MNLM for Nominal Outcomes

Objectives

Introduce the MNLM as an extension of the BLM

Derive the model as a nonlinear probability model

Illustrate the difficulties in interpretation due to the large number of parameters and comparisons

Introduce graphical methods that make interpretation simpler

(Rethinking) the BLM

The BLM describes the relative probability of one outcome compared to a base or reference outcome

For example, being in the labor force compared to being out of the labor force

Think of the BLM as having two sets of $\beta's$

One set is associated with y=1 compared to y=0

The other set is associated with y=0 compared to y=1

Only J-1 sets are estimated

Binary Logit Model (new notation)

$$\ln\left[\frac{\Pr(y=1\mid\mathbf{x})}{\Pr(y=0\mid\mathbf{x})}\right]=\mathbf{x}\boldsymbol{\beta}\Rightarrow$$

$$\ln\left[\frac{\Pr(y=A\mid\mathbf{x})}{\Pr(y=B\mid\mathbf{x})}\right]=\mathbf{x}\boldsymbol{\beta}_{A|B}$$

For a model with three independent variables

$$\ln\left[\frac{\Pr(y=A\mid\mathbf{x})}{\Pr(y=B\mid\mathbf{x})}\right]=\beta_{0,A|B}+\beta_{1,A|B}x_1+\beta_{2,A|B}x_2+\beta_{3,A|B}x_3$$

The probability that y=1 (or A)

$$\Pr(y=1\,|\,\mathbf{x})=\frac{\exp(\mathbf{x}\boldsymbol{\beta})}{1+\exp(\mathbf{x}\boldsymbol{\beta})}\Rightarrow\Pr(y=A\,|\,\mathbf{x})=\frac{\exp(\mathbf{x}\boldsymbol{\beta}_{A|B})}{1+\exp(\mathbf{x}\boldsymbol{\beta}_{A|B})}$$

The probability that y=0 (or B)

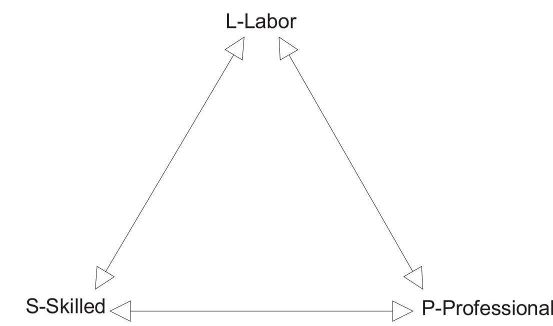
$$\Pr(y=0\,|\,\mathbf{x})=\frac{1}{1+\exp(\mathbf{x}\boldsymbol{\beta})}\Rightarrow\Pr(y=B\,|\,\mathbf{x})=\frac{1}{1+\exp(\mathbf{x}\boldsymbol{\beta}_{A|B})}$$

Question (for you)

Which is the base or reference category?

Three outcome categories

Consider y with categories L, S, and P



Think of this as three BLMs

The effect of *Ed* on the odds of L versus S:

$$\ln\left[\frac{\Pr(L\,|\,Ed)}{\Pr(S\,|\,Ed)}\right]=\beta_{0,L|S}+\beta_{1,L|S}Ed$$

For S versus P:

$$\ln\left[\frac{\Pr(S\,|\,Educ)}{\Pr(P\,|\,Educ)}\right]=\beta_{0,S|P}+\beta_{1,S|P}Ed$$

Question (for you)

What about the remaining comparison?

Redundancy

Using the property $\ln(a/b)=\ln(a)-\ln(b)$**:**

$$\ln\left[\frac{\Pr(L\mid Ed)}{\Pr(P\mid Ed)}\right]=\left[\ln\Pr(L\mid Ed)-\ln\Pr(P\mid Ed)\right]+[0]$$
$$=\left[\ln\Pr(L\mid Ed)-\ln\Pr(P\mid Ed)\right]+\left[\ln\Pr(S\mid Ed)-\ln\Pr(S\mid Ed)\right]$$
$$=\left[\ln\Pr(L\mid Ed)-\ln\Pr(S\mid Ed)\right]+\left[\ln\Pr(S\mid Ed)-\ln\Pr(P\mid Ed)\right]$$
$$=\ln\left[\frac{\Pr(L\mid Ed)}{\Pr(S\mid Ed)}\right]+\ln\left[\frac{\Pr(S\mid Ed)}{\Pr(P\mid Ed)}\right]$$

Thus, if we add equations 1 and 2, we get 3:

$$\ln\left[\frac{\Pr(L\mid Ed)}{\Pr(S\mid Ed)}\right]+\ln\left[\frac{\Pr(S\mid Ed)}{\Pr(P\mid Ed)}\right]=\ln\left[\frac{\Pr(L\mid Ed)}{\Pr(P\mid Ed)}\right]$$

Logical Relationship

You can find a coefficient for any comparison from a pair of other coefficients:

$$\beta_{L|P}=\beta_{L|S}+\beta_{S|P}$$
$$\beta_{L|S}=\beta_{L|P}-\beta_{S|P}$$
$$\beta_{S|P}=\beta_{L|P}-\beta_{L|S}$$

But...

Why won't the results from separate BLMs match those from MNLM exactly?

A Minimal Set of Coefficients

For J outcomes, J-1 comparisons

Different software might compute different minimal sets

Question (for you)

Is this a problem?

The MNLM as a Probability Model

Let y have J nominal outcomes numbered 1 through J

$\Pr(y = m \mid \mathbf{x})$ is a function of $\mathbf{x}\beta_{m|J}$

Take the exponential to ensure that the probabilities are non-negative

Divide by $\sum_{j=1}^{J} \exp(\mathbf{x}_i \beta_{j|J})$ to make the probabilities sum to 1

Which results in:

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \beta_{m|J})}{\sum_{j=1}^{J} \exp(\mathbf{x}_i \beta_{j|J})}$$

Identification

One of the β's is constrained to equal zero

For example,

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \beta_{m|J})}{\sum_{j=1}^{J} \exp(\mathbf{x}_i \beta_{j|J})} \quad \text{where } \beta_1 = 0$$

Can be written as:

$$\Pr(y_i = 1 \mid \mathbf{x}_i) = \frac{1}{\sum_{j=2}^{J} \exp(\mathbf{x}_i \beta_j)}$$

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \beta_m)}{\sum_{j=2}^{J} \exp(\mathbf{x}_i \beta_j)} \quad \text{for } m > 1$$

The Data

1982 General Social Survey

A sample of 337 currently employed men

Outcome

Respondents were asked to indicate their occupation

These occupations were recoded to correspond to Schmidt and Strauss (975) in an early application of the MNLM

Occupation categories:

- Menial jobs
- Blue-collar jobs
- Craft jobs
- White-collar jobs
- Professional jobs

Descriptive Information

```
. spex nomocc2

. use "http://www.indiana.edu/~jslsoc/stata/spex_data/nomocc2.dta", clear

. codebook, compact

Variable   Obs Unique      Mean  Min  Max  Label
-------------------------------------------------------------------------------
occ        337      5  3.397626    1    5  Occupation
white      337      2  .9169139    0    1  Race: 1=white 0=nonwhite
ed         337     16  13.09496    3   20  Years of education
exper      337     54  20.50148    2   66  Years of work experience
-------------------------------------------------------------------------------
```



```
. sum

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         occ |       337    3.397626    1.367913          1          5
       white |       337    .9169139    .2764227          0          1
          ed |       337    13.09496    2.946427          3         20
       exper |       337    20.50148    13.95936          2         66

. tab occ

 Occupation |      Freq.     Percent        Cum.
------------+-----------------------------------
     Menial |         31        9.20        9.20
    BlueCol |         69       20.47       29.67
      Craft |         84       24.93       54.60
   WhiteCol |         41       12.17       66.77
       Prof |        112       33.23      100.00
------------+-----------------------------------
      Total |        337      100.00
```


Descriptive Table

Name	Mean	StdDev	Min	Max	Description
OCC					Occupation: *M1*=menial; *B2*=blue collar; *C3*=craft; *W4*=white collar; *P5*=professional
WHITE	0.92	0.28	0.0	1.0	Race: 1= white; 0=another race
ED	13.10	2.95	3.0	20.0	Education: Number of years of formal education.
EXP	20.50	13.96	2.0	66.0	Possible years of work experience: Age minus years of education minus 5.

Note: N=337. OCC has categories: *M1*=menial; *B2*=blue collar; *C3*=craft; *W4*=white collar; *P5*=professional with marginal percentages 9, 21, 25, 12, and 33, respectively.

Estimating the MNLM

```
. mlogit occ white ed exper , base(1)

Iteration 0:   log likelihood = -509.84406
Iteration 1:   log likelihood = -432.18549
Iteration 2:   log likelihood = -426.88668
Iteration 3:   log likelihood = -426.80057
Iteration 4:   log likelihood = -426.80048
Iteration 5:   log likelihood = -426.80048
```



```
Multinomial logistic regression                   Number of obs   =        337
                                                  LR chi2(12)     =     166.09
                                                  Prob > chi2     =     0.0000
Log likelihood = -426.80048                       Pseudo R2       =     0.1629

------------------------------------------------------------------------------
         occ |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Menial       |  (base outcome)
-------------+----------------------------------------------------------------
BlueCol      |
       white |   1.236504   .7244352     1.71   0.088    -.1833631    2.656371
          ed |  -.0994247   .1022812    -0.97   0.331    -.2998922    .1010428
       exper |   .0047212   .0173984     0.27   0.786    -.0293789    .0388214
       _cons |   .7412336    1.51954     0.49   0.626     -2.23701    3.719477
-------------+----------------------------------------------------------------
Craft        |
       white |   .4723436   .6043097     0.78   0.434    -.7120817    1.656769
          ed |   .0938154    .097555     0.96   0.336    -.0973888    .2850197
       exper |   .0276838   .0166737     1.66   0.097     -.004996    .0603636
       _cons |  -1.091353   1.450218    -0.75   0.452    -3.933728    1.751022
-------------+----------------------------------------------------------------
WhiteCol     |
       white |   1.571385   .9027216     1.74   0.082    -.1979166    3.340687
          ed |   .3531577   .1172786     3.01   0.003     .1232959    .5830194
       exper |   .0345959   .0188294     1.84   0.066     -.002309    .0715007
       _cons |  -6.238608   1.899094    -3.29   0.001    -9.960764   -2.516453
-------------+----------------------------------------------------------------
Prof         |
       white |   1.774306   .7550543     2.35   0.019     .2944273    3.254186
          ed |   .7788519   .1146293     6.79   0.000     .5541826    1.003521
       exper |   .0356509    .018037     1.98   0.048      .000299    .0710028
       _cons |  -11.51833   1.849356    -6.23   0.000      -15.143   -7.893659
```


Interpretation

In even a simple MNLM there are a lot of parameters

Too often, the MNLM is estimated, the parameters are listed, and statistical significance is noted, while the magnitudes and even directions of the effects are ignored

We will consider:

- Factor change in the odds (odds ratio)
- Predicted probabilities

Factor change coefficients

For a model with three independent variables,

$$\Omega_{m|n}(\mathbf{x},x_2) = e^{\beta_{0,m|n}} e^{\beta_{1,m|n}x_1} e^{\beta_{2,m|n}x_2} e^{\beta_{3,m|n}x_3}$$

A change of one unit in x_2 can be measured by the ratio of the odds:

$$\frac{\Omega_{m|n}(\mathbf{x},x_2+1)}{\Omega_{m|n}(\mathbf{x},x_2)} = \frac{e^{\beta_{0,m|n}} \ e^{\beta_{1,m|n}x_1} \ e^{\beta_{2,m|n}x_2} \ e^{\beta_{2,m|n}} \ e^{\beta_{3,m|n}x_3}}{e^{\beta_{0,m|n}} \ e^{\beta_{1,m|n}x_1} \ e^{\beta_{2,m|n}x_2} \ e^{\beta_{3,m|n}x_3}} = e^{\beta_{2,m|n}}$$

Interpretation

For a unit change in x_k, the odds are expected to change by a factor of $\exp(\beta_{k,m|n})$, holding all other variables constant

For a standard deviation change in x_k, the odds are expected to change by a factor of $\exp(\beta_{k,m|n} \times s_k)$, holding

Computing factor change

```
. listcoef, help
mlogit (N=337): Factor Change in the Odds of occ

Variable: white (sd=.27642268)
Odds comparing    |
Alternative 1     |
to Alternative 2  |      b         z     P>|z|     e^b   e^bStdX
------------------+---------------------------------------------
BlueCol -Craft    |   0.76416    1.208   0.227   2.1472   1.2352
BlueCol -WhiteCol |  -0.33488   -0.359   0.720   0.7154   0.9116
BlueCol -Prof     |  -0.53780   -0.673   0.501   0.5840   0.8619
BlueCol -Menial   |   1.23650    1.707   0.088   3.4436   1.4075
Craft   -BlueCol  |  -0.76416   -1.208   0.227   0.4657   0.8096
Craft   -WhiteCol |  -1.09904   -1.343   0.179   0.3332   0.7380
Craft   -Prof     |  -1.30196   -2.011   0.044   0.2720   0.6978
Craft   -Menial   |   0.47234    0.782   0.434   1.6037   1.1395
WhiteCol-BlueCol  |   0.33488    0.359   0.720   1.3978   1.0970
WhiteCol-Craft    |   1.09904    1.343   0.179   3.0013   1.3550
WhiteCol-Prof     |  -0.20292   -0.233   0.815   0.8163   0.9455
WhiteCol-Menial   |   1.57139    1.741   0.082   4.8133   1.5440
Prof    -BlueCol  |   0.53780    0.673   0.501   1.7122   1.1603
Prof    -Craft    |   1.30196    2.011   0.044   3.6765   1.4332
Prof    -WhiteCol |   0.20292    0.233   0.815   1.2250   1.0577
Prof    -Menial   |   1.77431    2.350   0.019   5.8962   1.6331
Menial  -BlueCol  |  -1.23650   -1.707   0.088   0.2904   0.7105
Menial  -Craft    |  -0.47234   -0.782   0.434   0.6235   0.8776
Menial  -WhiteCol |  -1.57139   -1.741   0.082   0.2078   0.6477
Menial  -Prof     |  -1.77431   -2.350   0.019   0.1696   0.6123
----------------------------------------------------------------
Variable: ed (sd=2.9464271)
```



```
Odds comparing    |
Alternative 1     |
to Alternative 2  |      b         z     P>|z|     e^b   e^bStdX
------------------+---------------------------------------------
BlueCol -Craft    |  -0.19324   -2.494   0.013   0.8243   0.5659
BlueCol -WhiteCol |  -0.45258   -4.425   0.000   0.6360   0.2636
BlueCol -Prof     |  -0.87828   -8.735   0.000   0.4155   0.0752
BlueCol -Menial   |  -0.09942   -0.972   0.331   0.9054   0.7461
Craft   -BlueCol  |   0.19324    2.494   0.013   1.2132   1.7671
Craft   -WhiteCol |  -0.25934   -2.773   0.006   0.7716   0.4657
Craft   -Prof     |  -0.68504   -7.671   0.000   0.5041   0.1329
Craft   -Menial   |   0.09382    0.962   0.336   1.0984   1.3184
WhiteCol-BlueCol  |   0.45258    4.425   0.000   1.5724   3.7943
WhiteCol-Craft    |   0.25934    2.773   0.006   1.2961   2.1471
WhiteCol-Prof     |  -0.42569   -4.616   0.000   0.6533   0.2853
WhiteCol-Menial   |   0.35316    3.011   0.003   1.4236   2.8308
Prof    -BlueCol  |   0.87828    8.735   0.000   2.4067  13.3002
Prof    -Craft    |   0.68504    7.671   0.000   1.9838   7.5264
Prof    -WhiteCol |   0.42569    4.616   0.000   1.5307   3.5053
Prof    -Menial   |   0.77885    6.795   0.000   2.1790   9.9228
Menial  -BlueCol  |   0.09942    0.972   0.331   1.1045   1.3404
Menial  -Craft    |  -0.09382   -0.962   0.336   0.9105   0.7585
Menial  -WhiteCol |  -0.35316   -3.011   0.003   0.7025   0.3533
Menial  -Prof     |  -0.77885   -6.795   0.000   0.4589   0.1008
----------------------------------------------------------------
```



```
Variable: exper (sd=13.959364)

Odds comparing    |
Alternative 1     |
to Alternative 2  |      b         z     P>|z|     e^b   e^bStdX
------------------+---------------------------------------------
BlueCol -Craft    |  -0.02296   -1.829   0.067   0.9773   0.7258
BlueCol -WhiteCol |  -0.02987   -1.954   0.051   0.9706   0.6590
BlueCol -Prof     |  -0.03093   -2.147   0.032   0.9695   0.6494
BlueCol -Menial   |   0.00472    0.271   0.786   1.0047   1.0681
Craft   -BlueCol  |   0.02296    1.829   0.067   1.0232   1.3779
Craft   -WhiteCol |  -0.00691   -0.495   0.621   0.9931   0.9080
Craft   -Prof     |  -0.00797   -0.627   0.531   0.9921   0.8947
Craft   -Menial   |   0.02768    1.660   0.097   1.0281   1.4717
WhiteCol-BlueCol  |   0.02987    1.954   0.051   1.0303   1.5174
WhiteCol-Craft    |   0.00691    0.495   0.621   1.0069   1.1013
WhiteCol-Prof     |  -0.00106   -0.073   0.941   0.9989   0.9854
WhiteCol-Menial   |   0.03460    1.837   0.066   1.0352   1.6208
Prof    -BlueCol  |   0.03093    2.147   0.032   1.0314   1.5400
Prof    -Craft    |   0.00797    0.627   0.531   1.0080   1.1176
Prof    -WhiteCol |   0.00106    0.073   0.941   1.0011   1.0148
Prof    -Menial   |   0.03565    1.977   0.048   1.0363   1.6449
Menial  -BlueCol  |  -0.00472   -0.271   0.786   0.9953   0.9362
Menial  -Craft    |  -0.02768   -1.660   0.097   0.9727   0.6795
Menial  -WhiteCol |  -0.03460   -1.837   0.066   0.9660   0.6170
Menial  -Prof     |  -0.03565   -1.977   0.048   0.9650   0.6079
----------------------------------------------------------------
```

Computing all contrasts at a given p value and for one x

```
. listcoef white, pval(.05)

mlogit (N=337): Factor Change in the Odds of occ when P>|z| < 0.05

Variable: white (sd=.27642268)

Odds comparing    |
Alternative 1     |
to Alternative 2  |      b         z     P>|z|     e^b   e^bStdX
------------------+---------------------------------------------
Menial   -Prof    |  -1.77431   -2.350   0.019   0.1696   0.6123
Craft    -Prof    |  -1.30196   -2.011   0.044   0.2720   0.6978
Prof     -Menial  |   1.77431    2.350   0.019   5.8962   1.6331
Prof     -Craft   |   1.30196    2.011   0.044   3.6765   1.4332
----------------------------------------------------------------
```


Predicted probabilities

As before

`predict`,

`prtab`,

`prchange`,

`prgen`, and

`prvalue` can be used

Something new

A discrete change plot

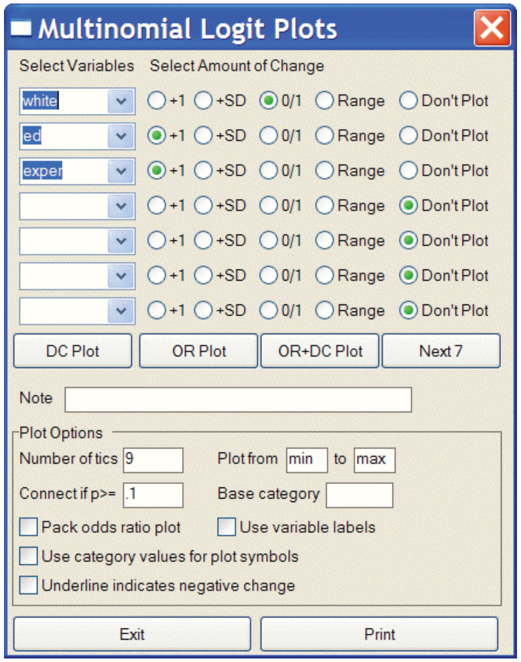
The steps:

Run `mlogit`, for example: `mlogit occ exper ed white`

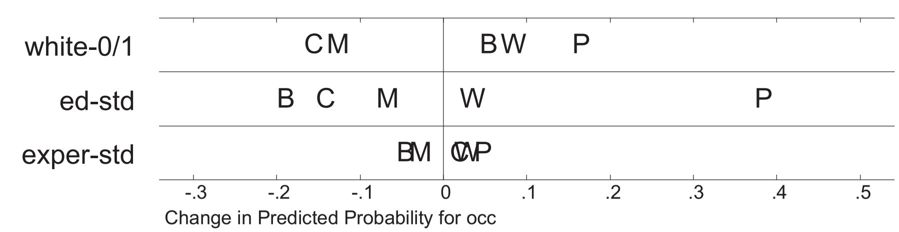
Run `prchange`, for example: `prchange, rest(mean)`

Type `mlogview` (this step doesn't work on Mac) and you will see:


```
. quietly mlogit occ white ed exper , base(1)

. prchange, rest(mean)

mlogit: Changes in Probabilities for occ

white
        Avg|Chg|     BlueCol       Craft    WhiteCol        Prof      Menial
0->1   .11623582   .04981799  -.15973434   .07971004    .1610615  -.13085523

ed
            Avg|Chg|     BlueCol       Craft    WhiteCol        Prof      Menial
Min->Max   .39242268  -.70077323  -.15010394   .02425591   .95680079  -.13017954
   -+1/2   .05855425  -.06831616  -.05247185   .01250795   .13387768  -.02559762
  -+sd/2    .1640657  -.19310513  -.14576758   .03064777   .37951647  -.07129153
MargEfct   .05894859  -.06870635  -.05287415   .01282041   .13455107  -.02579097

exper
            Avg|Chg|     BlueCol       Craft    WhiteCol        Prof      Menial
Min->Max   .12193559  -.18947365   .03115708   .09478889   .17889298  -.11536534
   -+1/2   .00233425  -.00356567   .00105992    .0016944   .00308132  -.00226997
  -+sd/2   .03253578  -.04966453   .01479983   .02360725   .04293236  -.03167491
MargEfct   .00233427  -.00356571   .00105992   .00169442   .00308134  -.00226997

           BlueCol       Craft    WhiteCol        Prof      Menial
Pr(y|x)  .18419114   .29411051   .16112968   .26630062   .09426806

          white        ed     exper
    x=  .916914    13.095   20.5015
 sd_x=  .276423   2.94643   13.9594
```



Discrete Change Plot



Adding CI to the DC

```
. quietly prvalue , x(white=0) rest(mean) save

. prvalue , x(white=1) rest(mean) diff

mlogit: Change in Predictions for occ

Confidence intervals by delta method

                     Current     Saved    Change    95% CI for Change
  Pr(y=BlueCol|x):    0.1862    0.1363    0.0498   [-0.0488,   0.1484]
  Pr(y=Craft|x):      0.2790    0.4387   -0.1597   [-0.3074,  -0.0120]
  Pr(y=WhiteCol|x):   0.1674    0.0877    0.0797   [-0.0104,   0.1698]
  Pr(y=Prof|x):       0.2814    0.1204    0.1611   [ 0.0668,   0.2554]
  Pr(y=Menial|x):     0.0860    0.2168   -0.1309   [-0.2544,  -0.0073]

              white         ed       exper
Current=          1  13.094955   20.501484
  Saved=          0  13.094955   20.501484
   Diff=          1          0           0
```



```
. sum ed

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
          ed |       337    13.09496    2.944237          3         20

. local lbsd = r(mean) - (r(sd)/2)
. local ubsd = r(mean) + (r(sd)/2)

. quietly prvalue , x(ed=`lbsd') rest(mean) save
. prvalue , x(ed=`ubsd') rest(mean) diff

mlogit: Change in Predictions for occ

Confidence intervals by delta method

                     Current     Saved    Change    95% CI for Change
  Pr(y=BlueCol|x):    0.0936    0.2866   -0.1930   [-0.2287,  -0.1573]
  Pr(y=Craft|x):      0.1986    0.3443   -0.1457   [-0.1916,  -0.0998]
  Pr(y=WhiteCol|x):   0.1594    0.1288    0.0306   [-0.0090,   0.0703]
  Pr(y=Prof|x):       0.4930    0.1137    0.3793   [ 0.3271,   0.4314]
  Pr(y=Menial|x):     0.0554    0.1267   -0.0712   [-0.0975,  -0.0450]

              white         ed       exper
Current=  .91691395  14.567074   20.501484
  Saved=  .91691395  11.622837   20.501484
   Diff=          0  2.9442373           0
```

```
. sum exper

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       exper |       337    20.50148    13.94898          2         66

. local lbsd = r(mean) - (r(sd)/2)
. local ubsd = r(mean) + (r(sd)/2)

. quietly prvalue , x(exper=`lbsd') rest(mean) save
. prvalue , x(exper=`ubsd') rest(mean) diff

mlogit: Change in Predictions for occ

Confidence intervals by delta method

                     Current     Saved    Change   95% CI for Change
  Pr(y=BlueCol|x):    0.1602    0.2099   -0.0496  [-0.0814,  -0.0179]
  Pr(y=Craft|x):      0.3003    0.2855    0.0148  [-0.0239,   0.0535]
  Pr(y=WhiteCol|x):   0.1727    0.1491    0.0236  [-0.0083,   0.0555]
  Pr(y=Prof|x):       0.2874    0.2445    0.0429  [ 0.0011,   0.0847]
  Pr(y=Menial|x):     0.0794    0.1110   -0.0317  [-0.0561,  -0.0072]

             white          ed       exper
Current=  .91691395  13.094955  27.475978
  Saved=  .91691395  13.094955  13.526989
   Diff=          0          0  13.948989
```



```
mlogplot white ed exper, dc std(0ss) ///
    min(-.3) max(.5) ntics(9) ///
    note(P=Prof; W=WhiteCol; C=Craft; B=BlueCol; M=Menial)
graph export mnlm-01-DCplot.wmf , replace
```

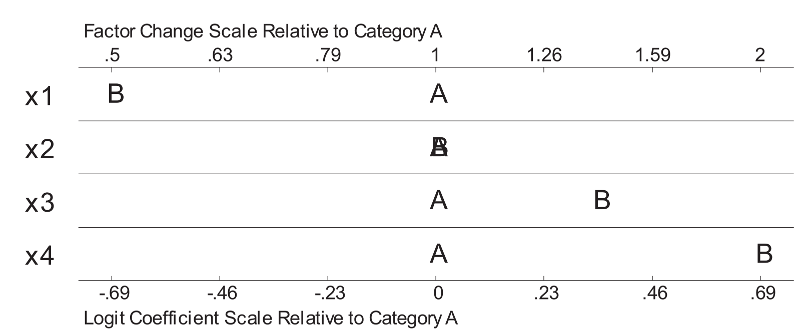

Getting the Odds Ratio out of the "doghouse"

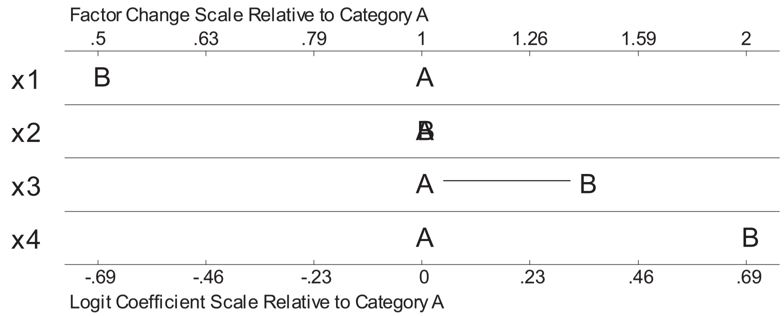
Discrete change does not indicate the dynamics among the dependent outcomes

For example, a decrease in education increases the probability of both blue collar and craft jobs, but, how does it affect the odds of a person choosing a craft job relative to a blue collar job?

To answer these questions, consider the factor change in the odds

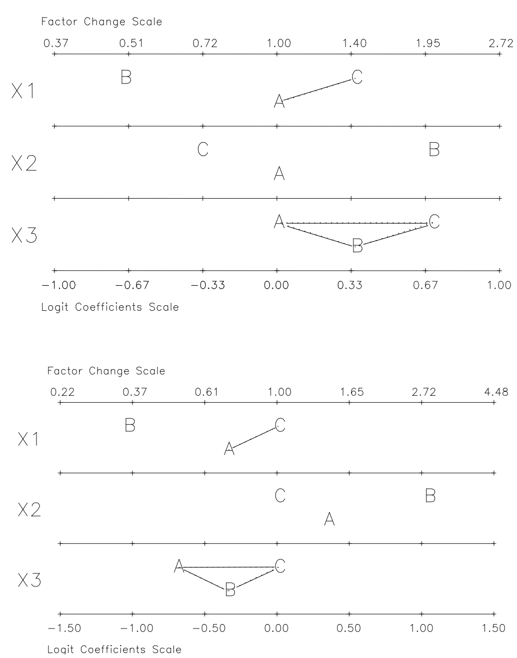
	$\beta_{B\|A}$	$\exp(\beta_{B\|A})$	p-value
x_1	-0.693	0.500	0.02
x_2	0.000	1.000	0.99
x_3	0.347	1.414	0.11
x_4	0.693	2.000	0.04





Consider a hypothetical model with three outcomes:

Comparison		Logit Coefficient for x_1	x_2	x_3
$B \mid A$	$\beta_{B\mid A}$	-0.693	0.693	0.347
	$\exp(\beta_{B\mid A})$	0.500	2.000	1.414
	p	0.04	0.01	0.42
$C \mid A$	$\beta_{C\mid A}$	0.347	-0.347	0.693
	$\exp(\beta_{C\mid A})$	1.414	0.707	2.000
	p	0.21	0.04	0.37
$C \mid B$	$\beta_{C\mid B}$	1.040	-1.040	0.346
	$\exp(\beta_{C\mid B})$	2.828	0.354	1.414
	p	0.02	0.03	0.21



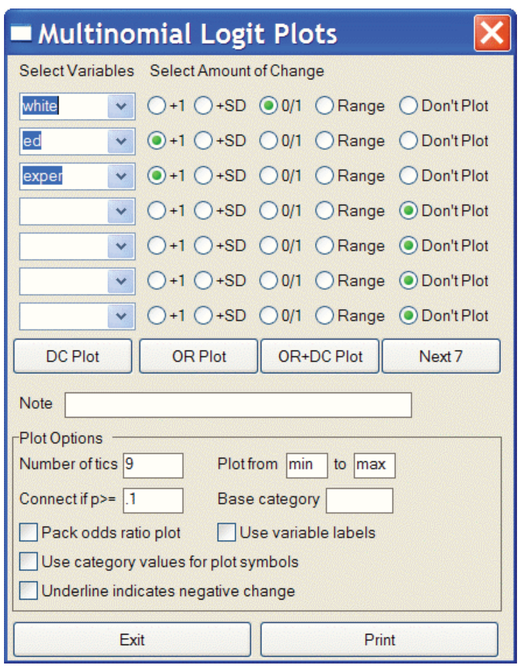

```
. listcoef, help
mlogit (N=337): Factor Change in the Odds of occ

Variable: white (sd=.27642268)
Odds comparing    |
Alternative 1     |
to Alternative 2  |      b         z     P>|z|     e^b   e^bStdX
------------------+---------------------------------------------
BlueCol -Craft    |   0.76416    1.208   0.227   2.1472   1.2352
BlueCol -WhiteCol |  -0.33488   -0.359   0.720   0.7154   0.9116
BlueCol -Prof     |  -0.53780   -0.673   0.501   0.5840   0.8619
BlueCol -Menial   |   1.23650    1.707   0.088   3.4436   1.4075
Craft   -BlueCol  |  -0.76416   -1.208   0.227   0.4657   0.8096
Craft   -WhiteCol |  -1.09904   -1.343   0.179   0.3332   0.7380
Craft   -Prof     |  -1.30196   -2.011   0.044   0.2720   0.6978
Craft   -Menial   |   0.47234    0.782   0.434   1.6037   1.1395
WhiteCol-BlueCol  |   0.33488    0.359   0.720   1.3978   1.0970
WhiteCol-Craft    |   1.09904    1.343   0.179   3.0013   1.3550
WhiteCol-Prof     |  -0.20292   -0.233   0.815   0.8163   0.9455
WhiteCol-Menial   |   1.57139    1.741   0.082   4.8133   1.5440
Prof    -BlueCol  |   0.53780    0.673   0.501   1.7122   1.1603
Prof    -Craft    |   1.30196    2.011   0.044   3.6765   1.4332
Prof    -WhiteCol |   0.20292    0.233   0.815   1.2250   1.0577
Prof    -Menial   |   1.77431    2.350   0.019   5.8962   1.6331
Menial  -BlueCol  |  -1.23650   -1.707   0.088   0.2904   0.7105
Menial  -Craft    |  -0.47234   -0.782   0.434   0.6235   0.8776
Menial  -WhiteCol |  -1.57139   -1.741   0.082   0.2078   0.6477
Menial  -Prof     |  -1.77431   -2.350   0.019   0.1696   0.6123
----------------------------------------------------------------
```

Variable: ed (sd=2.9464271)

Odds comparing Alternative 1 to Alternative 2	b	z	P>|z|	e^b	e^bStdX
BlueCol -Craft	-0.19324	-2.494	0.013	0.8243	0.5659
BlueCol -WhiteCol	-0.45258	-4.425	0.000	0.6360	0.2636
BlueCol -Prof	-0.87828	-8.735	0.000	0.4155	0.0752
BlueCol -Menial	-0.09942	-0.972	0.331	0.9054	0.7461
Craft -BlueCol	0.19324	2.494	0.013	1.2132	1.7671
Craft -WhiteCol	-0.25934	-2.773	0.006	0.7716	0.4657
Craft -Prof	-0.68504	-7.671	0.000	0.5041	0.1329
Craft -Menial	0.09382	0.962	0.336	1.0984	1.3184
WhiteCol-BlueCol	0.45258	4.425	0.000	1.5724	3.7943
WhiteCol-Craft	0.25934	2.773	0.006	1.2961	2.1471
WhiteCol-Prof	-0.42569	-4.616	0.000	0.6533	0.2853
WhiteCol-Menial	0.35316	3.011	0.003	1.4236	2.8308
Prof -BlueCol	0.87828	8.735	0.000	2.4067	13.3002
Prof -Craft	0.68504	7.671	0.000	1.9838	7.5264
Prof -WhiteCol	0.42569	4.616	0.000	1.5307	3.5053
Prof -Menial	0.77885	6.795	0.000	2.1790	9.9228
Menial -BlueCol	0.09942	0.972	0.331	1.1045	1.3404
Menial -Craft	-0.09382	-0.962	0.336	0.9105	0.7585
Menial -WhiteCol	-0.35316	-3.011	0.003	0.7025	0.3533
Menial -Prof	-0.77885	-6.795	0.000	0.4589	0.1008

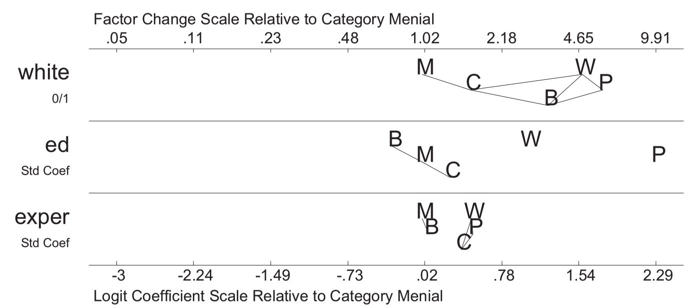
Variable: exper (sd=13.959364)

Odds comparing Alternative 1 to Alternative 2	b	z	P>|z|	e^b	e^bStdX
BlueCol -Craft	-0.02296	-1.829	0.067	0.9773	0.7258
BlueCol -WhiteCol	-0.02987	-1.954	0.051	0.9706	0.6590
BlueCol -Prof	-0.03093	-2.147	0.032	0.9695	0.6494
BlueCol -Menial	0.00472	0.271	0.786	1.0047	1.0681
Craft -BlueCol	0.02296	1.829	0.067	1.0232	1.3779
Craft -WhiteCol	-0.00691	-0.495	0.621	0.9931	0.9080
Craft -Prof	-0.00797	-0.627	0.531	0.9921	0.8947
Craft -Menial	0.02768	1.660	0.097	1.0281	1.4717
WhiteCol-BlueCol	0.02987	1.954	0.051	1.0303	1.5174
WhiteCol-Craft	0.00691	0.495	0.621	1.0069	1.1013
WhiteCol-Prof	-0.00106	-0.073	0.941	0.9989	0.9854
WhiteCol-Menial	0.03460	1.837	0.066	1.0352	1.6208
Prof -BlueCol	0.03093	2.147	0.032	1.0314	1.5400
Prof -Craft	0.00797	0.627	0.531	1.0080	1.1176
Prof -WhiteCol	0.00106	0.073	0.941	1.0011	1.0148
Prof -Menial	0.03565	1.977	0.048	1.0363	1.6449
Menial -BlueCol	-0.00472	-0.271	0.786	0.9953	0.9362
Menial -Craft	-0.02768	-1.660	0.097	0.9727	0.6795
Menial -WhiteCol	-0.03460	-1.837	0.066	0.9660	0.6170
Menial -Prof	-0.03565	-1.977	0.048	0.9650	0.6079



Odds Ratio Plot

P=Prof; W=WhiteCol; C=Craft; B=BlueCol; M=Menial



What do you see?

Question (for you)

Note the different ordering of categories for the different variables

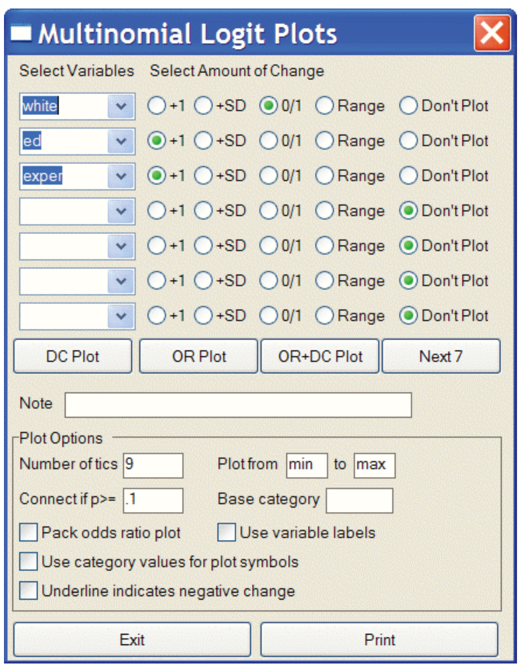
Would the OLM allow for this different ordering?

Why or why not?

Why predicted probabilities remain important

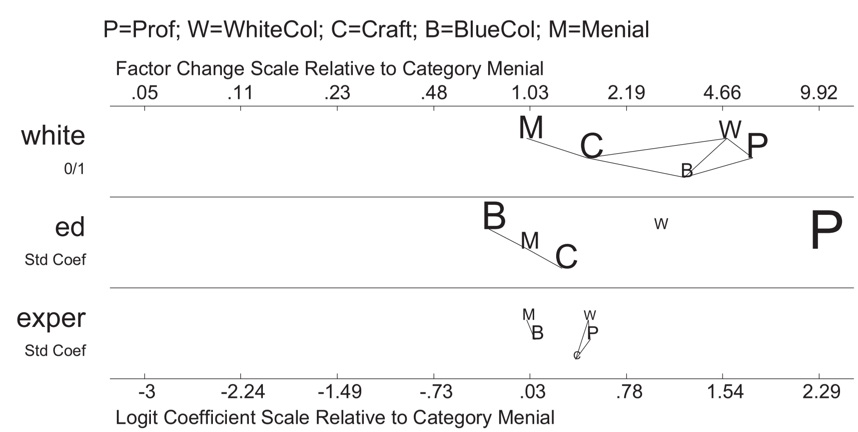
While the factor change in the odds is constant across the levels of all variables, the discrete changes get larger or smaller at different values of the variables

If the odds increase by a factor of ten but the current odds are 1 in 10,000, then the substantive impact is small.



Putting it all together

Incorporate information about the discrete change in the probability by making the height of the letter in the odds ratio plot proportional to the square root of the DC



Stata Code

```
. //OR plot
. mlogplot white ed exper, or std(0ss) prob(.1) ///
>         min(-3) max(.5) ntics(8) ///
>         note(P=Prof; W=WhiteCol; C=Craft; B=BlueCol; M=Menial)

. graph export mnlm-02-ORplot.wmf , replace


. //DC and OR combined
. mlogplot white ed exper, dc or std(0ss) prob(.1) ///
>         min(-3) max(.5) ntics(8) ///
>         note(P=Prof; W=WhiteCol; C=Craft; B=BlueCol; M=Menial)

. graph export mnlm-03-DCORplot.wmf , replace
```


Testing that a Variable Has No Effect

The hypothesis that x_k does not affect the dependent variable can be written as:

$$H_0: \beta_{k,B|M} = \beta_{k,C|M} = \beta_{k,W|M} = \beta_{k,P|M} = 0$$

LR test using `lrtest`

```
. quietly mlogit occ white ed exper, base(1) nolog
. estimates store base
. quietly mlogit occ white exper, base(1) nolog
. estimates store noed

. lrtest base noed

Likelihood-ratio test                                 LR chi2(4)  =    156.94
(Assumption: noed nested in base)                     Prob > chi2 =    0.0000
```

Wald test using `test`

```
. quietly mlogit occ white ed exper,base(1)
. test ed

 ( 1)  [Menial]o.ed = 0
 ( 2)  [BlueCol]ed = 0
 ( 3)  [Craft]ed = 0
 ( 4)  [WhiteCol]ed = 0
 ( 5)  [Prof]ed = 0
       Constraint 1 dropped

           chi2(  4) =   84.97
         Prob > chi2 =    0.0000
```


Either, using `mlogtest`

```
. quietly mlogit occ white ed exper,base(1)
. mlogtest ed , lr wald

**** Likelihood-ratio tests for independent variables (N=337)

 Ho: All coefficients associated with given variable(s) are 0.

             |       chi2    df   P>chi2
-------------+-------------------------
          ed |    156.937     4    0.000
---------------------------------------

**** Wald tests for independent variables (N=337)

 Ho: All coefficients associated with given variable(s) are 0.

             |       chi2    df   P>chi2
-------------+-------------------------
          ed |     84.968     4    0.000
---------------------------------------
```

Testing that outcome categories can be combined

The hypothesis that P and W are indistinguishable is

$$H_0: \beta_{1,P|W} = \beta_{2,P|W} = \beta_{3,P|W} = 0$$

A Wald test using `mlogtest`

```
. mlogtest, combine

**** Wald tests for combining outcome categories

 Ho: All coefficients except intercepts associated with given pair
     of outcomes are 0 (i.e., categories can be collapsed).

Categories tested |      chi2   df   P>chi2
------------------+------------------------
  Menial- BlueCol |     3.994    3    0.262
  Menial-   Craft |     3.203    3    0.361
  Menial-WhiteCol |    11.951    3    0.008
  Menial-    Prof |    48.190    3    0.000
 BlueCol-   Craft |     8.441    3    0.038
 BlueCol-WhiteCol |    20.055    3    0.000
 BlueCol-    Prof |    76.393    3    0.000
   Craft-WhiteCol |     8.892    3    0.031
   Craft-    Prof |    60.583    3    0.000
WhiteCol-    Prof |    22.203    3    0.000
-------------------------------------------
```


A LR test using `mlogtest`

```
. mlogtest, lrcom

**** LR tests for combining outcome categories

 Ho: All coefficients except intercepts associated with given pair
     of outcomes are 0 (i.e., categories can be collapsed).

Categories tested |      chi2   df   P>chi2
------------------+------------------------
  Menial- BlueCol |     4.095    3    0.251
  Menial-   Craft |     3.376    3    0.337
  Menial-WhiteCol |    13.223    3    0.004
  Menial-    Prof |    64.607    3    0.000
 BlueCol-   Craft |     9.176    3    0.027
 BlueCol-WhiteCol |    22.803    3    0.000
 BlueCol-    Prof |   125.699    3    0.000
   Craft-WhiteCol |     9.992    3    0.019
   Craft-    Prof |    95.889    3    0.000
WhiteCol-    Prof |    26.736    3    0.000
-------------------------------------------
```


Question (for you)

Do you notice any logical inconsistencies?

Specification Searches

Given the complexities in interpreting the MNLM, it is tempting to search for a more parsimonious model constructed by excluding variables or combining outcome categories

Tests for combining categories and that all coefficients for a variable are zero can guide a specification search, but great care is required to avoid over-fitting or misfitting the model

Independence of Irrelevant Alternatives

For a model with outcome categories M, N, and L

In the MNLM, the odds of M compared to N do not depend on :L

$$\frac{\Omega_{m|n}(\mathbf{x}, x_2+1)}{\Omega_{m|n}(\mathbf{x}, x_2)} = \frac{e^{\beta_{0,m|n}} \quad e^{\beta_{1,m|n}x_1} \quad e^{\beta_{2,m|n}x_2} \quad e^{\beta_{2,m|n}} \quad e^{\beta_{3,m|n}x_3}}{e^{\beta_{0,m|n}} \quad e^{\beta_{1,m|n}x_1} \quad e^{\beta_{2,m|n}x_2} \quad e^{\beta_{3,m|n}x_3}} = e^{\beta_{2,m|n}}$$

In other words, outcome L is irrelevant to the comparison of M to N

This property is called the *independence of irrelevant alternatives* (IIA)

McFadden's Classic example of IIA

A person has two choices:

$\Pr(\text{car}) = 1/2$ and $\Pr(\text{red bus}) = 1/2$

Odds of taking the car versus the red bus are

$$\frac{\Pr(\text{car})}{\Pr(\text{red bus})} = \frac{1/2}{1/2} = 1$$

A new bus company opens with identical service to the red bus

IIA requires

$\Pr(\text{car}) = 1/3;$ $\Pr(\text{red bus}) = 1/3;$ $\Pr(\text{blue bus}) = 1/3$

So that the original odds can be maintained

$$\frac{\Pr(\text{car})}{\Pr(\text{red bus})} = 1 = \frac{1/3}{1/3}$$

What makes sense,

$\Pr(\text{car}) = 1/2; \quad \Pr(\text{red bus}) = 1/4; \quad \Pr(\text{blue bus}) = 1/4$

Violates IIA,

$$\frac{\Pr(\text{car})}{\Pr(\text{red bus})} = \frac{1/2}{1/4} = 2$$

This implies that MNLM should only be used in cases where the outcome categories can plausibly be assumed to be distinct

Red bus and Blue bus can be viewed as "perfect substitutes"

Care in specifying the model to involve distinct outcomes that are not substitutes for one another seems to be reasonable advice

But, many reviewers like to see formal tests of IIA…

Formal tests of IIA

Hausman-type test

Comparison of two estimators of the same parameter
- One estimator is consistent and efficient if the null hypothesis is true
- The second estimator is consistent but inefficient.

Question (for you)

What would be a consistent but inefficient estimator?


```
. set seed 112

. mlogtest , iia

**** Hausman tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |      chi2   df   P>chi2   evidence
---------+------------------------------------
 BlueCol |     0.333   12    1.000   for Ho
   Craft |   -14.436   12     ---    ---
WhiteCol |    -7.764   12     ---    ---
    Prof |    -0.119   12     ---    ---
----------------------------------------------
 Note: If chi2<0, the estimated model does not
 meet asymptotic assumptions of the test.

**** suest-based Hausman tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |      chi2   df   P>chi2   evidence
---------+------------------------------------
 BlueCol |     3.513   12    0.991   for Ho
   Craft |    25.388   12    0.013   against Ho
WhiteCol |     6.541   12    0.886   for Ho
    Prof |     6.850   12    0.867   for Ho
----------------------------------------------
```

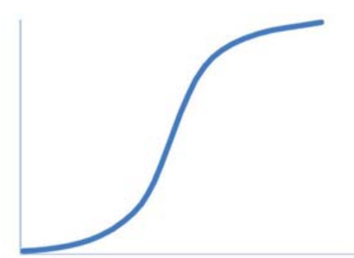
```
**** Small-Hsiao tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |  lnL(full)  lnL(omit)    chi2   df   P>chi2   evidence
---------+------------------------------------------------------------
 BlueCol |   -173.821   -143.396  60.850   12    0.000   against Ho
   Craft |   -169.717   -137.746  63.942   12    0.000   against Ho
WhiteCol |   -172.867   -163.953  17.829   12    0.121   for Ho
    Prof |   -176.871   -148.980  55.781   12    0.000   against Ho
----------------------------------------------------------------------
```



```
. set seed 1821

. mlogtest , iia

**** Hausman tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |      chi2   df   P>chi2   evidence
---------+------------------------------------
 BlueCol |     0.333   12    1.000   for Ho
   Craft |   -14.436   12     ---    ---
WhiteCol |    -7.764   12     ---    ---
    Prof |    -0.119   12     ---    ---
----------------------------------------------
 Note: If chi2<0, the estimated model does not
 meet asymptotic assumptions of the test.

**** suest-based Hausman tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |      chi2   df   P>chi2   evidence
---------+------------------------------------
 BlueCol |     3.513   12    0.991   for Ho
   Craft |    25.388   12    0.013   against Ho
WhiteCol |     6.541   12    0.886   for Ho
    Prof |     6.850   12    0.867   for Ho
----------------------------------------------
```



```
**** Small-Hsiao tests of IIA assumption (N=337)

 Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

 Omitted |  lnL(full)  lnL(omit)    chi2   df   P>chi2   evidence
---------+------------------------------------------------------------
 BlueCol |   -129.048   -125.766   6.564   12    0.885   for Ho
   Craft |   -122.878   -118.103   9.549   12    0.655   for Ho
WhiteCol |   -151.232   -147.833   6.797   12    0.871   for Ho
    Prof |   -129.995   -126.344   7.302   12    0.837   for Ho
----------------------------------------------------------------------
```

End MNLM