Name: Khoa Tran

Assignment 10

1.

Text

Description automatically generated

2.

(1)

data <- transform(data,ecobuy = as.numeric(ecolbs>0))

(2)

> #percent correctly predicted for LPM, probit, logit models

> #define hhsize as a categorical variable

> data <- transform(data,hhsize.f = as.factor(hhsize))

> #LPM

> lpm = lm(ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,data=data)

> data<-transform(data,ecobuy\_fitted\_lp=fitted.values(lpm))

> data<-transform(data,ecobuy\_tilde\_lp=as.numeric(ecobuy\_fitted\_lp>=0.5))

> data<-transform(data,matched\_lp=as.numeric(ecobuy\_tilde\_lp==ecobuy))

> obsn <- lpm$df.residual+nrow(summary(lpm)$coef)

> correctly\_pred\_lpm=sum(data$matched\_lp)/obsn

> correctly\_pred\_lpm

[1] 0.6727273

> #probit

> probit <- glm(ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,family=binomial(link='probit'),data=data)

> data<-transform(data,ecobuy\_fitted\_probit=fitted.values(probit))

> data<-transform(data,ecobuy\_tilde\_probit=as.numeric(ecobuy\_fitted\_probit>=0.5))

> data<-transform(data,matched\_probit=as.numeric(ecobuy\_tilde\_probit==ecobuy))

> obsn <- probit$df.residual+nrow(summary(probit)$coef)

> correctly\_pred\_prob=sum(data$matched\_probit)/obsn

> correctly\_pred\_prob

[1] 0.6727273

> #logit

> logit <- glm(ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,family=binomial(link='logit'),data=data)

> data<-transform(data,ecobuy\_fitted\_logit=fitted.values(logit))

> data<-transform(data,ecobuy\_tilde\_logit=as.numeric(ecobuy\_fitted\_logit>=0.5))

> data<-transform(data,matched\_logit=as.numeric(ecobuy\_tilde\_logit==ecobuy))

> obsn <- logit$df.residual+nrow(summary(logit)$coef)

> correctly\_pred\_log=sum(data$matched\_logit)/obsn

> correctly\_pred\_log

[1] 0.6727273

(3)

Probit AME for ecoprc:

> ame\_prob=probitmfx(formula=ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,data=data,atmean=FALSE)

> ame\_prob$mfxest[1]

[1] -0.7651859

Probit PEA for ecoprc:

> pea\_prob=probitmfx(formula=ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,data=data,atmean=TRUE)

> pea\_prob$mfxest[1]

[1] -0.8399909

Logit AME for ecoprc:

> ame\_log=logitmfx(formula=ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,data=data,atmean=FALSE)

> ame\_log$mfxest[1]

[1] -0.7525779

Logit PEA for ecoprc:

> pea\_log=logitmfx(formula=ecobuy~ecoprc+regprc+faminc+hhsize.f+educ+age,data=data,atmean=TRUE)

> pea\_log$mfxest[1]

[1] -0.8348828

-> Discussion: AME measures the average marginal effect i.e. sum up all the marginal effect for all observations and then take the average. While the PEA measures the marginal effect of the average measurements (i.e. first take the average for all observations then calculate the marginal effect.)

(4)

> stargazer(lpm, probit, logit,

+ add.lines=list(c("% correctly predicted LPM:",correctly\_pred\_lpm),

+ c("% correctly predicted Probit:",correctly\_pred\_prob),

+ c("% correctly predicted Logit:",correctly\_pred\_log)),

+ type="text",title="Table. Ecobuy")

Table. Ecobuy

=========================================================================

Dependent variable:

-------------------------------------------

ecobuy

OLS probit logistic

(1) (2) (3)

-------------------------------------------------------------------------

ecoprc -0.789\*\*\* -2.244\*\*\* -3.643\*\*\*

(0.110) (0.321) (0.536)

regprc 0.704\*\*\* 1.995\*\*\* 3.227\*\*\*

(0.132) (0.383) (0.636)

faminc 0.001 0.001 0.002

(0.001) (0.002) (0.003)

hhsize.f2 0.102\* 0.291\* 0.463\*

(0.056) (0.160) (0.263)

hhsize.f3 0.121\* 0.351\*\* 0.558\*

(0.062) (0.178) (0.293)

hhsize.f4 0.099 0.276 0.450

(0.065) (0.184) (0.304)

hhsize.f5 0.151\* 0.431\* 0.712\*

(0.079) (0.229) (0.379)

hhsize.f6 0.293\*\* 0.950\*\* 1.553\*\*

(0.122) (0.399) (0.695)

hhsize.f7 0.179 0.516 0.875

(0.143) (0.432) (0.738)

hhsize.f8 -0.020 -0.040 -0.057

(0.180) (0.503) (0.815)

hhsize.f9 0.415 4.371 13.162

(0.464) (92.126) (535.411)

educ 0.024\*\*\* 0.072\*\*\* 0.118\*\*\*

(0.008) (0.025) (0.041)

age -0.0002 -0.0004 -0.001

(0.001) (0.004) (0.006)

Constant 0.390\*\* -0.353 -0.585

(0.166) (0.478) (0.790)

-------------------------------------------------------------------------

% correctly predicted LPM: 0.672727272727273

% correctly predicted Probit: 0.672727272727273

% correctly predicted Logit: 0.672727272727273

Observations 660 660 660

R2 0.118

Adjusted R2 0.100

Log Likelihood -395.763 -395.999

Akaike Inf. Crit. 819.526 819.998

Residual Std. Error 0.460 (df = 646)

F Statistic 6.634\*\*\* (df = 13; 646)

=========================================================================

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

> ame\_prob

Call:

probitmfx(formula = ecobuy ~ ecoprc + regprc + faminc + hhsize.f +

educ + age, data = data, atmean = FALSE)

Marginal Effects:

dF/dx Std. Err. z P>|z|

ecoprc -0.76518591 0.09701446 -7.8873 3.087e-15 \*\*\*

regprc 0.68027759 0.12257318 5.5500 2.857e-08 \*\*\*

faminc 0.00046835 0.00053693 0.8723 0.3830601

hhsize.f2 0.09690087 0.05153333 1.8804 0.0600600 .

hhsize.f3 0.11495741 0.05514450 2.0847 0.0371004 \*

hhsize.f4 0.09088561 0.05820736 1.5614 0.1184268

hhsize.f5 0.13695856 0.06600644 2.0749 0.0379933 \*

hhsize.f6 0.25914609 0.07521609 3.4454 0.0005703 \*\*\*

hhsize.f7 0.15890221 0.11574948 1.3728 0.1698110

hhsize.f8 -0.01358926 0.17325854 -0.0784 0.9374833

hhsize.f9 0.37593568 0.01798386 20.9041 < 2.2e-16 \*\*\*

educ 0.02451798 0.00824475 2.9738 0.0029417 \*\*

age -0.00015282 0.00124322 -0.1229 0.9021667

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

dF/dx is for discrete change for the following variables:

[1] "hhsize.f2" "hhsize.f3" "hhsize.f4" "hhsize.f5" "hhsize.f6" "hhsize.f7"

[7] "hhsize.f8" "hhsize.f9"

> pea\_prob

Call:

probitmfx(formula = ecobuy ~ ecoprc + regprc + faminc + hhsize.f +

educ + age, data = data, atmean = TRUE)

Marginal Effects:

dF/dx Std. Err. z P>|z|

ecoprc -0.83999086 0.12692460 -6.6180 3.640e-11 \*\*\*

regprc 0.74678185 0.14793395 5.0481 4.463e-07 \*\*\*

faminc 0.00051414 0.00059095 0.8700 0.3842869

hhsize.f2 0.10609800 0.05697154 1.8623 0.0625611 .

hhsize.f3 0.12530284 0.06057828 2.0684 0.0385982 \*

hhsize.f4 0.09938087 0.06391179 1.5550 0.1199534

hhsize.f5 0.14808928 0.07118476 2.0804 0.0374934 \*

hhsize.f6 0.26990053 0.07747779 3.4836 0.0004947 \*\*\*

hhsize.f7 0.16998799 0.12046575 1.4111 0.1582181

hhsize.f8 -0.01493544 0.19064118 -0.0783 0.9375551

hhsize.f9 0.36286832 0.01958192 18.5308 < 2.2e-16 \*\*\*

educ 0.02691487 0.00931072 2.8907 0.0038434 \*\*

age -0.00016776 0.00136485 -0.1229 0.9021732

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

dF/dx is for discrete change for the following variables:

[1] "hhsize.f2" "hhsize.f3" "hhsize.f4" "hhsize.f5" "hhsize.f6" "hhsize.f7"

[7] "hhsize.f8" "hhsize.f9"

> ame\_log

Call:

logitmfx(formula = ecobuy ~ ecoprc + regprc + faminc + hhsize.f +

educ + age, data = data, atmean = FALSE)

Marginal Effects:

dF/dx Std. Err. z P>|z|

ecoprc -0.75257787 0.12850856 -5.8562 4.734e-09 \*\*\*

regprc 0.66673040 0.14346774 4.6472 3.364e-06 \*\*\*

faminc 0.00051246 0.00054693 0.9370 0.348771

hhsize.f2 0.09352340 0.05115031 1.8284 0.067489 .

hhsize.f3 0.11038562 0.05468167 2.0187 0.043519 \*

hhsize.f4 0.08964789 0.05762416 1.5557 0.119771

hhsize.f5 0.13600974 0.06492274 2.0949 0.036176 \*

hhsize.f6 0.25146090 0.07583177 3.3160 0.000913 \*\*\*

hhsize.f7 0.16065994 0.11489840 1.3983 0.162030

hhsize.f8 -0.01189831 0.17029586 -0.0699 0.944298

hhsize.f9 0.37636280 0.01774099 21.2143 < 2.2e-16 \*\*\*

educ 0.02448012 0.00871959 2.8075 0.004993 \*\*

age -0.00019522 0.00123967 -0.1575 0.874871

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

dF/dx is for discrete change for the following variables:

[1] "hhsize.f2" "hhsize.f3" "hhsize.f4" "hhsize.f5" "hhsize.f6" "hhsize.f7"

[7] "hhsize.f8" "hhsize.f9"

> pea\_log

Call:

logitmfx(formula = ecobuy ~ ecoprc + regprc + faminc + hhsize.f +

educ + age, data = data, atmean = TRUE)

Marginal Effects:

dF/dx Std. Err. z P>|z|

ecoprc -0.83488276 0.23020402 -3.6267 0.0002871 \*\*\*

regprc 0.73964667 0.22575210 3.2764 0.0010515 \*\*

faminc 0.00056850 0.00061910 0.9183 0.3584786

hhsize.f2 0.10298244 0.06253898 1.6467 0.0996214 .

hhsize.f3 0.12060314 0.06873501 1.7546 0.0793260 .

hhsize.f4 0.09834209 0.06861382 1.4333 0.1517809

hhsize.f5 0.14633668 0.08318172 1.7592 0.0785366 .

hhsize.f6 0.25657040 0.12987188 1.9756 0.0482040 \*

hhsize.f7 0.16978436 0.13014491 1.3046 0.1920361

hhsize.f8 -0.01323028 0.18981189 -0.0697 0.9444308

hhsize.f9 0.36031779 0.02008400 17.9405 < 2.2e-16 \*\*\*

educ 0.02715736 0.01132513 2.3980 0.0164861 \*

age -0.00021657 0.00137609 -0.1574 0.8749471

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

dF/dx is for discrete change for the following variables:

[1] "hhsize.f2" "hhsize.f3" "hhsize.f4" "hhsize.f5" "hhsize.f6" "hhsize.f7"

[7] "hhsize.f8" "hhsize.f9"

(5)

Based on the AME and PEA for faminc in both logit and probit models and the LPM model presented above, we observe the estimates for faminc are always positive, which suggest that eco apples are normal goods. Yet because they are all statistically insignificant (p-value > 0.2), thus we cannot conclude this. We perform upper tail test using the LPM model:

> #t-statistic of faminc

> #from lpm we have t-value for faminc

> faminc\_stat = 0.976

> #critical value for one sided test

> alpha = 0.1

> qt(1-alpha,lpm$df.residual)

[1] 1.282863

> #since faminc\_stat < critical value, we fail to reject H0.

In relation to the regular apple, we observe that AME and PEA estimates for regprc as well as LPM estimate for regprc are positive and statistically significant. Thus, the eco apple is likely to be a substitution for regular apple.

(6)

> #LPM

> linearHypothesis(lpm,c('faminc=0',

+ 'hhsize.f2=0',

+ 'hhsize.f3=0',

+ 'hhsize.f4=0',

+ 'hhsize.f5=0',

+ 'hhsize.f6=0',

+ 'hhsize.f7=0',

+ 'hhsize.f8=0',

+ 'hhsize.f9=0'

+ ))

Linear hypothesis test

Hypothesis:

faminc = 0

hhsize.f2 = 0

hhsize.f3 = 0

hhsize.f4 = 0

hhsize.f5 = 0

hhsize.f6 = 0

hhsize.f7 = 0

hhsize.f8 = 0

hhsize.f9 = 0

Model 1: restricted model

Model 2: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age

Res.Df RSS Df Sum of Sq F Pr(>F)

1 655 138.92

2 646 136.58 9 2.3419 1.2308 0.2728

|  |
| --- |
| > #Probit  > probitR <- glm(ecobuy~ecoprc+regprc+educ+age,family=binomial(link='probit'),data=data)  > waldtest(probit,probitR,test='Chisq')  Wald test  Model 1: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age  Model 2: ecobuy ~ ecoprc + regprc + educ + age  Res.Df Df Chisq Pr(>Chisq)  1 646  2 655 -9 10.674 0.2987  > lrtest(probit,probitR)  Likelihood ratio test  Model 1: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age  Model 2: ecobuy ~ ecoprc + regprc + educ + age  #Df LogLik Df Chisq Pr(>Chisq)  1 14 -395.76  2 5 -401.55 -9 11.58 0.238  > anova(probit,probitR,test='Rao')  Analysis of Deviance Table  Model 1: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age  Model 2: ecobuy ~ ecoprc + regprc + educ + age  Resid. Df Resid. Dev Df Deviance Rao Pr(>Chi)  1 646 791.53  2 655 803.11 -9 -11.58 -11.056 0.2719 |
|  |
| |  | | --- | | > #Logit  > logitR <- glm(ecobuy~ecoprc+regprc+educ+age,family=binomial(link='logit'),data=data)  > waldtest(logit,logitR,test='Chisq')  Wald test  Model 1: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age  Model 2: ecobuy ~ ecoprc + regprc + educ + age  Res.Df Df Chisq Pr(>Chisq)  1 646  2 655 -9 10.301 0.3266  > lrtest(logit,logitR)  Likelihood ratio test  Model 1: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age  Model 2: ecobuy ~ ecoprc + regprc + educ + age  #Df LogLik Df Chisq Pr(>Chisq)  1 14 -396.00  2 5 -401.74 -9 11.489 0.2437  > anova(logit,logitR,test='Rao')  Analysis of Deviance Table  Model 1: ecobuy ~ ecoprc + regprc + faminc + hhsize.f + educ + age  Model 2: ecobuy ~ ecoprc + regprc + educ + age  Resid. Df Resid. Dev Df Deviance Rao Pr(>Chi)  1 646 792.00  2 655 803.49 -9 -11.489 -11.039 0.2731 | |

* All the tests agree that faminc and hhsize.f variables are jointly statistically insignificant at 5% level.