Discrete Choice Models

Kristopher C. Toll

February 8, 2018

## Descriptive Stats

## [1] "Sample Size" "2694"

## Var1 Var2 Freq  
## 1 Renters Cache 88  
## 2 Owners Cache 142  
## 3 Renters WFRC-MAG 538  
## 4 Owners WFRC-MAG 1490  
## 5 Renters Dixie 15  
## 6 Owners Dixie 133  
## 7 Renters Utah Other 35  
## 8 Owners Utah Other 253

## Var1 Freq  
## 1 -1 225  
## 2 1 76  
## 3 2 224  
## 4 3 230  
## 5 4 401  
## 6 5 593  
## 7 6 444  
## 8 7 346  
## 9 8 88  
## 10 9 31  
## 11 10 36

## Models

### Full Model

##   
## Call:  
## glm(formula = choice ~ NominalPrice + commute + destinations +   
## homes + streets + transit + ParkingDriveway + ParkingOffStreet +   
## alt, data = LogitData)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.91876 -0.41037 -0.03403 0.42698 0.85070   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 0.42008489835473 0.00822887232898 51.050  
## NominalPrice -0.00000000003011 0.00000000006588 -0.457  
## commute -0.09066704199367 0.00404825586560 -22.397  
## destinations -0.10226766308032 0.00403384756585 -25.352  
## homes 0.13689625249530 0.00453494499059 30.187  
## streets 0.07897149251385 0.00405010248170 19.499  
## transit -0.06813930419282 0.00404093847881 -16.862  
## ParkingDriveway 0.25073557976939 0.00453280522655 55.316  
## ParkingOffStreet -0.02574915262749 0.00592725021633 -4.344  
## alt 0.01603454632277 0.00401580683133 3.993  
## Pr(>|t|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## NominalPrice 0.648   
## commute < 0.0000000000000002 \*\*\*  
## destinations < 0.0000000000000002 \*\*\*  
## homes < 0.0000000000000002 \*\*\*  
## streets < 0.0000000000000002 \*\*\*  
## transit < 0.0000000000000002 \*\*\*  
## ParkingDriveway < 0.0000000000000002 \*\*\*  
## ParkingOffStreet 0.0000140 \*\*\*  
## alt 0.0000654 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.2170472)  
##   
## Null deviance: 13470 on 53879 degrees of freedom  
## Residual deviance: 11692 on 53870 degrees of freedom  
## AIC: 70608  
##   
## Number of Fisher Scoring iterations: 2

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: FullModel$y, FullModel$fitted.values  
## X-squared = 164.75, df = 8, p-value < 0.00000000000000022

### Renters Model

##   
## Call:  
## glm(formula = choice ~ NominalPrice + commute + destinations +   
## homes + streets + transit + ParkingDriveway + ParkingOffStreet +   
## alt, data = subset(LogitData, rent\_own == "1"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.03663 -0.44177 -0.06233 0.44516 0.88022   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.49094735 0.01660862 29.560 <0.0000000000000002 \*\*\*  
## NominalPrice -0.00052921 0.00003777 -14.012 <0.0000000000000002 \*\*\*  
## commute -0.12676190 0.00822622 -15.409 <0.0000000000000002 \*\*\*  
## destinations -0.12141576 0.00817468 -14.853 <0.0000000000000002 \*\*\*  
## homes 0.10378015 0.00918533 11.298 <0.0000000000000002 \*\*\*  
## streets 0.08667322 0.00821763 10.547 <0.0000000000000002 \*\*\*  
## transit -0.07827002 0.00818435 -9.563 <0.0000000000000002 \*\*\*  
## ParkingDriveway 0.16717441 0.00917869 18.213 <0.0000000000000002 \*\*\*  
## ParkingOffStreet -0.01250742 0.01194621 -1.047 0.2951   
## alt 0.01870657 0.00813336 2.300 0.0215 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.2231853)  
##   
## Null deviance: 3380.0 on 13519 degrees of freedom  
## Residual deviance: 3015.2 on 13510 degrees of freedom  
## AIC: 18103  
##   
## Number of Fisher Scoring iterations: 2

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: Renters$y, Renters$fitted.values  
## X-squared = 14.863, df = 8, p-value = 0.06187

### Owner's Model

##   
## Call:  
## glm(formula = choice ~ NominalPrice + commute + destinations +   
## homes + streets + transit + ParkingDriveway + ParkingOffStreet +   
## alt, data = subset(LogitData, rent\_own == "2"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.92704 -0.39392 -0.03466 0.40109 0.85773   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 0.40019262125962 0.00942812987391 42.447  
## NominalPrice -0.00000000002997 0.00000000006523 -0.460  
## commute -0.08130483304942 0.00463136857542 -17.555  
## destinations -0.09471164517478 0.00461457580568 -20.524  
## homes 0.14701265671341 0.00518748679688 28.340  
## streets 0.07453484980451 0.00463128013547 16.094  
## transit -0.06366272658951 0.00462257902183 -13.772  
## ParkingDriveway 0.27747783281579 0.00518599597827 53.505  
## ParkingOffStreet -0.03214644659794 0.00679202931308 -4.733  
## alt 0.01390763649796 0.00459415617770 3.027  
## Pr(>|t|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## NominalPrice 0.64586   
## commute < 0.0000000000000002 \*\*\*  
## destinations < 0.0000000000000002 \*\*\*  
## homes < 0.0000000000000002 \*\*\*  
## streets < 0.0000000000000002 \*\*\*  
## transit < 0.0000000000000002 \*\*\*  
## ParkingDriveway < 0.0000000000000002 \*\*\*  
## ParkingOffStreet 0.00000222 \*\*\*  
## alt 0.00247 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.2127481)  
##   
## Null deviance: 10090.0 on 40359 degrees of freedom  
## Residual deviance: 8584.4 on 40350 degrees of freedom  
## AIC: 52086  
##   
## Number of Fisher Scoring iterations: 2

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: Owners$y, Owners$fitted.values  
## X-squared = 73.256, df = 8, p-value = 0.000000000001102

### Low Income

##   
## Call:  
## glm(formula = choice ~ NominalPrice + commute + destinations +   
## homes + streets + transit + ParkingDriveway + ParkingOffStreet +   
## alt, data = subset(LogitData, income == "1" | income == "2" |   
## income == "3"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.8761 -0.4353 -0.0122 0.4328 0.8434   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 0.47665310145 0.01884686109 25.291  
## NominalPrice 0.00000003080 0.00000008044 0.383  
## commute -0.10109563771 0.00932464309 -10.842  
## destinations -0.10518118048 0.00929011906 -11.322  
## homes 0.09826977424 0.01042393902 9.427  
## streets 0.08568655345 0.00934260831 9.172  
## transit -0.08595913178 0.00930773489 -9.235  
## ParkingDriveway 0.17813229419 0.01040565508 17.119  
## ParkingOffStreet -0.04389894548 0.01363652790 -3.219  
## alt 0.01792062224 0.00924856421 1.938  
## Pr(>|t|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## NominalPrice 0.70179   
## commute < 0.0000000000000002 \*\*\*  
## destinations < 0.0000000000000002 \*\*\*  
## homes < 0.0000000000000002 \*\*\*  
## streets < 0.0000000000000002 \*\*\*  
## transit < 0.0000000000000002 \*\*\*  
## ParkingDriveway < 0.0000000000000002 \*\*\*  
## ParkingOffStreet 0.00129 \*\*   
## alt 0.05269 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.2263844)  
##   
## Null deviance: 2650.0 on 10599 degrees of freedom  
## Residual deviance: 2397.4 on 10590 degrees of freedom  
## AIC: 14347  
##   
## Number of Fisher Scoring iterations: 2

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: LowI$y, LowI$fitted.values  
## X-squared = 17.876, df = 8, p-value = 0.02217

### Middle Income

##   
## Call:  
## glm(formula = choice ~ NominalPrice + commute + destinations +   
## homes + streets + transit + ParkingDriveway + ParkingOffStreet +   
## alt, data = subset(LogitData, income == "4" | income == "5"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.91888 -0.40234 -0.00181 0.41873 0.84965   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 0.432711529018 0.013537707626 31.963  
## NominalPrice -0.000000005387 0.000000003480 -1.548  
## commute -0.091833927524 0.006657542508 -13.794  
## destinations -0.104132025581 0.006631980039 -15.701  
## homes 0.140442147776 0.007462046457 18.821  
## streets 0.074015096286 0.006653625445 11.124  
## transit -0.064932570672 0.006640618128 -9.778  
## ParkingDriveway 0.255609345301 0.007450011729 34.310  
## ParkingOffStreet -0.026604620769 0.009743111974 -2.731  
## alt 0.007943459507 0.006603833919 1.203  
## Pr(>|t|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## NominalPrice 0.12162   
## commute < 0.0000000000000002 \*\*\*  
## destinations < 0.0000000000000002 \*\*\*  
## homes < 0.0000000000000002 \*\*\*  
## streets < 0.0000000000000002 \*\*\*  
## transit < 0.0000000000000002 \*\*\*  
## ParkingDriveway < 0.0000000000000002 \*\*\*  
## ParkingOffStreet 0.00633 \*\*   
## alt 0.22905   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.2165044)  
##   
## Null deviance: 4970.0 on 19879 degrees of freedom  
## Residual deviance: 4301.9 on 19870 degrees of freedom  
## AIC: 26010  
##   
## Number of Fisher Scoring iterations: 2

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: MiddleI$y, MiddleI$fitted.values  
## X-squared = 46.185, df = 8, p-value = 0.0000002192

### High Income

##   
## Call:  
## glm(formula = choice ~ NominalPrice + commute + destinations +   
## homes + streets + transit + ParkingDriveway + ParkingOffStreet +   
## alt, data = subset(LogitData, income == "6" | income == "7" |   
## income == "8" | income == "9" | income == "10"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.95751 -0.39542 -0.07055 0.40678 0.86939   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 0.38132472336 0.01379027073 27.652  
## NominalPrice -0.00000020344 0.00000007731 -2.631  
## commute -0.09491449077 0.00676281574 -14.035  
## destinations -0.09800926335 0.00672637761 -14.571  
## homes 0.14712396688 0.00755311293 19.479  
## streets 0.08431774958 0.00675240570 12.487  
## transit -0.06300481399 0.00673918774 -9.349  
## ParkingDriveway 0.28168892308 0.00756043971 37.258  
## ParkingOffStreet -0.01989985242 0.00990708694 -2.009  
## alt 0.02511682382 0.00669130308 3.754  
## Pr(>|t|)   
## (Intercept) < 0.0000000000000002 \*\*\*  
## NominalPrice 0.008510 \*\*   
## commute < 0.0000000000000002 \*\*\*  
## destinations < 0.0000000000000002 \*\*\*  
## homes < 0.0000000000000002 \*\*\*  
## streets < 0.0000000000000002 \*\*\*  
## transit < 0.0000000000000002 \*\*\*  
## ParkingDriveway < 0.0000000000000002 \*\*\*  
## ParkingOffStreet 0.044589 \*   
## alt 0.000175 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 0.2111883)  
##   
## Null deviance: 4725.0 on 18899 degrees of freedom  
## Residual deviance: 3989.3 on 18890 degrees of freedom  
## AIC: 24258  
##   
## Number of Fisher Scoring iterations: 2

##   
## Hosmer and Lemeshow goodness of fit (GOF) test  
##   
## data: HighI$y, HighI$fitted.values  
## X-squared = 47.219, df = 8, p-value = 0.0000001393