Assignment_3_RutgerGeelen_WashantvanDam

Rutger Geelen (13510517) and Washant van Dam (13510053)

December 8th, 2020

Preparation

Clear workspace and load the required packages:

```
rm(list=ls())
suppressMessages(library("lmtest"))
suppressMessages(library("car"))
library("sandwich")
suppressMessages(library("ggplot2"))
library("rmarkdown")
library ("survival")
suppressMessages (library ("stargazer"))
library ("rbibutils")
library ("plm")
library ("vars")
```

```
## Loading required package: MASS
```

```
## Loading required package: strucchange
```

```
## Loading required package: urca
```

```
library("urca")
library("dynlm")
```

Q1

We analyse annual data on the S&P 500 stock market index and related time series, with observations over the period 1871–2015. The data are given in SP500.csv, and the variables are defined as follows:
- P S&P 500 index (value-weighted average of stock prices of 500 important US companies)
- D DAnnual (value-weighted average) dividend per share on the stocks in the index
- Rs Short-term (one-year) US interest rate
- Rl Long-term (10-year) US interest rate
- CPI Consumper price index
- lrp = ln(P/CPI), logarithm of "real" index, i.e., corrected for changes in consumer prices
- lrd = ln(D/CPI), logarithm of real dividends
- lpd = ln(P/D), logarithm of price-dividend ratio
- ret = (Pt + Dt - Pt-1)/Pt-1, annual return on the index, including dividends
- TSpr = Rl-Rs, difference between long- and short-term interest rate (term spread)The file also contains up to 3 lags of the (lower case) variables, indicated by the extension "j",j= 1,2,3.

Read data:

```
MyQ1Data <- read.csv("https://raw.githubusercontent.com/rutgerg/econometrics_assigment_3/master/SP500.cs
MyQ1Data = na.omit(MyQ1Data)
```

Q1A

Test for a unit root in lrp, lrd, Rs and Rl. Motivate your choice between either a constant only, or a constant and a linear trend in the test regression. Report and interpret the outcome of the test.

If all roots are greater than 1 in absolute value, the AR(p) series is stationary. If at least one root equals 1, the AR(p) is said to have a unit root and thus has a stochastic trend.

We use The ADF test for a unit autoregressive root tests the hypothesis H0:d=0 (stochastic trend) against the one-sided alternative H1:d<0 (stationarity) using the usual OLS t-statistic.

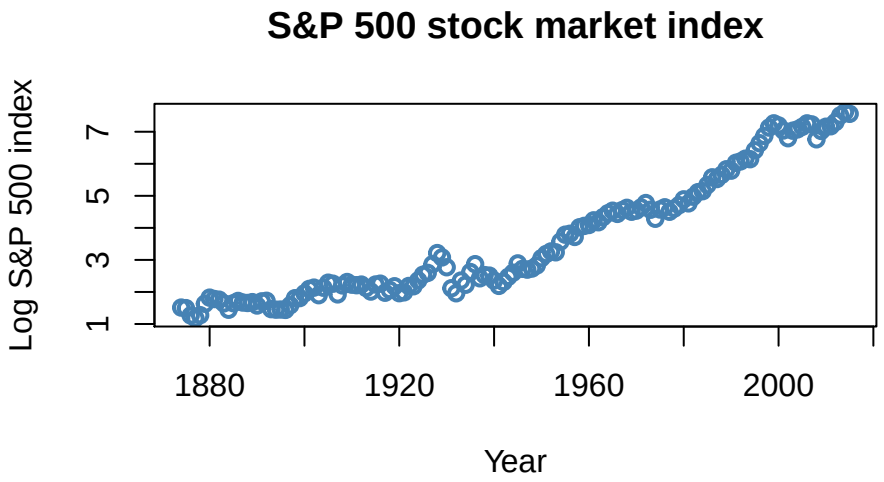
General specigfication: deltaPt = b0 + (a * t) + d * Pt-1 + g1 * deltaPt-1 + ... + gp-1 * deltaPt-p+1 + ut

b0 is intercept (a * t) is linear time trend

d = B1 + .. + Bp - 1

H0: d = 0 (unit root) H1: d < 1 (stationarity)

Intercept only specification only if there is not long term growth in the series (eg interest, inflation, unemployment). Since this is index data we expect intrecept and linear time trend. To check we first plot the data:



Since the trend is upward we choose the intercept and time trend specification.

Let's test:

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -452.10  -12.62   -2.84    6.96  264.37
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -15.36025   15.21563  -1.010  0.31464
```

```
## z.lag.1       0.05012    0.02440   2.054  0.04202 *
## tt            0.36472    0.22830   1.598  0.11260
## z.diff.lag1   0.08607    0.08899   0.967  0.33528
## z.diff.lag2  -0.28056    0.08830  -3.177  0.00186 **
## z.diff.lag3  -0.10939    0.09229  -1.185  0.23812
## z.diff.lag4  -0.14654    0.08990  -1.630  0.10555
## z.diff.lag5  -0.42753    0.09094  -4.701 6.59e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 73.86 on 128 degrees of freedom
## Multiple R-squared:  0.2661, Adjusted R-squared:  0.226
## F-statistic:  6.63 on 7 and 128 DF,  p-value: 1.08e-06
##
##
## Value of test-statistic is: 2.0539 8.1619 9.0781
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

We need to recalculate the p-value since we can not use the p-value from the linear regression above which assumes normal distribution and 2 sided test.

```
## p-value of ADF test: 0.9999995
```

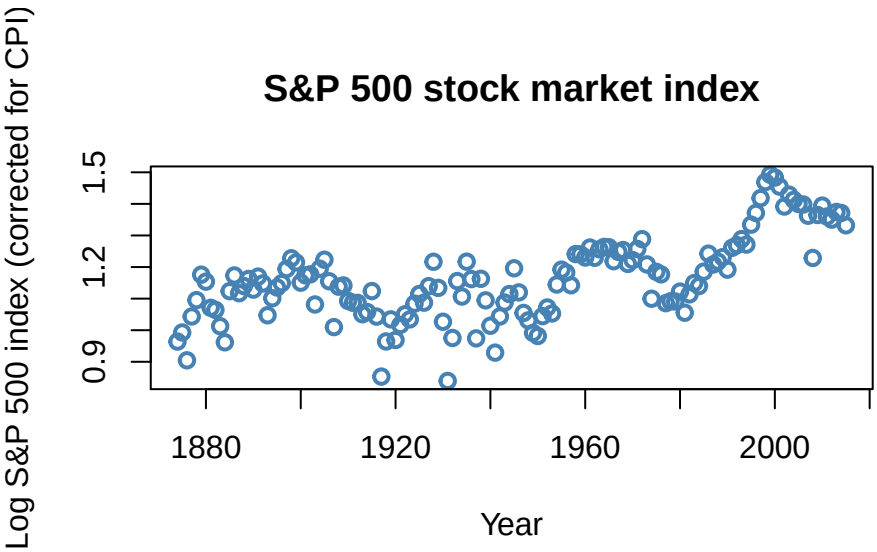
Since Padf-value is 0.99 we do not reject H0 so there is non stationary data and have unit root.

Q1B

Particular financial theories imply that log (real) stock prices and log (real) dividends should be cointegrated, with a coefficient of 1. Explain that, under the assumption that both prices and dividends have a unit root, this hypothesis can be tested with a unit root test on the variable lpd, and report and interpret the outcome of this test.(The assumption of a unit root in both series may not be supported by your answer to (a); this would have to be included in your discussion.)

Cointegration happens when two or more series share the same stochastic trend. In this csae that makes sense since stockprices rise (cp) with rising dividends because it assumes higher future cashflows. Cointegration says that if both series are non-stationary then difference can be stable or stationary and H0 should be reject in favor of H1.

Coefficient of 1 assumes thèta = 1



To check we first plot the data:
It looks like a positive trend; not stationary.

Let's test:

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62235 -0.11534  0.00943  0.11861  0.46315
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.3987696  0.1548224   2.576  0.01111 *
## z.lag.1     -0.1493394  0.0552352  -2.704  0.00777 **
## tt           0.0012557  0.0005625   2.232  0.02728 *
## z.diff.lag1 -0.0596104  0.0850817  -0.701  0.48478
## z.diff.lag2 -0.2728431  0.0833521  -3.273  0.00136 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1903 on 131 degrees of freedom
## Multiple R-squared:  0.1751, Adjusted R-squared:   0.15
## F-statistic: 6.954 on 4 and 131 DF,  p-value: 4.158e-05
##
##
## Value of test-statistic is: -2.7037 2.6161 3.8103
##
## Critical values for test statistics:
##       1pct  5pct 10pct
```

```
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

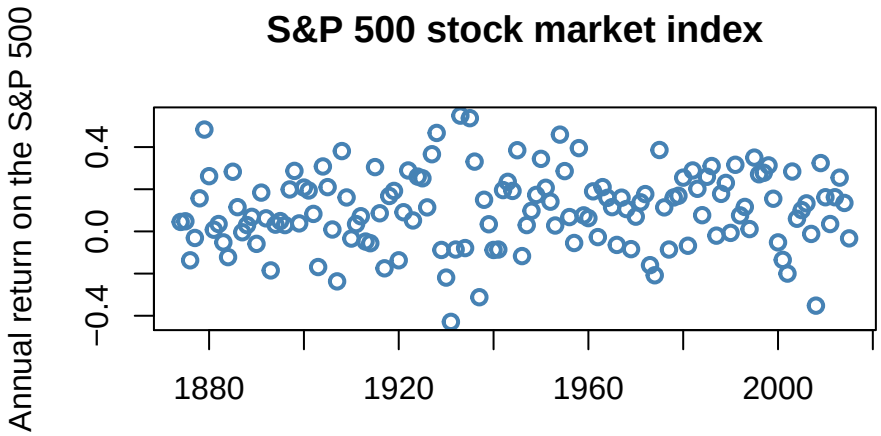
We need to recalculate the p-value since we can not use the p-value from the linear regression above which assumes normal distribution and 2 sided test.

```
## p-value of ADF test: 0.2349874
```

Since Padf-value is 0.23 we do not reject H0 so there is non stationary data and have unit root. There is not enough evidence to support cointegration. IS THIS A CORRECT CONCLUSION?

Q1C

We now wish to investigate if the annual returns on the S&P 500 index can be forecasted. First, plot the autocorrelation function of ret, and interpret the outcome. Next, estimate an AR(2) model for ret and test if the lagged returns have zero coefficients (jointly).



First we plot the data: The outcome suggest ret is constant and stationary; there is no trend.

Let's test:

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      2      2      2      2
##
## $criteria
##                  1           2
## AIC(n) -3.40220753 -3.42934189
## HQ(n)  -3.38513045 -3.40372627
## SC(n)  -3.36018407 -3.36630669
## FPE(n)  0.03329974  0.03240847
```

FUNCTION ABOVE GIVE ERRROR? IK WIL MATRIX VAN JAAR EN ret MAAR DAT LUKT NIET.

Q1D

Estimate an ADL model with ret as the dependent variable, and with three lags of ret and the term spread TSpr as explanatory variables. You may assume that TSpr is stationary. Carry out a Granger-causality test to see if the term spread Granger-causes returns. Interpret the outcome.

The Granger causality test Granger is an F test of the null hypothesis that all lags of a variable X included in a time series regression model do not have predictive power for Yt. The Granger causality test does not

test whether X actually causes Y but whether the included lags are informative in terms of predicting Y.

F-test for H0 : d1 = ... = dq = 0 (non-causality) in Yt =b0 +B1 * Yt-1 +...+bp * Yt-p + d1 * Xt-1 +...+ dq * Xt-q + ut.

ADL(3,2)

Uncommenting line 167 gives error: "Error in linearHypothesis.lm(aq1d, c("L(ret,3)=0","L(TSpr)=0"), vcov. = sandwich) : residual sum of squares is 0 (within rounding error)" must be soemthing wrong

Q1E

Select an appropriate lag order p (motivate your choice), and estimate the resulting VAR(p)model for ret and TSpr jointly. Report and interpret the outcomes, focusing in particular on the difference in predictability of the two time series.