# Illustration 6

Aims:

- Study the determinants of the choice between 2 tomato ketchup brands: Heinz and Hunts
- In particular, investigate whether the promotional activities developed by both brands impact on the probability of consumers choosing one instead of the other

Sample:

- 2798 purchases made during a period of about two years by a panel of 300 households in Springfield, Missouri
- Purchase data were collected automatically using an optical scanner; the shelf price of the other brand was also stored
- Record if at the time of purchase the brands had ongoing promotional activities

# Illustration 6

Promotional activities:

- Display only
- Newspaper feature only
- Both

Model:

$$Pr(Heinz = 1| \dots)$$

$$= G\left[\beta_0 + \beta_1 Dhei + \beta_2 Fhei + \beta_3 DFhei + \beta_4 Dhun\right.$$

$$\left. + \beta_5 Fhun + \beta_6 DFhun + \beta_7 log\left(\frac{Phei}{Phun}\right)\right]$$

Details:

- Franses and Paap (2001), ch. 4.4

# Illustration 6 – Question 1

```
. describe

------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
------------------------------------------------------------------------------
Family          int     %8.0g                 Family id
Heinz           byte    %8.0g                 =1 Heinz chosen
Hunts           byte    %8.0g                 =1 Hunts chosen
Phei            float   %9.0g                 Heinz price (US$/oz)
Phun            float   %9.0g                 Hunts price (US$/oz)
Dhei            byte    %8.0g                 =1 if Heinz was on display
                                                but not featured
Dhun            byte    %8.0g                 =1 if Hunts was on display
                                                but not featured
Fhei            byte    %8.0g                 =1 if Heinz was featured but
                                                not on display
Fhun            byte    %8.0g                 =1 if Hunts was featured but
                                                not on display
DFhei           byte    %8.0g                 =1 if Heinz was on display
                                                and featured
DFhun           byte    %8.0g                 =1 if Hunts was on display
                                                and featured
------------------------------------------------------------------------------
```

# Illustration 6 – Question 2

```
. summarize Heinz Hunts Dhei Dhun Fhei Fhun DFhei DFhun Phei Phun
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Heinz | 2798 | .8902788 | .3125978 | 0 | 1 |
| Hunts | 2798 | .1097212 | .3125978 | 0 | 1 |
| Dhei | 2798 | .159757 | .366446 | 0 | 1 |
| Dhun | 2798 | .0353824 | .1847774 | 0 | 1 |
| Fhei | 2798 | .124732 | .3304738 | 0 | 1 |
| Fhun | 2798 | .0364546 | .1874519 | 0 | 1 |
| DFhei | 2798 | .0375268 | .1900828 | 0 | 1 |
| DFhun | 2798 | .0092924 | .0959651 | 0 | 1 |
| Phei | 2798 | .0348276 | .0089737 | .001 | .061 |
| Phun | 2798 | .0335547 | .0053069 | .003 | .087 |

# Illustration 6 – Question 3

```
. gen lpp=ln(Phei/Phun)
. logit Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp

Logistic regression                                     Number of obs   =       2798
                                                        LR chi2(7)      =     584.96
                                                        Prob > chi2     =     0.0000
Log likelihood = -675.43973                             Pseudo R2       =     0.3022

------------------------------------------------------------------------------
       Heinz |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Dhei |   .5474406   .2427129     2.26   0.024      .071732    1.023149
        Dhun |  -.6207141   .2449583    -2.53   0.011    -1.100824   -.1406047
        Fhei |   .5785839   .3165889     1.83   0.068     -.041919    1.199087
        Fhun |  -.9971683   .3466031    -2.88   0.004    -1.676498   -.3178388
       DFhei |   .4452738   .4431288     1.00   0.315    -.4232427     1.31379
       DFhun |  -1.940127   .4802308    -4.04   0.000    -2.881362   -.9988919
         lpp |  -6.110695   .3858905   -15.84   0.000    -6.867026   -5.354363
       _cons |    3.27461   .1423418    23.01   0.000     2.995625    3.553595
------------------------------------------------------------------------------
. estimates store logit
```

# Illustration 6 – Question 3 (cont.)

```
. probit Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp

Probit regression                                 Number of obs   =       2798
                                                  LR chi2(7)      =     591.39
                                                  Prob > chi2     =     0.0000
Log likelihood = -672.22516                       Pseudo R2       =     0.3055

------------------------------------------------------------------------------
       Heinz |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Dhei |   .2768603   .1226386     2.26   0.024     .0364931    .5172275
        Dhun |  -.3599038   .1455494    -2.47   0.013    -.6451753   -.0746322
        Fhei |   .2396884   .1549172     1.55   0.122    -.0639437    .5433206
        Fhun |  -.5526825   .1904251    -2.90   0.004    -.9259089   -.1794561
       DFhei |   .2486314   .2282668     1.09   0.276    -.1987634    .6960262
       DFhun |  -1.068493      .2749    -3.89   0.000    -1.607287   -.5296985
         lpp |  -3.348536   .2094459   -15.99   0.000    -3.759042   -2.938029
       _cons |   1.840782   .0723702    25.44   0.000     1.698939    1.982625
------------------------------------------------------------------------------

. estimates store probit
```

# Illustration 6 – Question 3 (cont.)

```
. cloglog Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp

Complementary log-log regression                Number of obs     =       2798
                                                Zero outcomes     =        307
                                                Nonzero outcomes  =       2491

                                                LR chi2(7)        =     589.99
Log likelihood = -672.92517                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       Heinz |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Dhei |   .2088229   .0936429     2.23   0.026     .0252863    .3923596
        Dhun |  -.3518195   .1501518    -2.34   0.019    -.6461116   -.0575275
        Fhei |   .1227095   .1145124     1.07   0.284    -.1017307    .3471497
        Fhun |  -.4820495    .164347    -2.93   0.003    -.8041638   -.1599353
       DFhei |    .209065   .1765308     1.18   0.236     -.136929    .5550589
       DFhun |  -1.077127   .3208661    -3.36   0.001    -1.706013    -.448241
         lpp |  -2.802396   .1883769   -14.88   0.000    -3.171608   -2.433184
       _cons |   1.270849   .0566031    22.45   0.000     1.159909    1.381789
------------------------------------------------------------------------------
. estimates store cloglog
```

# Illustration 6 – Question 3 (cont.)

```
. estimates table logit probit cloglog, b star(0.1 0.05 0.01)

-------------------------------------------------------------
    Variable |     logit          probit          cloglog
-------------+-----------------------------------------------
        Dhei |  .54744056**     .27686031**     .20882294**
        Dhun | -.62071414**    -.35990375**    -.35181954**
        Fhei |  .57858393*      .23968844       .12270948
        Fhun | -.99716835***   -.55268247***   -.48204954***
       DFhei |   .4452738       .24863139       .20906496
       DFhun | -1.9401271***   -1.0684927***   -1.0771271***
         lpp | -6.1106949***   -3.3485359***   -2.8023956***
       _cons |  3.2746098***    1.8407819***    1.2708491***
-------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01
```

# Illustration 6 – Question 3 (cont.)

Main conclusions:

- There is no clear evidence on whether the promotional activities undertaken by Heinz increase significantly the probability of consumers purchasing its ketchup
- The promotional activities made by Hunts decrease significantly the probability of consumers purchasing Heinz ketchup ⟶ clearly, because Hunts is a smaller and less known brand, it benefits a lot from promotional activities, especially when display and feature activities are held at the same time
- Increasing the relative price of Heinz ketchup decreases the probability of consumers purchasing its ketchup

# Illustration 6 – Question 4

```
. estimates restore logit
. predict XBl, xb
. gen XBl2=XBl^2
. quietly logit Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp XBl2
. test XBl2
 ( 1)  [Heinz]XBl2 = 0

           chi2(  1) =    9.35
         Prob > chi2 =    0.0022
```

The null hypothesis of a well-specified functional form is rejected

```
. estimates restore probit
. predict XBp, xb
. gen XBp2=XBp^2
. quietly probit Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp XBp2
. test XBp2
 ( 1)  [Heinz]XBp2 = 0

           chi2(  1) =    0.73
         Prob > chi2 =    0.3933
```

The null hypothesis of a well-specified functional form cannot be rejected

# Illustration 6 – Question 4 (cont.)

```
. estimates restore cloglog
. predict XBcl, xb
. gen XBcl2=XBcl^2
. quietly cloglog Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp XBcl2
. test XBcl2
 ( 1)  [Heinz]XBcl2 = 0

           chi2(  1) =    9.63
         Prob > chi2 =    0.0019
```

The null hypothesis of a well-specified functional form is rejected

# Illustration 6 – Question 5.1

```
. quietly probit Heinz Dhei Dhun Fhei Fhun DFhei DFhun lpp XBp2

. estimates store probitR

. lrtest probit probitR

Likelihood-ratio test                                 LR chi2(1)  =      0.73
(Assumption: probit nested in probitR)                Prob > chi2 =    0.3922
```

The null hypothesis of a well-specified functional form is not rejected

# Illustration 6 – Question 5.2

```
. estimates restore probit
(results probit are active now)

. estat classification

Probit model for Heinz

              -------- True --------
Classified |         D             ~D  |      Total
-----------+--------------------------+-----------
     +     |      2462            232  |       2694
     -     |        29             75  |        104
-----------+--------------------------+-----------
   Total   |      2491            307  |       2798

Classified + if predicted Pr(D) >= .5
True D defined as Heinz != 0
```

(continues in the next slide)

# Illustration 6 – Question 5.2 (cont.)

| | | |
|---|---|---|
| Sensitivity | Pr( +\| D) | 98.84% |
| Specificity | Pr( -\|~D) | 24.43% |
| Positive predictive value | Pr( D\| +) | 91.39% |
| Negative predictive value | Pr(~D\| -) | 72.12% |
| False + rate for true ~D | Pr( +\|~D) | 75.57% |
| False - rate for true D | Pr( -\| D) | 1.16% |
| False + rate for classified + | Pr(~D\| +) | 8.61% |
| False - rate for classified - | Pr( D\| -) | 27.88% |
| Correctly classified | | 90.67% |

→ % 1's correctly predicted

→ % 0's correctly predicted

→ % correct predictions

# Illustration 6 – Question 6.1

```
. summarize Phei Phun

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        Phei |      2798    .0348276    .0089737       .001       .061
        Phun |      2798    .0335547    .0053069       .003       .087

. scalar lppm=log(0.0348276/0.0335547)

. display normal(_b[_cons]+_b[lpp]*lppm)
.9569286

. display normal(_b[_cons]+_b[DFhei]+_b[lpp]*lppm)
.97527763

. display normal(_b[_cons]+_b[DFhun]+_b[lpp]*lppm)
.74138217

. display normal(_b[_cons]+_b[DFhei]+_b[DFhun]+_b[lpp]*lppm)
.81493872
```

# Illustration 6 – Question 6.1 (cont.)

| | I | II | III | IV |
|---|---|---|---|---|
| Dhei | 0 | 0 | 0 | 0 |
| Fhei | 0 | 0 | 0 | 0 |
| DFhei | 0 | 1 | 0 | 1 |
| Dhun | 0 | 0 | 0 | 0 |
| Fhun | 0 | 0 | 0 | 0 |
| DFhun | 0 | 0 | 1 | 1 |
| Phei | 3.48 | 3.48 | 3.48 | 3.48 |
| Phun | 3.36 | 3.36 | 3.36 | 3.36 |
| $Pr(Heinz = 1 \mid \ldots)$ | 95.69 | 97.53 | 74.14 | 81.49 |

# Illustration 6 – Question 6.2

```
. margins, dydx(_all)

Average marginal effects                        Number of obs   =       2798
Model VCE    : OIM

Expression   : Pr(Heinz), predict()
dy/dx w.r.t. : Dhei Dhun Fhei Fhun DFhei DFhun lpp

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Dhei |   .0363993   .0161128     2.26   0.024     .0048189    .0679798
        Dhun |  -.0473172   .0190541    -2.48   0.013    -.0846626   -.0099719
        Fhei |   .0315123    .020376     1.55   0.122     -.008424    .0714486
        Fhun |  -.0726622    .024952    -2.91   0.004    -.1215672   -.0237572
       DFhei |    .032688   .0299977     1.09   0.276    -.0261065    .0914825
       DFhun |  -.1404768   .0358536    -3.92   0.000    -.2107485    -.070205
         lpp |  -.4402384    .024884   -17.69   0.000      -.48901   -.3914667
------------------------------------------------------------------------------
```

# Illustration 6 – Question 6.3

```
. margins, dydx(DFhun) at(Dhei=0 Fhei=0 DFhei=0 Dhun=0 Fhun=0 DFhun=0 lpp=0)

Conditional marginal effects                       Number of obs   =       2798
Model VCE    : OIM

Expression   : Pr(Heinz), predict()
dy/dx w.r.t. : DFhun
at           : Dhei            =           0
               Dhun            =           0
               Fhei            =           0
               Fhun            =           0
               DFhei           =           0
               DFhun           =           0
               lpp             =           0

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       DFhun |  -.0783219   .0220713    -3.55   0.000    -.1215808   -.0350629
------------------------------------------------------------------------------
```

# Illustration 6 – Question 6.4

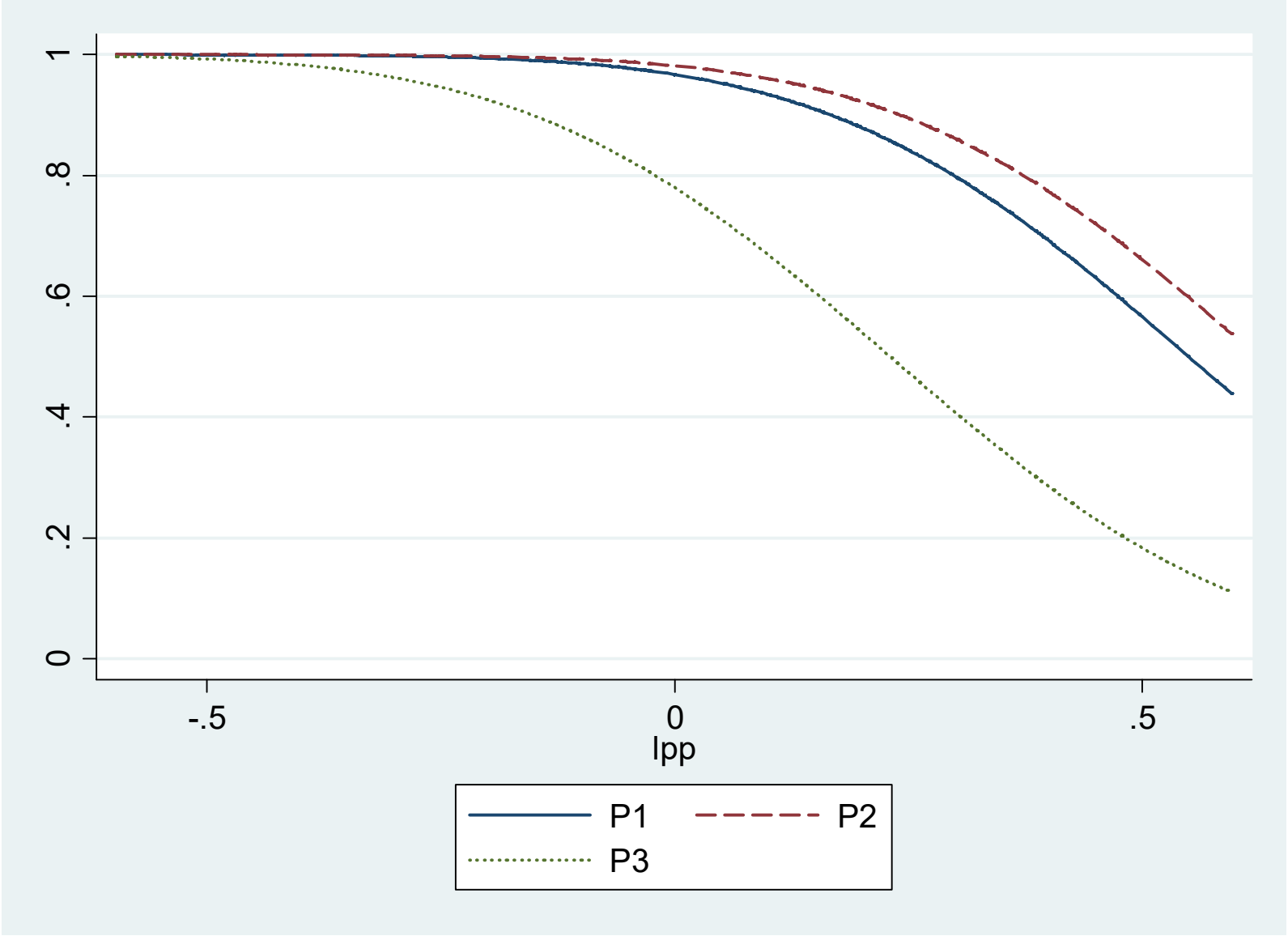
```
. gen P1=normal(_b[_cons]+_b[lpp]*lpp)

. gen P2=normal(_b[_cons]+_b[DFhei]+_b[lpp]*lpp)

. gen P3=normal(_b[_cons]+_b[DFhun]+_b[lpp]*lpp)

. line P1 P2 P3 lpp if lpp > -0.6 & lpp < 0.6, lpattern(solid dash dot)
```

# Illustration 6 – Question 6.4 (cont.)



# Illustration 6 – Question 6.4 (cont.)

Main conclusions:

- The higher the price of Heinz relative to Hunts, the less the probability of Heinz being purchased (and vice-versa)
- Heinz promotional activities do not allow this brand to substantially increase its price or market share
- Hunts promotional activities are more effective, with two possible effects:
  - If Hunts does not change its price, the probability of consumers purchasing its ketchup increases substantially
  - If Hunts opts for keeping its market share, it is able to sell the same quantity but at a higher price

# Illustration 7 - Question 1

Description of share2:

```
. summarize share2
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      share2 |      2724    .0122429     .024919          0   .1927609

. count if share2==0
 1688

. display 1688/2724
.61967695
```

# Illustration 7 – Question 2

```
. tobit share2 age nadults nkids nkids2 lnx agelnx nadlnx, ll(0)
Tobit regression                                  Number of obs   =       2724
                                                  LR chi2(7)      =     170.18
                                                  Prob > chi2     =     0.0000
Log likelihood =  758.70053                       Pseudo R2       =    -0.1263
------------------------------------------------------------------------------
      share2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.1258528   .0241782    -5.21   0.000    -.1732624   -.0784432
     nadults |     .01537   .0380475     0.40   0.686    -.0592349     .089975
       nkids |   .0042697   .0013247     3.22   0.001     .0016723    .0068671
      nkids2 |  -.0099719   .0054713    -1.82   0.068    -.0207002    .0007565
         lnx |  -.0444314   .0068893    -6.45   0.000    -.0579402   -.0309225
      agelnx |   .0088221   .0017832     4.95   0.000     .0053256    .0123187
      nadlnx |  -.0006007   .0027501    -0.22   0.827    -.0059933    .0047918
       _cons |   .5899797   .0934268     6.31   0.000     .4067849    .7731745
-------------+----------------------------------------------------------------
      /sigma |   .0479951   .0011832                      .0456751    .0503151
------------------------------------------------------------------------------
  Obs. summary:       1688  left-censored observations at share2<=0
                      1036     uncensored observations
                         0 right-censored observations
```

# Illustration 7 – Question 3 (additional topic)

```
. probit d2 age nadults nkids nkids2 lnx agelnx nadlnx
(…)
Probit regression                                 Number of obs   =       2724
                                                  LR chi2(7)      =     101.68
                                                  Prob > chi2     =     0.0000
Log likelihood = -1758.5011                       Pseudo R2       =     0.0281
------------------------------------------------------------------------------
          d2 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -2.479476   .5573451    -4.45   0.000    -3.571853     -1.3871
     nadults |   .6532982   .8664414     0.75   0.451    -1.044896    2.351492
       nkids |   .0878242   .0305731     2.87   0.004      .027902    .1477463
      nkids2 |  -.2184835   .1224591    -1.78   0.074    -.4584988    .0215319
         lnx |  -.6162177   .1615998    -3.81   0.000    -.9329476   -.2994879
      agelnx |   .1732215   .0410448     4.22   0.000     .0927751    .2536678
      nadlnx |  -.0365886   .0625613    -0.58   0.559    -.1592064    .0860293
       _cons |   8.077725   2.197003     3.68   0.000     3.771679    12.38377
------------------------------------------------------------------------------
```

# Illustration 7 – Question 3 (additional topic)

```
. reg share2 age nadults nkids nkids2 lnx agelnx nadlnx if share2>0
      Source |       SS       df       MS              Number of obs =    1036
-------------+------------------------------           F(  7,  1028) =   26.73
       Model |   .15794176     7  .022563109            Prob > F      =  0.0000
    Residual |  .867670386  1028  .000844037            R-squared     =  0.1540
-------------+------------------------------           Adj R-squared =  0.1482
       Total |  1.02561215  1035   .00099093            Root MSE      =  .02905
------------------------------------------------------------------------------
      share2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0314662   .0205632    -1.53   0.126    -.0718168    .0088844
     nadults |  -.0130266   .0324149    -0.40   0.688    -.0766334    .0505803
       nkids |   .0012847   .0010541     1.22   0.223    -.0007837    .0033531
      nkids2 |  -.0034369    .004556    -0.75   0.451    -.0123771    .0055033
         lnx |  -.0335767   .0054672    -6.14   0.000    -.0443049   -.0228484
      agelnx |   .0022097    .001516     1.46   0.145     -.000765    .0051844
      nadlnx |   .0011125    .002345     0.47   0.635     -.003489    .0057141
       _cons |   .4896596   .0740595     6.61   0.000     .3443345    .6349847
------------------------------------------------------------------------------
```