Elección del tipo de interés de un préstamo hipotecario

J. Ramajo

2020

library(readr)  
library(AER)

## Loading required package: car

## Loading required package: carData

## Loading required package: lmtest

## Loading required package: zoo

##   
## Attaching package: 'zoo'

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

## Loading required package: sandwich

## Loading required package: survival

library(alr4)

## Loading required package: 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(margins)  
PREST\_HIP <- read\_csv("PREST\_HIP.csv")

## Parsed with column specification:  
## cols(  
## BA = col\_double(),  
## BS = col\_double(),  
## CB = col\_double(),  
## FI = col\_double(),  
## FTB = col\_double(),  
## LA = col\_double(),  
## MARG = col\_double(),  
## MAT = col\_double(),  
## MC = col\_double(),  
## MOB = col\_double(),  
## NW = col\_double(),  
## PTS = col\_double(),  
## SE = col\_double(),  
## STL = col\_double(),  
## Y = col\_double(),  
## YLD = col\_double()  
## )

attach(PREST\_HIP)  
# Modelo logit  
ti.logit <- glm(Y ~ FI + MARG + YLD + PTS + MAT + BA + BS + FTB + CB + MC + SE + MOB + NW + LA + STL, family=binomial(link="logit"))  
summary(ti.logit)

##   
## Call:  
## glm(formula = Y ~ FI + MARG + YLD + PTS + MAT + BA + BS + FTB +   
## CB + MC + SE + MOB + NW + LA + STL, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1639 -0.4782 -0.1384 0.6414 2.1456   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.23887 10.47904 -0.405 0.6858   
## FI 1.71750 0.72773 2.360 0.0183 \*  
## MARG -1.18952 0.48545 -2.450 0.0143 \*  
## YLD -4.62670 2.91972 -1.585 0.1130   
## PTS -1.49980 0.71993 -2.083 0.0372 \*  
## MAT -2.05707 1.63112 -1.261 0.2073   
## BA -0.01048 0.07569 -0.138 0.8899   
## BS -0.19825 0.17245 -1.150 0.2503   
## FTB 0.11092 0.98371 0.113 0.9102   
## CB -1.75974 1.24217 -1.417 0.1566   
## MC -1.92294 1.18297 -1.626 0.1041   
## SE -2.20851 2.80100 -0.788 0.4304   
## MOB -0.15308 0.09700 -1.578 0.1145   
## NW 0.24406 0.18503 1.319 0.1872   
## LA 0.02207 0.06101 0.362 0.7176   
## STL 0.03156 0.05172 0.610 0.5417   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 105.604 on 77 degrees of freedom  
## Residual deviance: 61.792 on 62 degrees of freedom  
## AIC: 93.792  
##   
## Number of Fisher Scoring iterations: 6

# Modelo probit  
ti.probit <- glm(Y ~ FI + MARG + YLD + PTS + MAT + BA + BS + FTB + CB + MC + SE + MOB + NW + LA + STL, family=binomial(link="probit"))  
summary(ti.probit)

##   
## Call:  
## glm(formula = Y ~ FI + MARG + YLD + PTS + MAT + BA + BS + FTB +   
## CB + MC + SE + MOB + NW + LA + STL, family = binomial(link = "probit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.08188 -0.49604 -0.07784 0.64347 2.08997   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.107496 5.877216 -0.529 0.5970   
## FI 1.008059 0.410667 2.455 0.0141 \*   
## MARG -0.705235 0.272291 -2.590 0.0096 \*\*  
## YLD -2.525164 1.588075 -1.590 0.1118   
## PTS -0.830267 0.397719 -2.088 0.0368 \*   
## MAT -1.164395 0.894643 -1.302 0.1931   
## BA -0.003978 0.042864 -0.093 0.9261   
## BS -0.108266 0.099792 -1.085 0.2780   
## FTB 0.143454 0.558299 0.257 0.7972   
## CB -1.066544 0.692153 -1.541 0.1233   
## MC -1.058573 0.672743 -1.574 0.1156   
## SE -1.127484 1.559724 -0.723 0.4698   
## MOB -0.093035 0.054982 -1.692 0.0906 .   
## NW 0.128775 0.105279 1.223 0.2213   
## LA 0.014615 0.034992 0.418 0.6762   
## STL 0.016129 0.028296 0.570 0.5687   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 105.604 on 77 degrees of freedom  
## Residual deviance: 61.459 on 62 degrees of freedom  
## AIC: 93.459  
##   
## Number of Fisher Scoring iterations: 7

# Significación de las características personales  
linearHypothesis ( ti.logit , c("BA = 0", "BS = 0","FTB = 0","CB = 0","MC = 0","SE = 0","MOB = 0","NW = 0","LA = 0","STL = 0"))

## Linear hypothesis test  
##   
## Hypothesis:  
## BA = 0  
## BS = 0  
## FTB = 0  
## CB = 0  
## MC = 0  
## SE = 0  
## MOB = 0  
## NW = 0  
## LA = 0  
## STL = 0  
##   
## Model 1: restricted model  
## Model 2: Y ~ FI + MARG + YLD + PTS + MAT + BA + BS + FTB + CB + MC + SE +   
## MOB + NW + LA + STL  
##   
## Res.Df Df Chisq Pr(>Chisq)  
## 1 72   
## 2 62 10 12.728 0.2393

#  
ti.logit.1 <- glm(Y ~ FI + MARG + YLD + PTS + MAT, family=binomial(link="logit"))  
summary(ti.logit.1)

##   
## Call:  
## glm(formula = Y ~ FI + MARG + YLD + PTS + MAT, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1131 -0.8118 -0.4449 0.9597 2.4977   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -6.7318 7.0594 -0.954 0.34029   
## FI 1.2646 0.4540 2.785 0.00535 \*\*  
## MARG -0.7178 0.3138 -2.287 0.02218 \*   
## YLD -4.8275 1.9588 -2.465 0.01372 \*   
## PTS -0.3590 0.4234 -0.848 0.39641   
## MAT -0.5503 1.0366 -0.531 0.59549   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 105.604 on 77 degrees of freedom  
## Residual deviance: 82.946 on 72 degrees of freedom  
## AIC: 94.946  
##   
## Number of Fisher Scoring iterations: 4

linearHypothesis ( ti.logit.1 , c("PTS = 0", "MAT = 0"))

## Linear hypothesis test  
##   
## Hypothesis:  
## PTS = 0  
## MAT = 0  
##   
## Model 1: restricted model  
## Model 2: Y ~ FI + MARG + YLD + PTS + MAT  
##   
## Res.Df Df Chisq Pr(>Chisq)  
## 1 74   
## 2 72 2 0.959 0.6191

ti.logit.2 <- glm(Y ~ FI + MARG + YLD , family=binomial(link="logit"))  
S(ti.logit.2)

## Call: glm(formula = Y ~ FI + MARG + YLD, family = binomial(link = "logit"))  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -8.1920 6.8807 -1.191 0.23382   
## FI 1.1846 0.4387 2.700 0.00693 \*\*  
## MARG -0.6608 0.2852 -2.317 0.02049 \*   
## YLD -4.0305 1.7131 -2.353 0.01864 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 105.604 on 77 degrees of freedom  
## Residual deviance: 83.959 on 74 degrees of freedom  
##   
## logLik df AIC BIC   
## -41.98 4 91.96 101.39   
##   
## Number of Fisher Scoring iterations: 4  
##   
## Exponentiated Coefficients and Confidence Bounds  
## Estimate 2.5 % 97.5 %  
## (Intercept) 0.000276863 2.023320e-10 150.0833242  
## FI 3.269534462 1.444690e+00 8.2514763  
## MARG 0.516458335 2.809716e-01 0.8737826  
## YLD 0.017765498 4.893026e-04 0.4380433

coeftest(ti.logit.2, vcov. = vcovHC, type = "HC1")

##   
## z test of coefficients:  
##   
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -8.19199 8.03329 -1.0198 0.30784   
## FI 1.18465 0.57315 2.0669 0.03874 \*  
## MARG -0.66076 0.30026 -2.2006 0.02776 \*  
## YLD -4.03050 1.83833 -2.1925 0.02834 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Anova(ti.logit.2)

## Analysis of Deviance Table (Type II tests)  
##   
## Response: Y  
## LR Chisq Df Pr(>Chisq)   
## FI 8.3553 1 0.003846 \*\*  
## MARG 6.2076 1 0.012720 \*   
## YLD 6.1767 1 0.012945 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

c(dev=deviance(ti.logit.2), df=df.residual(ti.logit.2))

## dev df   
## 83.95927 74.00000

#  
# Ajuste del modelo  
#  
pseudoR2 <- 1 - (ti.logit.2$deviance) / (ti.logit.2$null.deviance)  
pseudoR2

## [1] 0.2049648

#  
ti.logit.0 <- glm(Y ~ 1 , family=binomial(link="logit"))  
1 - logLik(ti.logit.2)[1]/logLik(ti.logit.0)[1]

## [1] 0.2049648

#  
# Tabla de éxito-fracaso  
#  
table(true=Y, predicted=round(fitted(ti.logit.2)))

## predicted  
## true 0 1  
## 0 37 9  
## 1 11 21

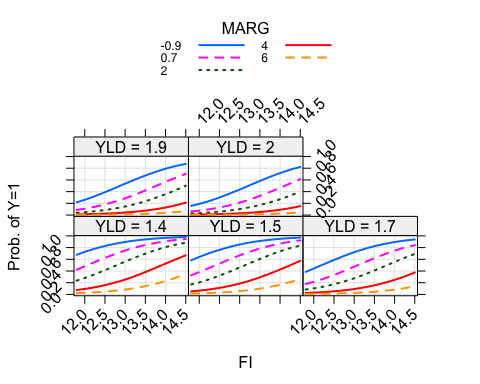
#  
# Efectos marginales  
#  
margins(ti.logit.2)

## Average marginal effects

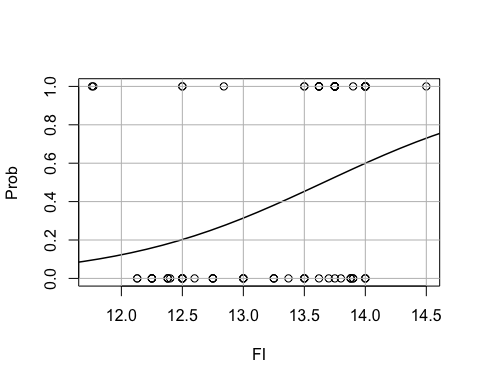
## glm(formula = Y ~ FI + MARG + YLD, family = binomial(link = "logit"))

## FI MARG YLD  
## 0.2138 -0.1192 -0.7272

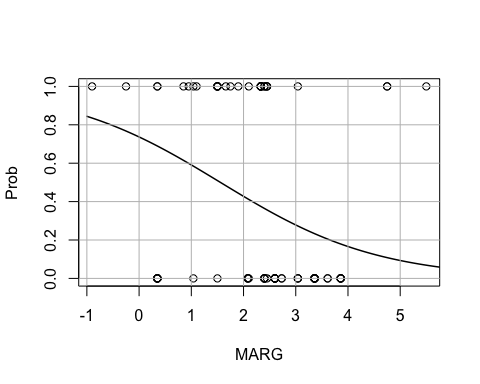
#  
# Gráficas de efectos marginales  
#  
effs <- Effect(c("FI", "MARG","YLD"), ti.logit.2)  
plot(effs, multiline=TRUE, grid=TRUE, lines=c(1, 2, 3),   
 xlab="FI",main="", rotx=45, roty = 45,  
 ylab="Prob. of Y=1", rescale.axis=FALSE, rug=FALSE)



#  
# Gráfica de la función de probabilidad estimada (para la variable FI)  
#   
plot( Y ~ FI, PREST\_HIP, xlab="FI", ylab="Prob", ylim=c(0,1))  
FInew <- seq(10, 20, length=78)  
lines(FInew, predict(ti.logit.2, newdata=data.frame(FI=FInew, MARG=rep(mean(MARG), 78), YLD=rep(mean(YLD), 78)), type="response"), lwd=1.5)  
grid(col="gray", lty="solid")



#  
# Gráfica de la función de probabilidad estimada (para la variable MARG)  
#   
plot( Y ~ MARG, PREST\_HIP, xlab="MARG", ylab="Prob", ylim=c(0,1))  
MARGnew <- seq(-1, 6, length=78)  
lines(MARGnew, predict(ti.logit.2, newdata=data.frame(FI=rep(mean(FI), 78), MARG=MARGnew, YLD=rep(mean(YLD), 78)), type="response"), lwd=1.5)  
grid(col="gray", lty="solid")



#  
# Gráfica de la función de probabilidad estimada (para la variable YLD)  
#   
plot( Y ~ YLD, PREST\_HIP, xlab="YLD", ylab="Prob", ylim=c(0,1))  
YLDnew <- seq(1, 2.5, length=78)  
lines(YLDnew, predict(ti.logit.2, newdata=data.frame(FI=rep(mean(FI), 78), MARG=rep(mean(MARG), 78), YLD=YLDnew), type="response"), lwd=1.5)  
grid(col="gray", lty="solid")

