Heteroscedasticidad y MCP

J. Ramajo

2020

library(readr)  
library(car)

## Loading required package: carData

library(MASS)  
library(effects)

## Registered S3 methods overwritten by 'lme4':  
## method from  
## cooks.distance.influence.merMod car   
## influence.merMod car   
## dfbeta.influence.merMod car   
## dfbetas.influence.merMod car

## lattice theme set by effectsTheme()  
## See ?effectsTheme for details.

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v dplyr 1.0.2  
## v tibble 3.0.4 v stringr 1.4.0  
## v tidyr 1.1.2 v forcats 0.5.0  
## v purrr 0.3.4

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()  
## x dplyr::recode() masks car::recode()  
## x dplyr::select() masks MASS::select()  
## x purrr::some() masks car::some()

library(gvlma)  
library(tseries)

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

library(stats)  
library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

#  
DEM\_HET <- read\_csv("DEM\_HET.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## PA = col\_double(),  
## PB = col\_double(),  
## PC = col\_double(),  
## QA = col\_double(),  
## Y = col\_double()  
## )

View(DEM\_HET)  
head(DEM\_HET, n=10)

## # A tibble: 10 x 5  
## PA PB PC QA Y  
## <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 100 100 100 100 100   
## 2 102. 84.6 85.8 93.7 108.  
## 3 89.2 97.9 91.8 119. 111.  
## 4 97.1 94.8 99.8 98.9 117.  
## 5 105. 90.4 104. 99.6 119.  
## 6 103. 112. 98.1 117. 127.  
## 7 113. 94.2 105. 99.9 133.  
## 8 113. 94.7 100. 101. 135.  
## 9 112. 94.2 103. 115. 154.  
## 10 114. 86.8 129. 95.7 155.

dim(DEM\_HET)

## [1] 30 5

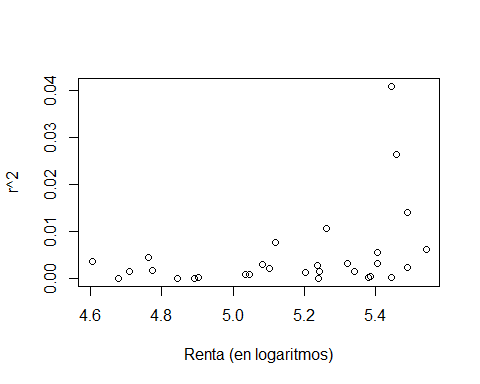
summary(DEM\_HET)

## PA PB PC QA   
## Min. : 89.15 Min. : 51.04 Min. : 85.85 Min. : 93.01   
## 1st Qu.:102.52 1st Qu.: 74.72 1st Qu.:103.85 1st Qu.: 99.65   
## Median :111.77 Median : 81.49 Median :130.84 Median :105.93   
## Mean :109.02 Mean : 81.91 Mean :127.68 Mean :109.59   
## 3rd Qu.:114.53 3rd Qu.: 90.13 3rd Qu.:145.07 3rd Qu.:117.28   
## Max. :124.85 Max. :111.66 Max. :182.04 Max. :140.97   
## Y   
## Min. :100.0   
## 1st Qu.:139.5   
## Median :188.4   
## Mean :180.6   
## 3rd Qu.:221.0   
## Max. :254.9

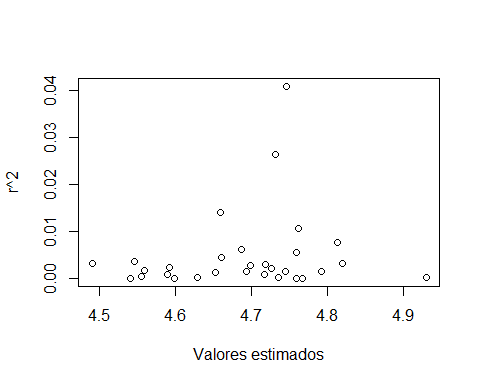
#  
# Función de demanda estándar  
#  
S(lm\_dem <- lm(log(QA) ~ log(PA) + log(PB) + log(PC)+ log(Y), data = DEM\_HET))

## Call: lm(formula = log(QA) ~ log(PA) + log(PB) + log(PC) + log(Y), data =  
## DEM\_HET)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.2747 1.2052 3.547 0.001570 \*\*   
## log(PA) -0.8901 0.2157 -4.126 0.000358 \*\*\*  
## log(PB) 0.5542 0.1300 4.262 0.000252 \*\*\*  
## log(PC) -0.3491 0.1831 -1.906 0.068154 .   
## log(Y) 0.7440 0.1352 5.503 1.02e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard deviation: 0.07619 on 25 degrees of freedom  
## Multiple R-squared: 0.6643  
## F-statistic: 12.37 on 4 and 25 DF, p-value: 1.105e-05   
## AIC BIC   
## -62.81 -54.40

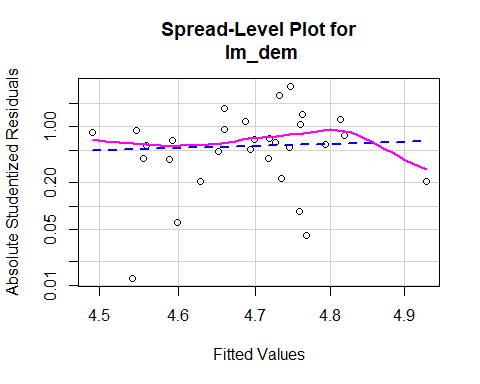
#  
# Chequeo de varianza no constante: heteroscedasticidad  
r2 <- resid(lm\_dem)^2  
yhat <- fitted(lm\_dem)  
l\_Y <- log(DEM\_HET$Y)  
plot(l\_Y,r2, xlab="Renta (en logaritmos)", ylab="r^2")



plot(yhat,r2, xlab="Valores estimados", ylab="r^2")



#  
spreadLevelPlot(lm\_dem)



##   
## Suggested power transformation: -1.784909

#  
ncvTest(lm\_dem, var.formula = ~ log(Y)) # Test de Breusch-Pagan (Score test)

## Non-constant Variance Score Test   
## Variance formula: ~ log(Y)   
## Chisquare = 6.205347, Df = 1, p = 0.012736

ncvTest(lm\_dem)

## Non-constant Variance Score Test   
## Variance formula: ~ fitted.values   
## Chisquare = 1.28873, Df = 1, p = 0.25628

#  
# Test de Breusch-Pagan (versión clásica)  
alpha <- 0.05  
# Regresión auxiliar:  
S(lm\_r2 <- lm(r2~ log(Y), data=DEM\_HET))

## Call: lm(formula = r2 ~ log(Y), data = DEM\_HET)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.05395 0.02832 -1.905 0.0671 .  
## log(Y) 0.01139 0.00548 2.078 0.0469 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard deviation: 0.008199 on 28 degrees of freedom  
## Multiple R-squared: 0.1337  
## F-statistic: 4.32 on 1 and 28 DF, p-value: 0.04695   
## AIC BIC   
## -199.16 -194.96

N <- nobs(lm\_r2)  
p <- 1   
slm\_r2 <- summary(lm\_r2)  
R2\_lm\_r2 <- slm\_r2$r.squared  
BP <- N\*R2\_lm\_r2  
# Contraste Chi-cuadrado   
chisqcr <- qchisq(1-alpha, p)  
pval <- 1-pchisq(BP,p)  
BP ; chisqcr

## [1] 4.009848

## [1] 3.841459

pval

## [1] 0.04523524

#  
bptest(lm\_dem) # Breusch-Pagan robusto (variante de Koenker)

##   
## studentized Breusch-Pagan test  
##   
## data: lm\_dem  
## BP = 4.3822, df = 4, p-value = 0.3567

bptest(lm\_dem, varformula = ~ log(Y), data=DEM\_HET)

##   
## studentized Breusch-Pagan test  
##   
## data: lm\_dem  
## BP = 4.0098, df = 1, p-value = 0.04524

bptest(lm\_dem, studentize = FALSE) # Breusch-Pagan estándar (escalado)

##   
## Breusch-Pagan test  
##   
## data: lm\_dem  
## BP = 6.7816, df = 4, p-value = 0.1479

bptest(lm\_dem, studentize = FALSE, varformula = ~ log(Y), data=DEM\_HET)

##   
## Breusch-Pagan test  
##   
## data: lm\_dem  
## BP = 6.2053, df = 1, p-value = 0.01274

#  
# Corrección de la heteroscedasticidad  
#  
# Mínimos cuadrados corregidos (White)  
#  
S(lm\_dem, vcov.=hccm(lm\_dem, type = "hc1")) # (Corrección de White: errores estándar robustos)

## Call: lm(formula = log(QA) ~ log(PA) + log(PB) + log(PC) + log(Y), data =  
## DEM\_HET)  
## Standard errors computed by hccm(lm\_dem, type = "hc1")   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.27468 0.89931 4.753 7.06e-05 \*\*\*  
## log(PA) -0.89012 0.13561 -6.564 7.08e-07 \*\*\*  
## log(PB) 0.55419 0.08331 6.652 5.70e-07 \*\*\*  
## log(PC) -0.34909 0.15058 -2.318 0.0289 \*   
## log(Y) 0.74395 0.11879 6.263 1.49e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard deviation: 0.07619 on 25 degrees of freedom  
## Multiple R-squared: 0.6643  
## F-statistic: 26.61 on 4 and 25 DF, p-value: 1.086e-08   
## AIC BIC   
## -62.81 -54.40

#  
# Mínimos cuadrados ponderados (MCP)  
#  
# Regresión auxiliar para la varianza estimada:  
S(lm\_lr2 <- lm(log(r2)~ log(Y), data=DEM\_HET))

## Call: lm(formula = log(r2) ~ log(Y), data = DEM\_HET)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -26.658 7.565 -3.524 0.00148 \*\*  
## log(Y) 3.846 1.464 2.628 0.01379 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard deviation: 2.19 on 28 degrees of freedom  
## Multiple R-squared: 0.1978  
## F-statistic: 6.905 on 1 and 28 DF, p-value: 0.01379   
## AIC BIC   
## 136.1 140.3

s2 <- exp(fitted(lm\_lr2))  
S(lm\_dem\_het <- lm(log(QA) ~ log(PA) + log(PB) + log(PC)+ log(Y), weights = 1/s2, data = DEM\_HET))

## Call: lm(formula = log(QA) ~ log(PA) + log(PB) + log(PC) + log(Y), data =  
## DEM\_HET, weights = 1/s2)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.48077 0.87248 5.136 2.62e-05 \*\*\*  
## log(PA) -0.90716 0.14931 -6.076 2.38e-06 \*\*\*  
## log(PB) 0.54594 0.10265 5.318 1.64e-05 \*\*\*  
## log(PC) -0.24882 0.12394 -2.007 0.0556 .   
## log(Y) 0.63130 0.08844 7.138 1.76e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard deviation: 2.026 on 25 degrees of freedom  
## Multiple R-squared: 0.7734  
## F-statistic: 21.33 on 4 and 25 DF, p-value: 9.299e-08   
## AIC BIC   
## -70.17 -61.77

#