Intermediate Social Statistics Hilary 2009 Lecture 6: Multinomial Logit

Raymond Duch

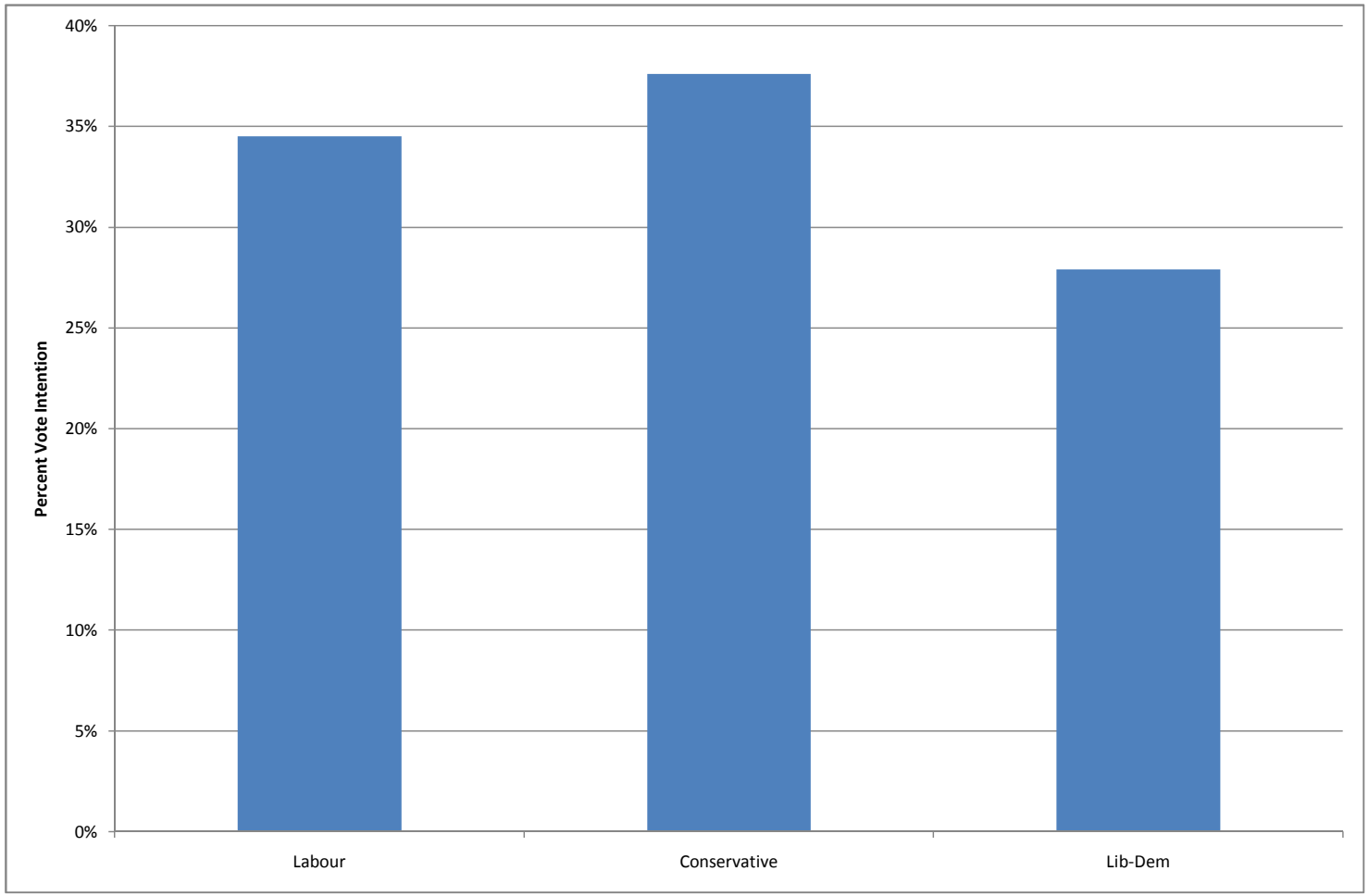
Nuffield College

February 25, 2009

A Multinomial Dependent Variable

- Lets return to the vote variable from the UK 2004 election study
- Instead of two, the dependent variable takes on three values
- lets say that we want to model a vote dependent variable that has a value of 1 for Labour, 2 for Conservative and 3 for Liberal Democrats
- Note these values are not ordinal

Vote Intention Data



A Multinomial Dependent Variable

- we have n decision makers - in this case n respondents to the survey indexed by i
- there are J alternatives or choices ranging from 1 through J
  - j=1 (Labour)
  - j=2 (Conservative)
  - j=3 (Lib-Dem)
- we observe $y = 1$ if i chooses j
- we observe a vector of characteristics z_i for decision maker i (each respondent in the survey)
  - $z_{i,1}$=income
  - $z_{i,2}$=left-right identification
  - etc.

The Data

	vote	age	male	urban	own	union	relig	hhincome	retnat	class
1.	labour	74	0	1	0	0	none	460	worse	working
2.	.	45	1	1	1	0	4	0	better	working
3.	labour	34	0	1	1	0	several times a week	1	same	middle
4.	.	41	0	1	0	0	once a year	1000	same	working
5.	.	68	1	1	1	0	once a year	0	better	working
6.	lib-dem	19	1	1	0	0	none	1	same	working
7.	.	50	0	1	0	0	none	400	same	working
8.	lib-dem	48	1	1	0	0	once a year	0	better	working
9.	con	24	1	1	1	1	4	1	same	lower middle
10.	con	43	0	1	0	0	none	400	worse	working
11.	labour	55	1	0	1	0	none	1500	same	working
12.	con	31	1	1	1	0	once a year	1500	worse	working
13.	lib-dem	38	1	1	0	1	none	2400	better	working
14.	.	43	0	1	1	0	3	800	same	working
15.	.	52	0	1	1	0	3	0	worse	working

Generating Choice Probabilities

We want to figure out the probability that an individual i votes for party j given their particular characteristics:

$$P_{ij} = Pr(y_i = j|z_i) \tag{1}$$

where $P_{ij} \infty \exp^{(\mathbf{x_i}\beta_\mathbf{j})}$ with constraint: $P_{ij} \in [0, 1]$

Generating Choice Probabilities

The probabilities across choices must sum to one so

$$\sum_{j=1}^{J} P_{ij} = 1 \Rightarrow P_{ij} \frac{\exp^{(\mathbf{x_i}\beta_\mathbf{j})}}{\sum_{k=1}^{J} \exp^{(\mathbf{x_i}\beta_\mathbf{k})}} \tag{2}$$

Generating Choice Probabilities

- Note the unknown parameters in this model are: $\beta_1, \beta_2, \ldots \beta_j$
- We cannot estimate all these parameters
- But with $J = 3$ alternatives in our election example corresponding to 3 parties
- Once we know choice probabilities of $J - 1$ of these alternatives we know probability of choosing remaining alternative
- We only need to estimate probability of choosing 2 parties – probability of 3rd is given

Generating Choice Probabilities

- But with $J = 3$ alternatives in our election example corresponding to 3 parties
- We can only estimate $J - 1$ identifiable parameter vectors $(\beta_{\mathbf{k}})$
- Once we have $J - 1$ then J is given
- We can eliminate one of the probabilities as an unknown piece of information by simply setting $\beta_{\mathbf{1}} = 0$

Generating Choice Probabilities

$$P_{ij} = \frac{\frac{\exp^{(\mathbf{x_i}\beta_\mathbf{j})}}{\exp^{(\mathbf{x_i}\beta_\mathbf{1})}}}{\sum_{k=1}^{J} \frac{\exp^{(\mathbf{x_i}\beta_\mathbf{j})}}{\exp^{(\mathbf{x_i}\beta_\mathbf{1})}}} \tag{3}$$

$$P_{ij} = \frac{\exp^{(\mathbf{x_i}[\beta_\mathbf{j}-\beta_\mathbf{1}])}}{\exp^{(\mathbf{x_i}\beta_\mathbf{1})} + \exp^{(\mathbf{x_i}\beta_\mathbf{2})} + \ldots + \exp^{(\mathbf{x_i}\beta_\mathbf{j})}} \tag{4}$$

$$P_{ij} = \frac{\exp^{(\mathbf{x_i}[\beta_\mathbf{j}-\beta_\mathbf{1}])}}{1 + \sum_{k=2}^{J} \exp^{(\mathbf{x_i}[\beta_\mathbf{k}-\beta_\mathbf{1}])}} \tag{5}$$

$$P_{ij} = \frac{\exp^{(\mathbf{x_i}[\beta_\mathbf{j}'])}}{1 + \sum_{k=2}^{J} \exp^{(\mathbf{x_i}[\beta_\mathbf{k}'])}} \tag{6}$$

Estimation

First define a set of outcome variables, d_{ij}, where $d_{ij} = 1$ if $y_i = j$ and $d_{ij} = 0$ otherwise

$$LnL = \sum_{i=1}^{n} \sum_{j=1}^{J} d_{ij} ln P_{ij} \tag{7}$$

Estimating Multinomial Logit Models

```
Multinomial logistic regression                 Number of obs   =        785
                                                LR chi2(14)     =     233.16
                                                Prob > chi2     =     0.0000
Log likelihood = -736.13494                     Pseudo R2       =     0.1367
------------------------------------------------------------------------------
        vote |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
labour       |
      retnat |  -.9268795   .1203081    -7.70   0.000    -1.162679   -.6910799
       class |  -.4233506   .0985861    -4.29   0.000    -.6165758   -.2301253
       union |   .7310438   .2297756     3.18   0.001     .2806919    1.181396
   southwest |  -.3820443    .410329    -0.93   0.352    -1.186274    .4221858
       urban |   .5507096   .2061001     2.67   0.008     .1467609    .9546583
      lrself |  -.3056163   .0448015    -6.82   0.000    -.3934257   -.2178069
         own |  -.8494617   .2302677    -3.69   0.000    -1.300778   -.3981452
       _cons |   4.503524   .4934129     9.13   0.000     3.536452    5.470596
-------------+----------------------------------------------------------------
lib-dem      |
      retnat |  -.4896131   .1222184    -4.01   0.000    -.7291568   -.2500695
       class |  -.0825112   .0999756    -0.83   0.409    -.2784597    .1134373
       union |    .490678   .2413226     2.03   0.042     .0176944    .9636615
   southwest |   .9437263   .3353735     2.81   0.005     .2864062    1.601046
       urban |   .4391652   .2103857     2.09   0.037     .0268168    .8515136
      lrself |  -.3596167   .0474539    -7.58   0.000    -.4526245   -.2666088
         own |   -.563892   .2448292    -2.30   0.021    -1.043748   -.0840356
       _cons |   2.819705   .5077238     5.55   0.000     1.824585    3.814826
------------------------------------------------------------------------------
(vote==con is the base outcome)
```

Interpreting the Coefficients: Predicted Probabilities

- P_{ij} where the ith individual is a high status voter and the following characteristics:
  - retnat=1 (satisfied with the economy)
  - class=4 (upper middle class)
  - union=0 (nonunion)
  - southwest=1 (lives in southwest)
  - urban=1 (urban)
  - lrself=7 (centre-right)
  - own=1 (owns their own home)

Interpreting the Coefficients: Predicted Probabilities

- j in this case represents a Labour vote choice
- The probability of this person voting Labour can be calculated from the earlier equation:

(1.5)

$$\frac{\exp\left[1.96-.92(retnat)+.42(class)+.73(union)-.38(southwest)+.55(urban)-.31(lrself)-.85(own)\right]}{1+\exp\left[1.96-.92(retnat)+.42(class)+.73(union)-.38(southwest)+.55(urban)-.31(lrself)-.85(own)\right]+\exp\left[2.32-.49(retnat)+.08(class)+.49(union)+.94(southwest)+.43(urban)-.35(lrself)-.56(own)\right]}$$

Generating Point Predictions with Standard Errors

- Generate these predicted probabilities with spost or clarify
- Incorporate measures of uncertain

```
. simqi

      Quantity of Interest |      Mean       Std. Err.     [95% Conf. Interval]
---------------------------+--------------------------------------------------
           Pr(vote=labour) |   .3581655      .0889931      .1956686    .5431168
              Pr(vote=con) |   .1723673      .0504017      .0878398    .2827819
          Pr(vote=lib-dem) |   .4694672      .0877185      .3018294    .6350466
```

Generating Discrete Changes in Predicted Probabilities along with Standard Errors

- Define plausible changes in specific independent variable(s)
- Set other independent variables to plausible/interesting values

$$\frac{\Delta Pr(y = j|z)}{\Delta_{x_k}} = Pr(y = j|z, z_k = z_E) - Pr(y = j|z, z_k = z_S) \quad (8)$$

where z_E is an hypothesized high value on z_k and z_S is an hypothesized low value on z_k

Generating Discrete Changes in Predicted Probabilities along with Standard Errors

```
. simqi, fd(pr) changex(retnat 2 3)

First Difference: retnat 2 3

      Quantity of Interest |      Mean       Std. Err.     [95% Conf. Interval]
---------------------------+-----------------------------------------------------
             dPr(vote = 1) |  -.0956782      .0261542    -.1523756   -.0497648
             dPr(vote = 2) |   .1355281      .0303854     .0794868     .197017
             dPr(vote = 3) |    -.03985      .0332206    -.1056441    .0233147
```

Odds Ratios

$$\partial \frac{ln(\frac{Pr(y=1|X)}{1-Pr(y=1|X)})}{\partial x_k} = \partial \frac{(X_i\beta)}{\partial x_k} = \beta_k \tag{9}$$

Odds Ratio

- Allows nice way to summarize the impact of any variable on the odds ratio
- For a unit change in x_k, we can expect the log odds ratio to change by β_k, holding all other variables constant

For Logit use the logistic command

```
. logistic incumvote retnat class union southwest urban lrself own

Logistic regression                                     Number of obs   =       1161
                                                        LR chi2(7)      =     123.80
                                                        Prob > chi2     =     0.0000
Log likelihood = -572.50511                             Pseudo R2       =     0.0976

------------------------------------------------------------------------------
   incumvote | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      retnat |   .4670983    .043633    -8.15   0.000     .3889512    .5609466
       class |   1.289582   .1001786     3.27   0.001     1.107452    1.501664
       union |   1.509615   .2564867     2.42   0.015     1.082045    2.106139
   southwest |   .4187134   .1377687    -2.65   0.008     .2197086    .7979699
       urban |   1.488835   .2456241     2.41   0.016     1.077501    2.057195
      lrself |   .9181869   .0313531    -2.50   0.012     .8587472    .9817409
         own |   .6058125   .1018479    -2.98   0.003     .4357486    .8422489
```

For Multinomial Logit use ,rr

```
. mlogit  vote retnat class union southwest urban lrself own, rr
Multinomial logistic regression                 Number of obs   =        785
                                                LR chi2(14)     =     233.16
                                                Prob > chi2     =     0.0000
Log likelihood = -736.13494                     Pseudo R2       =     0.1367
------------------------------------------------------------------------------
        vote |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
labour       |
      retnat |   .3957868   .0476164    -7.70   0.000     .3126475    .5010347
       class |    1.52707   .1505479     4.29   0.000     1.258758    1.852574
       union |   2.077248   .4773007     3.18   0.001     1.324046    3.258919
   southwest |   .6824648   .2800351    -0.93   0.352     .3053568    1.525292
       urban |   1.734483   .3574771     2.67   0.008     1.158077    2.597783
      lrself |   .7366692   .0330039    -6.82   0.000     .6747414    .8042808
         own |   .4276451   .0984729    -3.69   0.000     .2723198    .6715645
-------------+----------------------------------------------------------------
lib-dem      |
      retnat |   .6128634   .0749032    -4.01   0.000     .4823155    .7787466
       class |   1.086011   .1085745     0.83   0.409     .8927601    1.321093
       union |   1.633423   .3941819     2.03   0.042     1.017852    2.621277
   southwest |   2.569538   .8617552     2.81   0.005     1.331633    4.958217
       urban |   1.551411   .3263948     2.09   0.037      1.02718    2.343191
      lrself |   .6979438   .0331201    -7.58   0.000     .6359569    .7659727
         own |   .5689902   .1393054    -2.30   0.021     .3521323    .9193985
------------------------------------------------------------------------------
(vote==con is the base outcome)
```

Odds Ratio

- The odds of voting Labour over Conservative are twice as high for union identifiers as opposed to non-union identifiers holding other variables constant
- The odds of voting Labour over Conservative are a half as high for those owning their house compared to those that do not own