Intermediate Social Statistics Hilary 2009 Lecture 6: Multinomial Logit

Raymond Duch

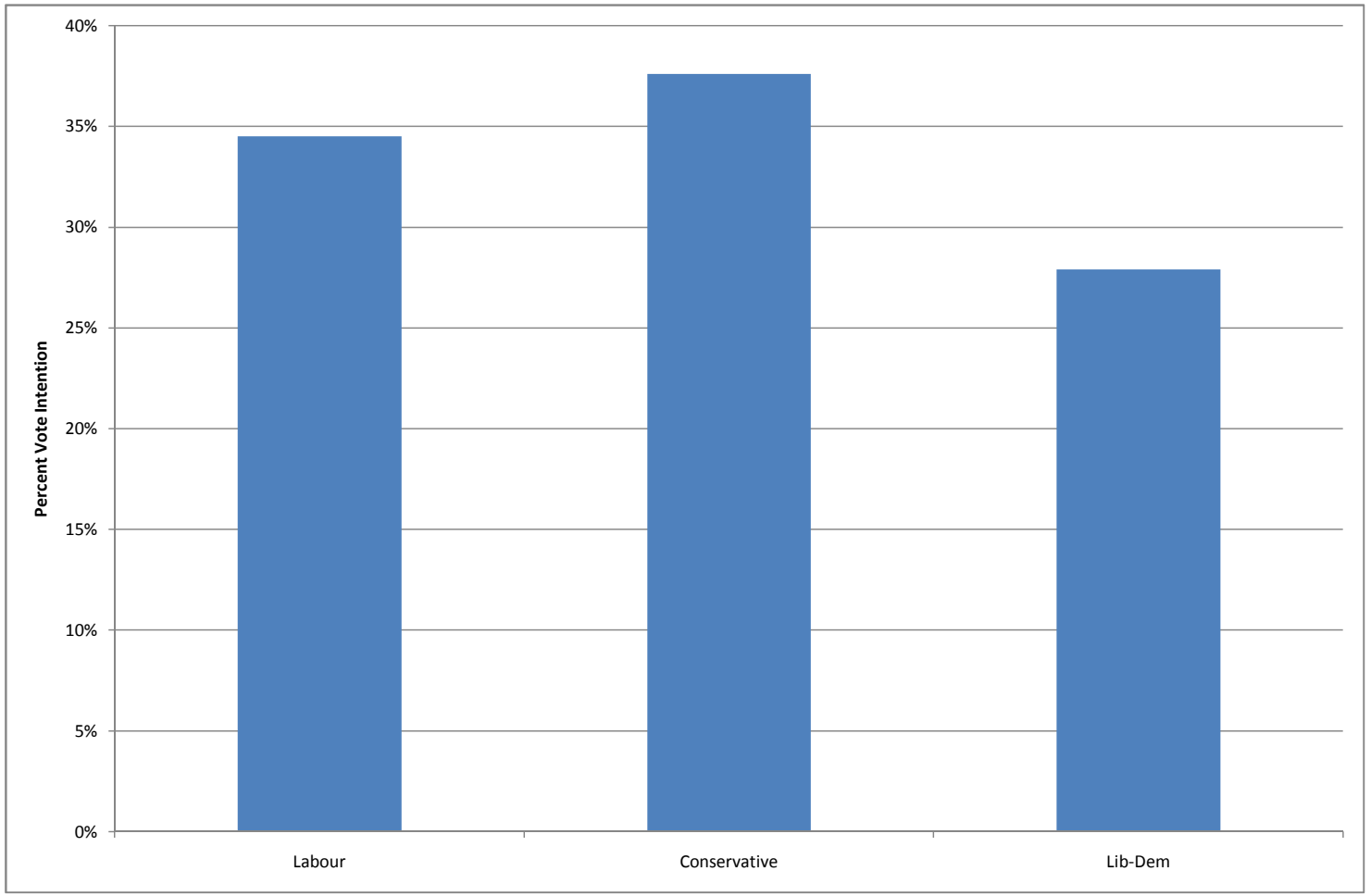
Nuffield College

February 23, 2010

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



What are you trying to learn from the data?

```
. tabulate vote retnat, column
+--------------------+
| Key                |
|--------------------|
|     frequency      |
| column percentage  |
+--------------------+
           |                retnat
      vote |    better       same      worse |     Total
-----------+---------------------------------+----------
    labour |       145        122         63 |       330
           |     47.70      39.87      18.86 |     34.96
-----------+---------------------------------+----------
       con |        72         98        182 |       352
           |     23.68      32.03      54.49 |     37.29
-----------+---------------------------------+----------
   lib-dem |        87         86         89 |       262
           |     28.62      28.10      26.65 |     27.75
-----------+---------------------------------+----------
     Total |       304        306        334 |       944
           |    100.00     100.00     100.00 |    100.00
```

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 Labour; $y = 2$ if i chooses Conservative
- we observe a vector of characteristics x_i for decision maker i (each respondent in the survey)
  - $x_{i,1}$=income
  - $x_{i,2}$=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:

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

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

- 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

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

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

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

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

Mulitinomial Logistic Probabilities

$$Prob(Y_i = 1) = \frac{1}{1 + \sum_{k=2}^{J} \exp^{(\mathbf{x_i}[\beta'_\mathbf{k}])}} \tag{7}$$

$$Prob(Y_i = j) = \frac{\exp^{(\mathbf{x_i}[\beta_\mathbf{j}{}'])}}{1 + \sum_{k=2}^{J} \exp^{(\mathbf{x_i}[\beta'_\mathbf{k}])}} \tag{8}$$

Estimation

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

$$LnL = \sum_{i=1}^{n} \sum_{j=1}^{J} d_{ij} ln P_{ij} \tag{9}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{J} d_{ij} ln \frac{\exp^{(\mathbf{x_i}[\beta_\mathbf{j}'])}}{1 + \sum_{k=2}^{J} \exp^{(\mathbf{x_i}[\beta_\mathbf{k}'])}} \tag{10}$$

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 ($j = 1$)
- The probability of this person voting Labour can be calculated from the earlier equation:

$$Prob(Y_i = j) = \frac{\exp^{(\mathbf{x_i}[\beta_\mathbf{j}'])}}{1 + \sum_{k=2}^{J} \exp^{(\mathbf{x_i}[\beta_\mathbf{k}'])}} \tag{11}$$

(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
```

Back to Multinomial Logit Model Results

```
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)
```

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

$$\frac{\Delta Pr(y = j|x)}{\Delta_{x_k}} = Pr(y = j|x, x_k = x_E) - Pr(y = j|x, x_k = x_S) \quad (12)$$

where x_E is an hypothesized high value on x_k and x_S is an hypothesized low value on x_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
```

Relative Risk and Relative Risk Ratios

The Relative Risk: the relative probability of $Y = 2$ to base category

$$\frac{Pr(y = 2|X)}{Pr(y = 1|X)} = \exp^{(X\beta^2)} \tag{13}$$

The Relative Risk Ratio: relative-risk ratio for a one-unit change in x_k

$$\exp^{(\beta_k^2)} \tag{14}$$

Risk of the outcome relative to the base outcome

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 Multinomial Logit use ,rrr

```
. 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 relative risk of voting Labour over Conservative is twice as high for union identifiers as opposed to non-union identifiers holding other variables constant
- The relative risk of voting Labour over Conservative is about .5 for those owning their house compared to those that do not own

Generating Substantive Effects for Whole Sample

- Calculate baseline probabilities for all individuals in sample – get mean probabilities
- Define plausible changes in specific independent variable(s)
- Change value of variable(s) for all observations in data set
- Calculate new mean probabilities for all observations in sample
- Calculate the difference in the two sets of mean probabilities

Generating Substantive Effects for Whole Sample

$$\frac{\Delta Pr(y = j|x)}{\Delta_{x_k}} = Pr(y = j|x) - Pr(y = j|x, x_k = x_k + \delta) \tag{15}$$

where δ is the substantive change in x_k whose effect on the probabilities you want to estimate.

Generating Predictions for Base Case x_k

```
. predict mnl_lab, outcome(1)
(option pr assumed; predicted probability)
(339 missing values generated)

. predict mnl_lib, outcome(3)
(option pr assumed; predicted probability)
(339 missing values generated)

. predict mnl_con, outcome(2)
(option pr assumed; predicted probability)
(339 missing values generated)

.
. summ mnl_lab mnl_lib mnl_con

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     mnl_lab |      1161    .3452281    .1802095    .0194829   .7878072
     mnl_lib |      1161    .2702638    .1129102    .0407086   .7474253
     mnl_con |      1161    .3845081    .2234257    .0242929   .9398085
```

Generating Effects of $x_k + 1$

```
. /* change everyones economic perception to whatever it was minus 1 */
.
. gen retnat_old=retnat
(46 missing values generated)

. replace retnat=retnat+1
(1454 real changes made)

. replace retnat=3 if retnat>3
(606 real changes made)

.
. predict mnl_lab_real1, outcome(1)
(option pr assumed; predicted probability)
(314 missing values generated)

. predict mnl_lib_real1, outcome(3)
(option pr assumed; predicted probability)
(314 missing values generated)

. predict mnl_con_real1, outcome(2)
(option pr assumed; predicted probability)
(314 missing values generated)

.
```

Generating Effects of $x_k + 1$

```
.
. gen mnl_difflab=mnl_lab_real1-mnl_lab
(339 missing values generated)

. summarize mnl_difflab if vote~=.

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
 mnl_difflab |       785   -.0906372    .0708667  -.1992379          0

.
. gen mnl_difflib=mnl_lib_real1-mnl_lib
(339 missing values generated)

. summarize mnl_difflib if vote~=.

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
 mnl_difflib |       785    .0008764    .0297978  -.0886205   .0886655

.
. gen mnl_diffcon=mnl_con_real1-mnl_con
(339 missing values generated)

. summarize mnl_diffcon if vote~=.

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
 mnl_diffcon |       785    .0897608    .0729932          0   .2004266

.
```

```
. list vote  mnl_lab- mnl_con mnl_lab_real1- mnl_diffcon in 1/25
```

	vote	mnl_lab	mnl_lib	mnl_con	mnl_la~1	mnl_li~1	mnl_co~1	mnl_di~ab	mnl_di~ib	mnl_
1.	labour	.4190372	.2773368	.3036259	.4190372	.2773368	.3036259	0	0	
2.	.	.5853969	.2036901	.210913	.408312	.2199956	.3716924	-.1770849	.0163055	.160
3.	labour	.3392915	.448593	.2121155	.2161289	.4424811	.34139	-.1231626	-.0061119	.129
4.	.	.5833635	.24934	.1672965	.4190372	.2773368	.3036259	-.1643263	.0279969	.136
5.	.	.5853969	.2036901	.210913	.408312	.2199956	.3716924	-.1770849	.0163055	.160
6.	lib-dem	.5229824	.2006484	.2763692	.3413817	.2028108	.4558075	-.1816007	.0021625	.179
7.	.	.6015419	.2713755	.1270826	.4479602	.3129295	.2391103	-.1535817	.041554	.112
8.	lib-dem	.727296	.2118915	.0608125	.6015419	.2713755	.1270826	-.1257541	.059484	.066
9.	con	.4713021	.2963524	.2323454	.3106309	.3024518	.3869173	-.1606712	.0060993	.154
10.	con	.4190372	.2773368	.3036259	.4190372	.2773368	.3036259	0	0	
11.	labour	.357278	.2271555	.4155665	.2031146	.1999681	.5969173	-.1541633	-.0271875	.181
12.	con	.2035572	.160901	.6355419	.2035572	.160901	.6355419	0	0	
13.	lib-dem	.7813636	.16068	.0579564	.6640832	.2114626	.1244541	-.1172804	.0507826	.066
14.	.	.5006336	.3171744	.182192	.344766	.3382244	.3170096	-.1558676	.02105	.134
15.	.	.3147	.2924995	.3928005	.3147	.2924995	.3928005	0	0	
16.	labour	.	.	.	.	.	.	.	.	
17.	.	.1460564	.0708887	.7830549	.06537	.049129	.885501	-.0806863	-.0217598	.102
18.	.	.	.	.	.	.	.	.	.	
19.	con	.1890971	.0999566	.7109463	.1890971	.0999566	.7109463	0	0	
20.	.	.4465703	.2539591	.2994705	.2797247	.2463242	.4739512	-.1668457	-.007635	.174
21.	lib-dem	.1967002	.3815351	.4217647	.1967002	.3815351	.4217647	0	0	
22.	lib-dem	.5329453	.2755689	.1914859	.3692123	.2956147	.3351729	-.1637329	.0200459	.143
23.	lib-dem	.5059353	.32795	.1661147	.5059353	.32795	.1661147	0	0	
24.	labour	.3987054	.2760825	.3252121	.3987054	.2760825	.3252121	0	0	
25.	lib-dem	.1981778	.3201742	.4816481	.1981778	.3201742	.4816481	0	0	