results are shown below.

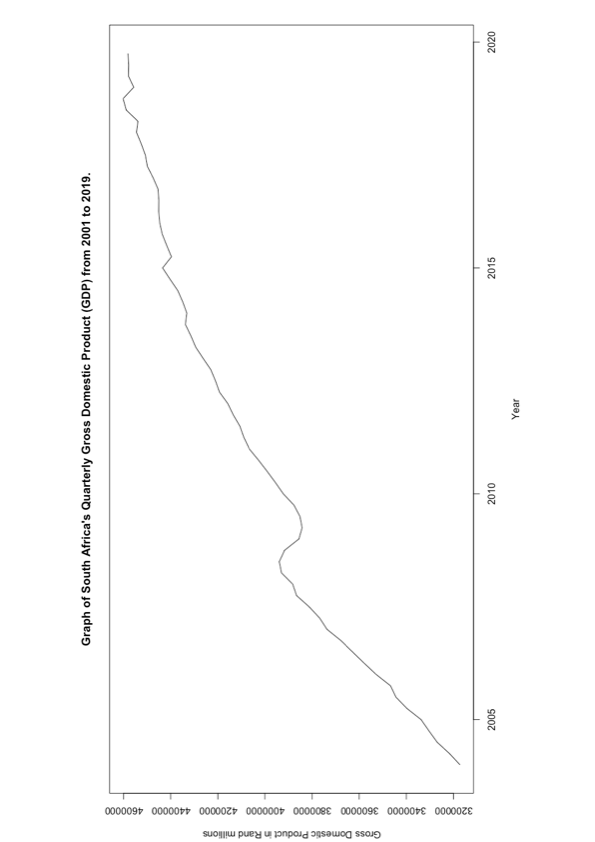


Figure 1: South Africa's Gross Domestic Product from 2001 to 2019

Chart, line chart

Description automatically generated

Figure 2: Graph of South Africa's Gross Domestic Product Growth Rate 2001 to 2019

Chart, line chart

Description automatically generated

Figure 3: South Africa's Government Bond Yield Spread from 2001 to 2019

Above shows Figures 1, 2 and 3, which are results from our dataset. Figure 1 is a line graph showing the gross domestic product of South Africa from 2001 to 2019. It as an increasing trend from 2001 to 2019, with a sharp decrease in 2008. This can be attributed to the 2008 global recession that affected all economies across the world. Figure 2 shows the growth rate of the country’s gross domestic product spanning the same time frame. A consistent downtrend in year-on-year gross domestic product growth can be observed, with the biggest decrease also coming from the year 2008. This decrease can also be explained by the 2008 global recession. During this year, a growth rate of -2% was observed. Figure 3 shows South Africa’s government bond yield spread for the same period. The bond yield spread of the country has been above 1% from 2001 to 2014, where it sharply dropped to -3%. After this sharp drop, however, the bond yield spread increased between 2015 and 2019, finally recovering to positive yield spreads from 2017 onwards. Furthermore, summary statistics were generated for the dataset, including the created variables.



Figure 4: Summary Statistics of the Given Data

From Figure 4, it can be observed that the three-month bond yield has an average of 7.267%, the five-year yield has an average of 7.881%, and the ten-year yield has an average of 8.364% from 2001 to 2019. The government bond’s yield spread has an average of 1.097% for the same period. A mean yield spread of 1.097% indicates that this would be a low-risk investment for buyers of government bonds. This is to be expected, as government bonds are investments much lower in risk compared to corporate bonds. Both the gross domestic product and its growth rate’s summary statistics are observed, with means of R4.079-trillion and 2.3% respectively. The relationship between the growth rate of the gross domestic product and the yield spread of the previous year was also analysed.

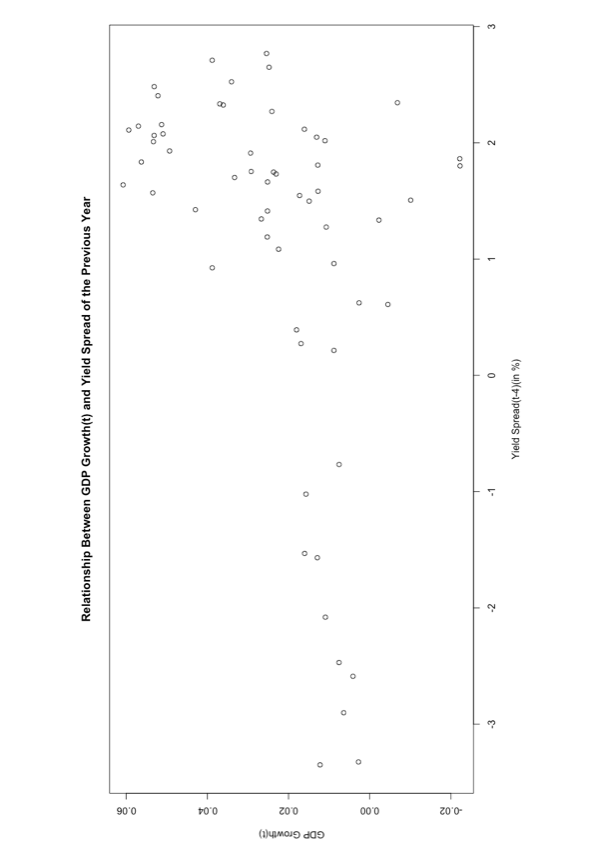


Figure 5: Scatter Plot Showing the Relationship Between GDP Growth and Yield Spread from the Previous Year

Figure 5 shows a scatter plot depicting the relationship between South Africa’s gross domestic product and its government bond yield spread from the previous year. From the scatter plot, it can be observed that there seems to be evidence of a strong positive relationship between the growth rate and the previous year’s yield spread. This scatter plot presents evidence for this article’s hypothesis that the yield curve does act as a leading indicator of economic growth.

Model and Results

Three models were selected to forecast the gross domestic product. The first one is an autoregressive model with lag 1 (AR (1)). This model has the gross domestic product’s growth rate as the dependent variable and its growth rate one year prior as the independent variable. The AR (1) model is formulated as:

∆GDPt = β0 + β1∆GDPt−4 + ut

The second model has the gross domestic product’s growth rate as the dependent variable and the lagged yield spread as the independent variable. This model is formulated follows:

∆GDPt = β0 + β1Yield Spreadt−4 + ut

The third model has the gross domestic product’s growth rate as the dependent variable, the lagged yield spread as one independent variable and the lagged growth rate as the other independent variable. This autoregressive distributed lag model is formulated as follows:

∆GDPt = β0 + β1∆GDPt−4 + β2Yield Spreadt−4 + ut

With regard to our models, it needs to be ensured that these models obey the time series assumptions. Where we may possibly have issues would be in ensuring that

E[u(t) |X] = 0, t = 0, 1, ..., n,

in other words, the error terms are uncorrelated with the explanatory variables for every period. Another possible issue that could arise would be from model where there is a deviation from the time series assumption that

Corr(ut , us) = 0 for all t ̸= s,

which states that there must be no correlation between the error terms over time.

The results of the three models can be seen below in Figure 6.

Table

Description automatically generated

Figure 6: Regression Models Used to Forecast the Growth Rate of GDP

For the first model, gdp\_growth\_lag’s coefficient is 0.921. This shows that a 1 unit increase in the lagged growth rate will result in a 0.921 unit increase in the growth rate. It also implies that there is a very strong autocorrelation between the GDP growth rate (gdp\_growth) and its lag gdp\_growth\_lag. With an R-squared value of 0.850, it can be observed that almost all the variation in gdp\_growth comes from the independent variable gdp\_growth\_lag.

For the second model, the coefficient of yield\_spread\_lag4 is 0.005. This shows that a one unit increase in the lagged yield spread will result in an insignificant increase of 0.005 units in the GDP growth rate. This implies that the independent variable of the lagged yield spread contains almost no information about the GDP growth rate. The second model has an R-squared value of 0.166, which suggests that only 16% of the variation in our dependent variable GDP growth is caused by the lagged yield spread.

The third and final model has the coefficients of the lagged GDP’s growth rate (gdp\_growth\_lag) and the lagged yield spread (yield\_spread\_lag4) as 0.923 and -0.00004 respectively. With an R-squared value of 0.850, it can be observed that 85% of the variation in the GDP growth rate can be explained by the independent variables (the lagged GDP growth rate and the lagged yield spread).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Models | Beta\_0 | Beta\_1 | Beta\_2 | Forecasted gdp growth rate | Forecast error |
| AR (1) Model | 0.0009255605 | 0.9211832 | - | -0.3212933 | 0.13 –  (-0.3212933)  =0.4512933 |
| Lagged Yield Spread model | 0.01791031 | 0.005184033 | - | 2.27712 | 0.13 –  (2.27712)  = -2.14712 |
| Ardl Model | 0.0009352716 | 0.922693 | 4.381712e-05 | -0.3251091 | 0.13 –  (-0.3251091)  = 0.4551091 |

Figure 7: Forecast Model Results for the Growth Rate of GDP

Therefore, it can be concluded that the additional distributed lags terms do not improve the autoregressive distributed lag models. Since the independent variable yield\_spread\_lag4 contains almost no information on the dependent variable gdp\_growth, both the lagged yield spread model

∆GDPt = β0 + β1Yield Spreadt−4 + ut

And the autoregressive distributed lag model

∆GDPt = β0 + β1∆GDPt−4 + β2Yield Spreadt−4 + ut

are the least preferable models. Therefore, the best model of the three is the AR (1) model

∆GDPt = β0 + β1∆GDPt−4 + ut.

Conclusion

From studying the current literature regarding yield curves, it is reasonable to conclude that yield curves are good predictors of economic activity. This is due to the fact that an increase or decrease in the yield curve normally precedes an economic expansion or recession. It also gives an indication as to the actions taken by central banks in response to the economic cycles.

The gross domestic product has increased consistently from 2001 to 2019, with the exception of the period during the 2008 global recession. The growth rate of gross domestic product, however, has decreased consistently, with South Africa reaching a negative rate during the 2008 recession. Government bond yield spreads have remained positive from 2001 until 2014 and again from 2017 to 2019. Government bond yields also averaged 1.097% throughout the 2001-2019 period, proving that they are a low-risk investment with minimal possibility of defaulting on the bond payments.

From the results of our modelling and forecasting, it can be concluded that the autoregressive model AR (1) is the best forecasting model for gross domestic product. The intuition behind this might lie in the fact that previous gross domestic product results will likely be the best forecast of immediate future results. As much as yield spreads are good predictors of economic growth, the predictive suitability relies on how long into the future it is trying forecast. It is most likely unsuitable for a period t +1 forecast, but for periods greater than t+1, it would be more suitable.

References

Argyropoulos, E. and Tzavalis, E., 2016. Forecasting economic activity from yield curve factors. The North American Journal of Economics and Finance, 36, pp.293-311.

Haubrich, J. G. and Dombrosky, A. M. (1996) ‘Predicting real growth using the yield curve’, Economic Review (00130281), 32(1), p. 26. Available at: <https://search-ebscohost-com.ezproxy.uct.ac.za/login.aspx?direct=true&db=aph&AN=9609200201&site=ehost-live>.

Kumar, R., Stauvermann, P. and Vu, H., 2021. The Relationship between Yield Curve and Economic Activity: An Analysis of G7 Countries. Journal of Risk and Financial Management, 14(2), p.62.

Appendix: R Code for Assignment

rm(list = ls())

# Removes all objects from the current workspace (R memory)

# Install readxl and stargazer

install.packages("stargazer")

install.packages("readxl")

install.packages("dplyr")

library(readxl)

library(stargazer)

library(dplyr)

setwd("~/Documents/University of Cape Town/Fourth Year/Quantitative Methods in Economics/Time Series Assignment/Dataset")

data <- read\_excel("assignment\_data.xlsx")

#Creating a GDP lag variable of 4 in order to calculate the year-on-year growth.

data$gdp\_lag4 <- dplyr::lag(data$gdp, 4)

#Calculating the year-on-year growth of GDP using the "gdp" and "gdp\_lag4" variables.

data$gdp\_growth <- (data$gdp - data$gdp\_lag4) / (data$gdp\_lag4)

#Creating a new variable called "yield\_spread", which takes the difference

#between the yield on the ten year government bond and the three month

#government bond.

data$yield\_spread <- (data$ten\_year - data$three\_month)

#Line graph of GDP

plot(y = data$gdp,

x = data$date,

type = "l",

main = "Graph of South Africa's Quarterly Gross Domestic Product (GDP) from 2001 to 2019.",

ylab = "Gross Domestic Product in Rand millions",

xlab = "Year")

#Line graph of growth rate of GDP

plot(y = data$gdp\_growth,

x = data$date,

type = "l",

main = "Graph of South Africa's Gross Domestic Product Growth Rate from 2001 to 2019.",

ylab = "Gross Domestic Product Growth Rate",

xlab = "Year")

#Line graph of yield spread

plot(y = data$yield\_spread,

x = data$date,

type = "l",

main = "South Africa's Government Bond Yield Spread from 2001 to 2019.",

ylab = "Yield Spread (in %)",

xlab = "Year")

#Creating summary statistics of all given and created variables.

summary(data[,c(2,3,4,5,7,8)])

stargazer(as.data.frame(data[,c(2,3,4,5,7,8)]), type = "text", out = "summary\_statistics.html")

#Finding median of all variables in data

median(data$three\_month)

median(data$five\_year)

median(data$ten\_year)

median(data$gdp)

median(data$gdp\_growth)

median(data$yield\_spread)

#Scatter Plot of relationship between GDP Growth and Yield Spread

data$yield\_spread\_lag4 <- dplyr::lag(data$yield\_spread, 4)

plot(y = data$gdp\_growth,

x = data$yield\_spread\_lag4,

main = "Relationship Between GDP Growth(t) and Yield Spread of the Previous Year",

ylab = "GDP Growth(t)",

xlab = "Yield Spread(t-4)(in %)")

#Generating an AR(1) model on gdp\_growth, but the lag of gdp\_growth is

#needed.

data$gdp\_growth\_lag <- dplyr::lag(data$gdp\_growth)

gdp\_growth\_ar1 <- lm(gdp\_growth ~ gdp\_growth\_lag, data = data)

stargazer::stargazer(gdp\_growth\_ar1, type = "text", out = "gdp\_growth\_ar1.html")

#Generating a lagged model with yield\_spread\_lag4 as an independent variable

gdp\_growth\_yield\_spread <- lm(gdp\_growth ~ yield\_spread\_lag4, data = data)

stargazer::stargazer(gdp\_growth\_yield\_spread, type = "text", out = "gdp\_growth\_yield\_spread.html")

#Generating a model for gdp\_growth with both lag of gdp\_growth and

#yield\_spread\_lag4

gdp\_growth\_ardl <- lm(gdp\_growth ~ gdp\_growth\_lag + yield\_spread\_lag4, data = data)

stargazer::stargazer(gdp\_growth\_ardl, type = "text", out = "gdp\_growth\_ardl.html")

#Combining all our models into one table for cohesion.

stargazer::stargazer(gdp\_growth\_ar1,

gdp\_growth\_yield\_spread,

gdp\_growth\_ardl,

type = "text",

out = "ar\_models.html")

#-------------------------------------------------------------------------------

#Creating a forecast for GDP growth using each model.

#-------------------------------------------------------------------------------

# AR(1) Model

ar1\_beta0 <- gdp\_growth\_ar1$coefficients[1]

ar1\_beta0 <- as.numeric(ar1\_beta0)

ar1\_beta1 <- gdp\_growth\_ar1$coefficients[2]

ar1\_beta1 <- as.numeric(ar1\_beta1)

forecast\_ar1 <- ar1\_beta0 + ar1\_beta1\*data$gdp\_growth[64]

forecast\_ar1 <- forecast\_ar1\*100

#-------------------------------------------------------------------------------

# Model with lagged yield spread

yield\_spread\_beta0 <- gdp\_growth\_yield\_spread$coefficients[1]

yield\_spread\_beta0 <- as.numeric(yield\_spread\_beta0)

yield\_spread\_beta1 <- gdp\_growth\_yield\_spread$coefficients[2]

yield\_spread\_beta1 <- as.numeric(yield\_spread\_beta1)

forecast\_yield\_spread <- yield\_spread\_beta0 + yield\_spread\_beta1\*data$yield\_spread[64]

forecast\_yield\_spread <- as.numeric(forecast\_yield\_spread)

forecast\_yield\_spread <- forecast\_yield\_spread\*100

#-------------------------------------------------------------------------------

# ARDL Model

ardl\_beta0 <- gdp\_growth\_ardl$coefficients[1]

ardl\_beta0 <- as.numeric(ardl\_beta0)

ardl\_beta1 <- gdp\_growth\_ardl$coefficients[2]

ardl\_beta1 <- as.numeric(ardl\_beta1)

ardl\_beta2 <- gdp\_growth\_ardl$coefficients[3]

ardl\_beta2 <- as.numeric(ardl\_beta2)

forecast\_ardl <- (ardl\_beta0 + ardl\_beta1\*data$gdp\_growth[64] +

ardl\_beta2\*data$yield\_spread[64])

forecast\_ardl <- forecast\_ardl\*100

#-------------------------------------------------------------------------------