Control for past alliances and conflicts

Re-analyzing Social Network Studies

Presentation 02/12/2015

Arno Fontaine

The six structural balance hypotheses with logistic regression - Maoz & al.

	+ +	- -	+ -
MID (Militarized Interstate Dispute)	↓ 5	↓ RH2	↑ RH4
Alliance	↑ 6	↑ RH1	↓ RH3

Names of the different variables in the data

```
> names(data_seminar)
 [1] "X"             "year"          "statea"        "stateb"        "abb_a"         "abb_b"         "cap_a"
 [8] "cap_b"         "lncaprat"      "allies"        "polity_a"      "polity_b"      "majpow_a"      "majpow_b"
[15] "trade"         "contig"        "distance"      "numstate"      "cgdpa"         "cgdpb"         "opena"
[22] "openb"         "total_ig"      "igos_i"        "rgdp96pca"     "rgdp96pcb"     "nextMID"       "mzmidnm"
[29] "mzkeynum"      "mzfatald1"     "mzcowwar1"     "logtrade"      "logdistance"   "diffmajp"      "summajp"
[36] "twomajpowers"  "onemajpowers"  "mingov"        "maxgov"        "meangov"       "nextAllies"    "mzmid1_NA"
[43] "MID"           "MID_NA"        "ff"            "ee"            "ef"            "dega_all"      "degb_all"
[50] "dega_mid"      "degb_mid"
```

Summary of the data

```
> summary(data_seminar)
       X                year          statea          stateb          abb_a           abb_b           cap_a
 Min.   :     1   Min.   :1885   Min.   :  2.0   Min.   : 20.0   USA    : 10750   THI    :  9945   Min.   :0.000
 1st Qu.:163240   1st Qu.:1965   1st Qu.:100.0   1st Qu.:420.0   MEX    :  9955   NEW    :  9180   1st Qu.:0.000
 Median :326480   Median :1981   Median :235.0   Median :616.0   GUA    :  9815   JPN    :  9118   Median :0.001
 Mean   :326480   Mean   :1975   Mean   :289.6   Mean   :579.7   CUB    :  9813   CHN    :  9117   Mean   :0.009
 3rd Qu.:489720   3rd Qu.:1993   3rd Qu.:435.0   3rd Qu.:740.0   DOM    :  9652   AUL    :  9069   3rd Qu.:0.005
 Max.   :652959   Max.   :2002   Max.   :987.0   Max.   :990.0   SAL    :  9593   NEP    :  8517   Max.   :0.384
                                                                 (Other):593381   (Other):598013   NA's   :18336
     cap_b          lncaprat          allies          polity_a        polity_b        majpow_a
 Min.   :0.000   Min.   : 0.0000   Min.   :0.00000   Min.   :-10.00   Min.   :-10.00   Min.   :0.00000
 1st Qu.:0.000   1st Qu.: 0.8414   1st Qu.:0.00000   1st Qu.: -7.00   1st Qu.: -7.00   1st Qu.:0.00000
 Median :0.001   Median : 1.9424   Median :0.00000   Median :  1.00   Median : -4.00   Median :0.00000
 Mean   :0.008   Mean   : 2.3717   Mean   :0.06372   Mean   :  1.06   Mean   : -1.07   Mean   :0.05744
 3rd Qu.:0.005   3rd Qu.: 3.4363   3rd Qu.:0.00000   3rd Qu.:  9.00   3rd Qu.:  7.00   3rd Qu.:0.00000
 Max.   :0.216   Max.   :11.9621   Max.   :1.00000   Max.   : 10.00   Max.   : 10.00   Max.   :1.00000
 NA's   :18336   NA's   :2072                        NA's   :80689    NA's   :68695
    majpow_b           trade              contig          distance        numstate          cgdpa
 Min.   :0.00000   Min.   :     0.0   Min.   :1.000   Min.   :    5   Min.   : 36.0   Min.   :     16
 1st Qu.:0.00000   1st Qu.:     0.0   1st Qu.:6.000   1st Qu.: 2589   1st Qu.:125.0   1st Qu.:   2033
 Median :0.00000   Median :     0.1   Median :6.000   Median : 4612   Median :159.0   Median :   9705
 Mean   :0.03993   Mean   :   147.4   Mean   :5.846   Mean   : 4790   Mean   :144.4   Mean   : 126839
 3rd Qu.:0.00000   3rd Qu.:     7.0   3rd Qu.:6.000   3rd Qu.: 6683   3rd Qu.:186.0   3rd Qu.:  50391
 Max.   :1.00000   Max.   :399075.0   Max.   :6.000   Max.   :12347   Max.   :192.0   Max.   :9810201
                   NA's   :92301                                                      NA's   :92052
     cgdpb             opena            openb          total_ig         igos_i        rgdp96pca       rgdp96pcb
 Min.   :     16   Min.   :0.00   Min.   :0.00   Min.   :  0.00   Min.   :  0.00   Min.   :  220   Min.   :  220
 1st Qu.:   2534   1st Qu.:0.12   1st Qu.:0.09   1st Qu.: 14.00   1st Qu.: 12.00   1st Qu.: 1946   1st Qu.: 1176
 Median :   9416   Median :0.21   Median :0.18   Median : 20.00   Median : 19.00   Median : 4315   Median : 2521
 Mean   :  87389   Mean   :0.32   Mean   :0.30   Mean   : 21.53   Mean   : 20.08   Mean   : 6525   Mean   : 5151
 3rd Qu.:  38326   3rd Qu.:0.38   3rd Qu.:0.34   3rd Qu.: 28.00   3rd Qu.: 27.00   3rd Qu.: 8633   3rd Qu.: 6083
 Max.   :4838492   Max.   :7.69   Max.   :7.69   Max.   :107.00   Max.   :107.00   Max.   :46065   Max.   :46065
 NA's   :97580     NA's   :92307  NA's   :98595  NA's   :164312   NA's   :3271     NA's   :79041   NA's   :83994
```

```
   nextMID          mzmidnm           mzkeynum          mzfatald1       mzcowwar1        logtrade      logdistance
Min.   :0       Min.   :   0.00   Min.   :   0.00   Min.   :0       Min.   :0       Min.   : 0.00   Min.   :1.609
1st Qu.:0       1st Qu.:   0.00   1st Qu.:   0.00   1st Qu.:0       1st Qu.:0       1st Qu.: 0.00   1st Qu.:7.859
Median :0       Median :   0.00   Median :   0.00   Median :0       Median :0       Median : 0.12   Median :8.436
Mean   :0       Mean   :  10.31   Mean   :  10.26   Mean   :0       Mean   :0       Mean   : 1.28   Mean   :8.233
3rd Qu.:0       3rd Qu.:   0.00   3rd Qu.:   0.00   3rd Qu.:0       3rd Qu.:0       3rd Qu.: 2.08   3rd Qu.:8.807
Max.   :1       Max.   :4343.00   Max.   :4343.00   Max.   :6       Max.   :1       Max.   :12.90   Max.   :9.421
NA's   :39706   NA's   :39706     NA's   :39706     NA's   :40048   NA's   :39706   NA's   :92301
   diffmajp          summajp         twomajpowers       onemajpowers      mingov          maxgov
Min.   :0.00000   Min.   :0.00000   Min.   :0.000000   Min.   :0.00000   Mode:logical   Mode:logical
1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0.00000   NA's:652959    NA's:652959
Median :0.00000   Median :0.00000   Median :0.000000   Median :0.00000
Mean   :0.09135   Mean   :0.09736   Mean   :0.003003   Mean   :0.09436
3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:0.00000
Max.   :1.00000   Max.   :2.00000   Max.   :1.000000   Max.   :1.00000
   meangov          nextAllies        mzmid1_NA           MID             MID_NA            ff
Min.   :-10.00   Min.   :0.00000   Min.   :0.000000   Min.   :0.00000   Min.   :0.000000   Min.   : 0.000
1st Qu.: -5.00   1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.: 0.000
Median :  0.50   Median :0.00000   Median :0.000000   Median :0.00000   Median :0.000000   Median : 0.000
Mean   : -0.08   Mean   :0.06331   Mean   :0.004092   Mean   :0.00409   Mean   :0.004091   Mean   : 1.287
3rd Qu.:  3.50   3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.: 0.000
Max.   : 10.00   Max.   :1.00000   Max.   :1.000000   Max.   :1.00000   Max.   :1.000000   Max.   :52.000
NA's   :124245   NA's   :200                          NA's   :200
      ee                ef              dega_all          degb_all          dega_mid          degb_mid
Min.   : 0.00000   Min.   : 0.00000   Min.   : 0.000   Min.   : 0.000   Min.   : 0.0000   Min.   : 0.0000
1st Qu.: 0.00000   1st Qu.: 0.00000   1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.: 0.0000   1st Qu.: 0.0000
Median : 0.00000   Median : 0.00000   Median : 0.000   Median : 0.000   Median : 0.0000   Median : 0.0000
Mean   : 0.01104   Mean   : 0.06941   Mean   : 3.151   Mean   : 2.711   Mean   : 0.3036   Mean   : 0.4296
3rd Qu.: 0.00000   3rd Qu.: 0.00000   3rd Qu.: 0.000   3rd Qu.: 0.000   3rd Qu.: 0.0000   3rd Qu.: 0.0000
Max.   :12.00000   Max.   :19.00000   Max.   :56.000   Max.   :49.000   Max.   :39.0000   Max.   :39.0000
```

Creation of different models in RStudio

- dyadic model : states a & b : relation in the next year

- basic model : next Allies, next MID, depends on ee, ff, ef

- degree model : deg_all & deg_MID

Model 1 : probability of next allies in t+1 when two countries are allies in t

```
> summary(model1)

Call:
glm(formula = nextAllies ~ allies, family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4797  -0.0733  -0.0733  -0.0733   3.4411

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.91796    0.02472  -239.4   <2e-16 ***
allies       8.94502    0.03401   263.0   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 308064  on 652758  degrees of freedom
Residual deviance:  38275  on 652757  degrees of freedom
  (200 observations deleted due to missingness)
AIC: 38279

Number of Fisher Scoring iterations: 8
```

Model 2 : nextMID when MID

```
> summary(model2)

Call:
glm(formula = nextMID ~ MID, family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8829  -0.0820  -0.0820  -0.0820   3.3754

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.69314    0.02211 -257.46   <2e-16 ***
MID          4.95197    0.05358   92.43   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 34380  on 613053  degrees of freedom
Residual deviance: 29892  on 613052  degrees of freedom
  (39905 observations deleted due to missingness)
AIC: 29896

Number of Fisher Scoring iterations: 8
```

Model 3 : next Alliance, just with ff

```
> summary(model3)

Call:
glm(formula = nextAllies ~ ff, family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-6.1584  -0.1408  -0.1408  -0.1408   3.0394

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.60918    0.01251  -368.4   <2e-16 ***
ff           0.48107    0.00215   223.7   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 308064  on 652758  degrees of freedom
Residual deviance:  83868  on 652757  degrees of freedom
  (200 observations deleted due to missingness)
AIC: 83872

Number of Fisher Scoring iterations: 7
```

Model 4 : next Alliance with ff, ef, ee

```
> summary(model4)

Call:
glm(formula = nextAllies ~ ff + ef + ee, family = "binomial",
    data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-6.3650  -0.1387  -0.1387  -0.1387   3.0491

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.638970   0.012712 -364.93   <2e-16 ***
ff           0.477266   0.002148  222.20   <2e-16 ***
ef           0.173825   0.015070   11.54   <2e-16 ***
ee           0.667969   0.039827   16.77   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 308064  on 652758  degrees of freedom
Residual deviance:  83446  on 652755  degrees of freedom
  (200 observations deleted due to missingness)
AIC: 83454

Number of Fisher Scoring iterations: 7
```

Model 5 : next MID

```
> summary(model5)

Call:
glm(formula = nextMID ~ ff + ef + ee, family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.6006  -0.0853  -0.0853  -0.0853   3.3518

Coefficients:
             Estimate Std. Error  z value Pr(>|z|)
(Intercept) -5.613665   0.021564 -260.324  < 2e-16 ***
ff           0.021093   0.002713    7.774 7.61e-15 ***
ef           0.595638   0.015938   37.372  < 2e-16 ***
ee           0.686466   0.046283   14.832  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 34382  on 613252  degrees of freedom
Residual deviance: 32735  on 613249  degrees of freedom
  (39706 observations deleted due to missingness)
AIC: 32743

Number of Fisher Scoring iterations: 8
```

Model 6 : MID depends on ee and ef

```
> summary(model6)

Call:
glm(formula = nextMID ~ ef + ee, family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.6981  -0.0871  -0.0871  -0.0871   3.3399

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.57363    0.02058 -270.78   <2e-16 ***
ef           0.61221    0.01606   38.12   <2e-16 ***
ee           0.69529    0.04667   14.90   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 34382  on 613252  degrees of freedom
Residual deviance: 32787  on 613250  degrees of freedom
  (39706 observations deleted due to missingness)
AIC: 32793

Number of Fisher Scoring iterations: 8
```

Model 7 : degree model

```
> summary(model7)

Call:
glm(formula = nextAllies ~ ff + ee + ef + dega_all + degb_all,
    family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-6.2763  -0.1078  -0.1078  -0.1078   3.2100

Coefficients:
             Estimate Std. Error  z value Pr(>|z|)
(Intercept) -5.146113   0.016382 -314.136  < 2e-16 ***
ff           0.390209   0.002213  176.351  < 2e-16 ***
ee           0.452239   0.039564   11.430  < 2e-16 ***
ef           0.082522   0.017696    4.663 3.11e-06 ***
dega_all     0.049858   0.001201   41.527  < 2e-16 ***
degb_all     0.042703   0.001680   25.416  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 308064  on 652758  degrees of freedom
Residual deviance:  73639  on 652753  degrees of freedom
  (200 observations deleted due to missingness)
AIC: 73651

Number of Fisher Scoring iterations: 8
```

Model 8 : deg_allies

```
> summary(model8)

Call:
glm(formula = nextAllies ~ dega_all + degb_all, family = "binomial",
    data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.8929  -0.1576  -0.1576  -0.1576   2.9648

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.382477   0.011125 -393.95   <2e-16 ***
dega_all     0.032020   0.000833   38.44   <2e-16 ***
degb_all     0.139890   0.001020  137.12   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 308064  on 652758  degrees of freedom
Residual deviance: 154057  on 652756  degrees of freedom
  (200 observations deleted due to missingness)
AIC: 154063

Number of Fisher Scoring iterations: 7
```

Model 9 : global compilation

```
> summary(model9)

Call:
glm(formula = nextAllies ~ allies + ef + ff + ee + dega_mid +
    degb_mid + dega_all + degb_all, family = "binomial", data = data_seminar)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.6752  -0.0411  -0.0411  -0.0411   4.2167

Coefficients:
             Estimate Std. Error  z value Pr(>|z|)
(Intercept) -7.075009   0.042856 -165.088  < 2e-16 ***
allies       9.588219   0.070059  136.860  < 2e-16 ***
ef          -0.142510   0.028025   -5.085 3.67e-07 ***
ff          -0.133163   0.003203  -41.573  < 2e-16 ***
ee          -0.754564   0.049615  -15.208  < 2e-16 ***
dega_mid     0.015996   0.007004    2.284   0.0224 *
degb_mid    -0.008268   0.005572   -1.484   0.1379
dega_all     0.015281   0.002016    7.578 3.50e-14 ***
degb_all     0.125521   0.002764   45.408  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 308064  on 652758  degrees of freedom
Residual deviance:  31505  on 652750  degrees of freedom
  (200 observations deleted due to missingness)
AIC: 31523

Number of Fisher Scoring iterations: 9
```