

MODELS USED IN FRONTIER'S PAPER

Version 1.0 - August 2021

Claude Grasland, Université de Paris

Model I : National effect / What =f(who-what)

Our first model will test the existence of comparable levels of interest for geopolitical topics in the different media of different countries. It is not really a research hypothesis but rather a control on the joint choice of media and topics linked to the choice of keywords. We assume therefore H1 :

- **H1 : The global level of interest of media for a geopolitical topic is not related to the country where the media is located.**

Migrants

In the case of the migrant topic, H1 is not fully verified and we can observe a national effect but with a low level of significance ($p = 0.047$). This effect is mainly related to an over-representation of the migrant topic in german outlets, with an average value of 2.67% of foreign news which is significantly higher than the levels observed in media outlets from Spain (1.66%), Italy and United Kingdom but not significantly different from the level observed in France.

Borders

```
##
## Call:
## lm(formula = estimate ~ country, data = dt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0048100 -0.0014394 -0.0002812  0.0016375  0.0052425
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0106625  0.0013585   7.849 1.09e-06 ***
## countryESP   0.0011875  0.0019212   0.618  0.5458
## countryFRA   0.0000575  0.0019212   0.030  0.9765
## countryGBR  -0.0018600  0.0019212  -0.968  0.3483
## countryITA  -0.0036950  0.0019212  -1.923  0.0736 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002717 on 15 degrees of freedom
## Multiple R-squared:  0.3485, Adjusted R-squared:  0.1748
## F-statistic: 2.006 on 4 and 15 DF,  p-value: 0.1453
```

In the case of the border topic, H1 is verified and no significant differences are observed between outlets according to their country of location ($p = 0.143$). The global level of interest is twice lower as in the case of migrants with only 1% of foreign news in average.

Pandemics (1)

```
##
## Call:
## lm(formula = estimate ~ country, data = dt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.018443 -0.002874  0.000170  0.003458  0.022738
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.005815   0.004463   1.303  0.21224
## countryESP   0.018638   0.006312   2.953  0.00988 **
## countryFRA   0.012082   0.006312   1.914  0.07485 .
## countryGBR   0.009185   0.006312   1.455  0.16621
## countryITA   0.011287   0.006312   1.788  0.09393 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.008926 on 15 degrees of freedom
## Multiple R-squared:  0.3772, Adjusted R-squared:  0.2112
## F-statistic: 2.272 on 4 and 15 DF,  p-value: 0.1098
```

In the case of pandemics, no significant differences can be observed for the whole sample of media, but significant differences could be observed between the samples of german and spanish newspapers when isolated.

Pandemics (2)

```
##
## Call:
## lm(formula = estimate ~ country, data = dt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.955e-03 -5.100e-04  5.875e-05  8.588e-04  1.582e-03
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0025850  0.0006611   3.910  0.00139 **
## countryESP   0.0019800  0.0009350   2.118  0.05132 .
## countryFRA   0.0023725  0.0009350   2.537  0.02276 *
## countryGBR   0.0019875  0.0009350   2.126  0.05055 .
## countryITA   0.0005525  0.0009350   0.591  0.56337
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.001322 on 15 degrees of freedom
## Multiple R-squared:  0.3963, Adjusted R-squared:  0.2353
## F-statistic: 2.461 on 4 and 15 DF,  p-value: 0.09024
```

The fact to exclude the first semester of 2020 does not really change the results.

synthesis

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
## % Date and time: Lun, août 23, 2021 - 17:55:19
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccc}
## \hline
## \hline
## & \multicolumn{4}{c}{\textit{Dependent variable:}} \\
## \cline{2-5}
## \hline
## & \multicolumn{4}{c}{estimate} \\
## \hline
## & (1) & (2) & (3) & (4) \\
## \hline
## countryESP & 0.001 &  $-\$0.010^{***}$  &  $0.019^{***}$  &  $0.002^{*}$  \\
## & (0.002) & (0.003) & (0.006) & (0.001) \\
## & & & & \\
## countryFRA & 0.0001 &  $-\$0.004$  &  $0.012^{*}$  &  $0.002^{**}$  \\
## & (0.002) & (0.003) & (0.006) & (0.001) \\
## & & & & \\
## countryGBR &  $-\$0.002$  &  $-\$0.008^{**}$  & 0.009 &  $0.002^{*}$  \\
## & (0.002) & (0.003) & (0.006) & (0.001) \\
## & & & & \\
## countryITA &  $-\$0.004^{*}$  &  $-\$0.008^{**}$  &  $0.011^{*}$  & 0.001 \\
## & (0.002) & (0.003) & (0.006) & (0.001) \\
## & & & & \\
## Constant &  $0.011^{***}$  &  $0.027^{***}$  & 0.006 &  $0.003^{***}$  \\
## & (0.001) & (0.002) & (0.004) & (0.001) \\
## & & & & \\
## \hline
## Observations & 20 & 20 & 20 & 20 \\
## R2 & 0.349 & 0.455 & 0.377 & 0.396 \\
## Adjusted R2 & 0.175 & 0.309 & 0.211 & 0.235 \\
## Residual Std. Error (df = 15) & 0.003 & 0.005 & 0.009 & 0.001 \\
## F Statistic (df = 4; 15) & 2.006 &  $3.125^{**}$  & 2.272 &  $2.461^{*}$  \\
## \hline
## \hline
## \textit{Note:} & \multicolumn{4}{r}{ $^{*}p < 0.1$ ;  $^{**}p < 0.05$ ;  $^{***}p < 0.01$ } \\
## \end{tabular}
## \end{table}
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Lun, août 23, 2021 - 17:54:46

Model 2 : National effect / What = (When, Who-Where)

Our second hypothesis is related to the temporal dimension of the geopolitical agenda. We assume that, all things being equal with the global level of interest for a geopolitical topic, the distribution of the period of interest is influenced by national agenda which produce significant differences in the timeline of the topic over the period of observation. H2 can therefore be formulated as :

- **H2 : The variation of the salience of a geopolitical topic through time is influenced by national agenda. The correlation of the timeline is higher between media of the same country than between media of different country**

Table 1:

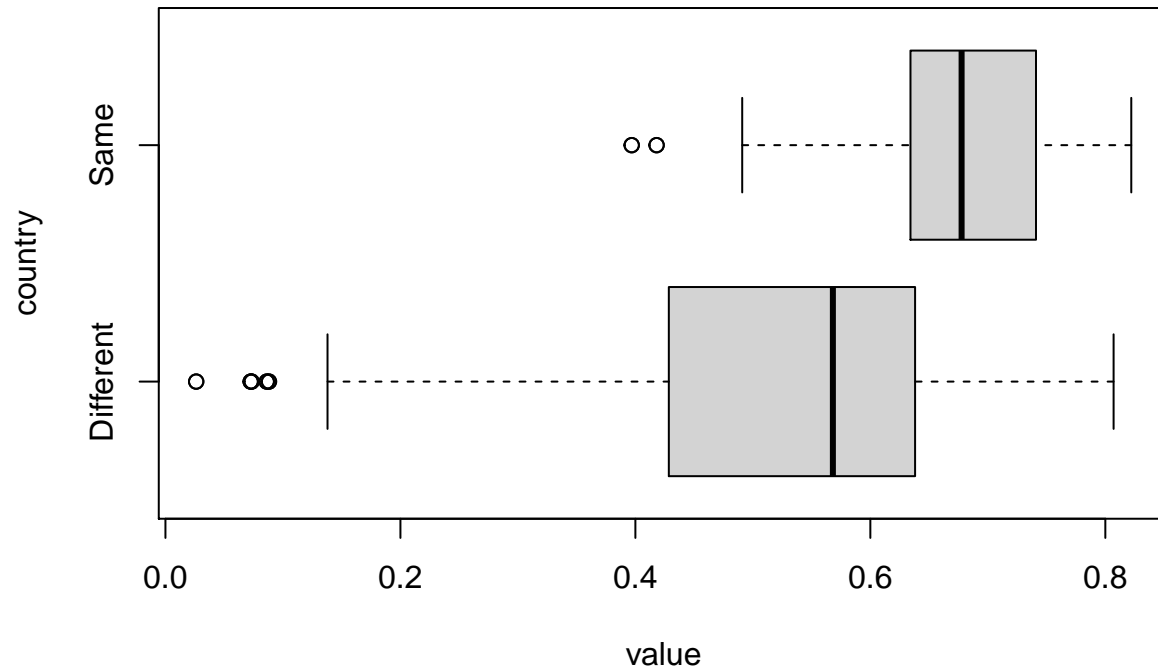
	<i>Dependent variable:</i>			
	estimate			
	(1)	(2)	(3)	(4)
countryESP	0.001 (0.002)	−0.010*** (0.003)	0.019*** (0.006)	0.002* (0.001)
countryFRA	0.0001 (0.002)	−0.004 (0.003)	0.012* (0.006)	0.002** (0.001)
countryGBR	−0.002 (0.002)	−0.008** (0.003)	0.009 (0.006)	0.002* (0.001)
countryITA	−0.004* (0.002)	−0.008** (0.003)	0.011* (0.006)	0.001 (0.001)
Constant	0.011*** (0.001)	0.027*** (0.002)	0.006 (0.004)	0.003*** (0.001)
Observations	20	20	20	20
R ²	0.349	0.455	0.377	0.396
Adjusted R ²	0.175	0.309	0.211	0.235
Residual Std. Error (df = 15)	0.003	0.005	0.009	0.001
F Statistic (df = 4; 15)	2.006	3.125**	2.272	2.461*

Note:

*p<0.1; **p<0.05; ***p<0.01

To verify H2, we build a correlation matrix between the levels of salience of the 20 media outlets (using pairwise complete observation when missing data are present) and we compare the distribution of the level of correlation for media located in the same country and media located in different countries with variance analysis and Fisher test.

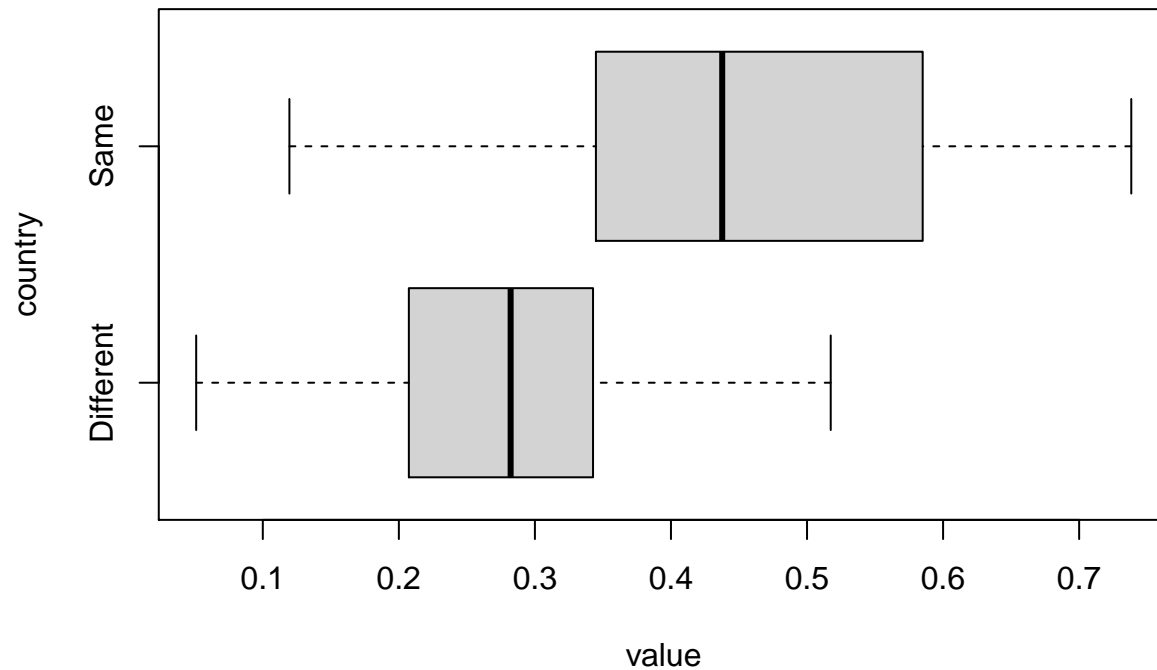
Migrant



```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50073 -0.09126  0.03862  0.10631  0.28014
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.526886   0.008765  60.115 < 2e-16 ***
## countrySame  0.135299   0.022057   6.134 2.16e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1568 on 378 degrees of freedom
## Multiple R-squared:  0.09053,    Adjusted R-squared:  0.08812
## F-statistic: 37.63 on 1 and 378 DF,  p-value: 2.161e-09
```

In the case of migrant topic, the average correlation between media of different countries is equal to +0.53 against +0.66 in the case of media located in the same country. Taking into account the variance observed in the sample, these differences appears strongly significant ($F = 37.6$, $df = (1,378)$, $p < 0.001$)

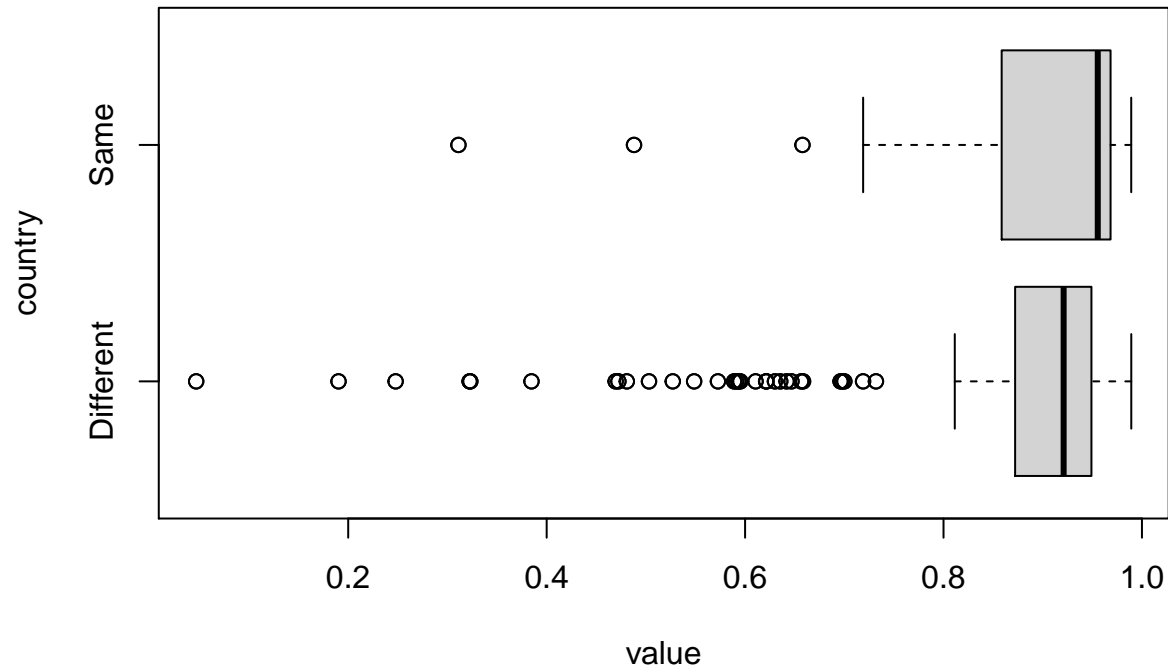
Borders



```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32356 -0.07790  0.00113  0.06977  0.29520
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.281033   0.006459  43.512  <2e-16 ***
## countrySame  0.162095   0.016254   9.973  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1155 on 378 degrees of freedom
## Multiple R-squared:  0.2083, Adjusted R-squared:  0.2062
## F-statistic: 99.45 on 1 and 378 DF, p-value: < 2.2e-16
```

In the case of border, the global level of correlation is lower but the national differences appears much more significant with only +0.28 in the case of media located in different countries against +0.44 for media located in different countries ($F = 99.4$, $df = (1,378)$, $p < 0.001$)

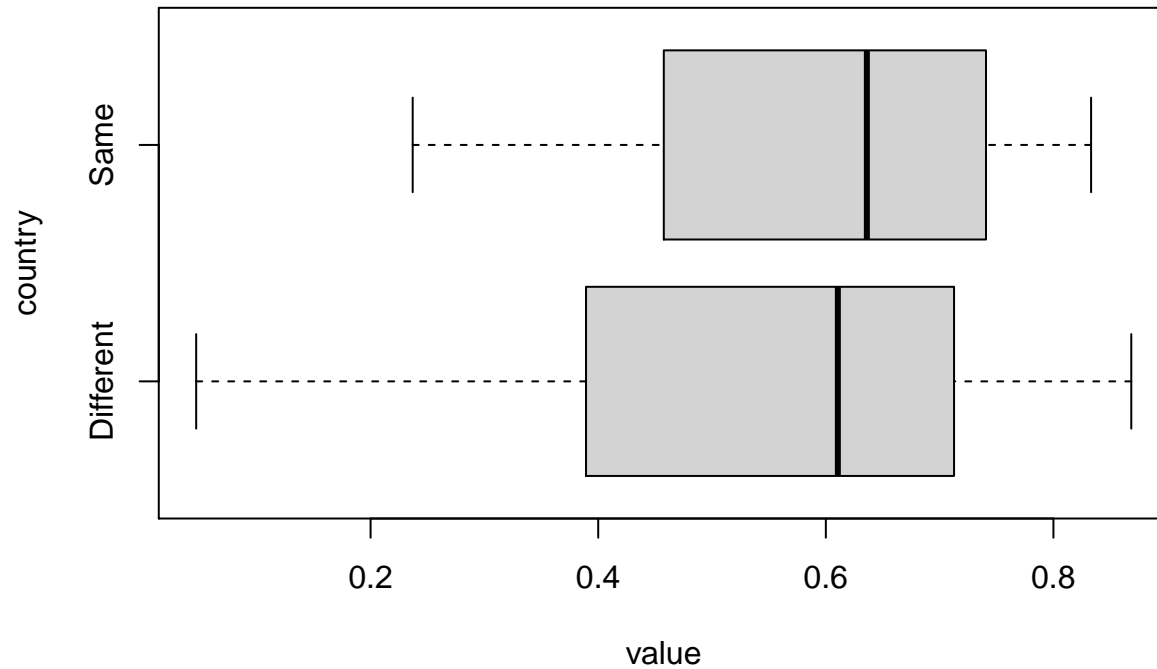
Pandemic (1)



```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.80644  0.01287  0.06872  0.09250  0.13606
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.853193   0.009524  89.585  <2e-16 ***
## countrySame  0.029441   0.023968   1.228    0.22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1704 on 378 degrees of freedom
## Multiple R-squared:  0.003976,    Adjusted R-squared:  0.001341
## F-statistic: 1.509 on 1 and 378 DF,  p-value: 0.2201
```

In the case of pandemic, the correlation is extremely high (0.86) because of the common perturbation observed in spring 2020 which appears in all newspaper as an exceptional jump. As a result, no significant differences can be observed between media located in the same country (+0.88) or in different countries (+0.83) and we can conclude to the existence of a global event which is not organized at national level in terms of geopolitical agenda.

Pandemic (2)



```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50325 -0.15708  0.05591  0.16062  0.31839
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.55000    0.01135  48.445  <2e-16 ***
## countrySame  0.03873    0.02857   1.355    0.176
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2031 on 378 degrees of freedom
## Multiple R-squared:  0.004836,    Adjusted R-squared:  0.002204
## F-statistic: 1.837 on 1 and 378 DF,  p-value: 0.1761
```

If we exclude the first semester of 2020, the conclusions are not modified because the pandemic curve of the countries remains strongly determined by the peaks of Ebola (2014) and Zyka (2017 ?). The correlation is of course lower but not significantly higher for media of the same country (+0.59) as compared to media of different countries (+0.55) because of strong dispersion.

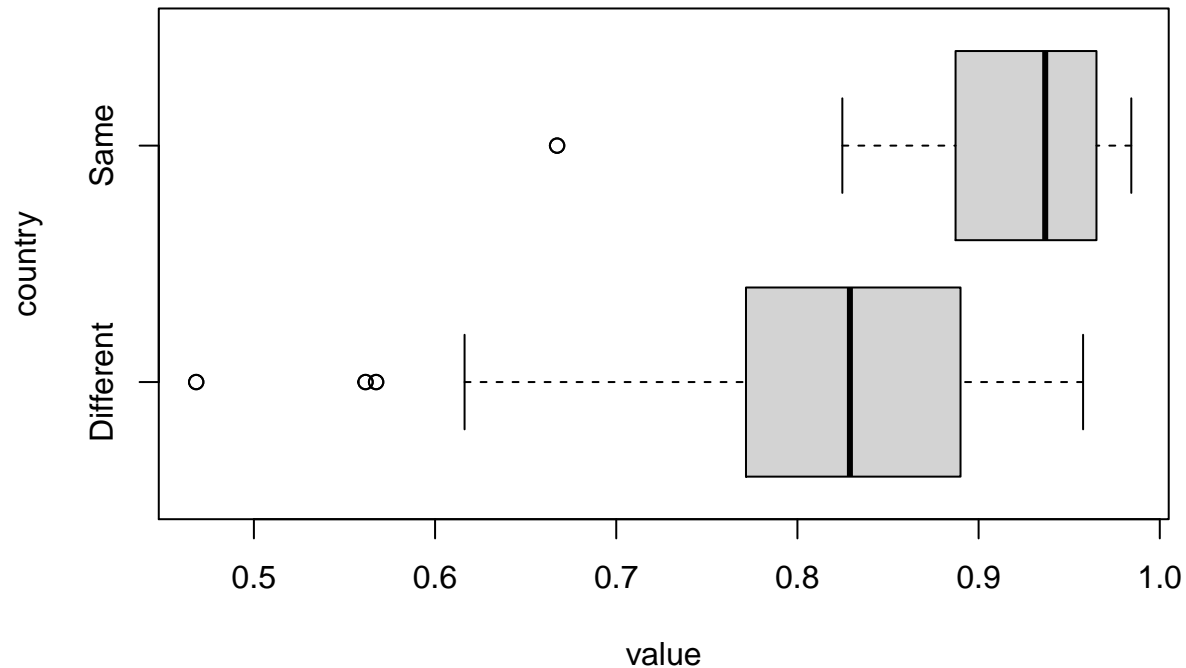
Model 3 : National effect / What = (Where, Who-Where)

We turn now to the second dimension of the geopolitical agenda which is related to the choice of foreign countries mentioned in news related to a geopolitical topic. The fact to use a relatively long period is particularly interesting in this case because it can help us to have a more structural view than an event oriented approach. It is clear for example that all media outlets has mentioned Turkey and Greece in relation with the migrant topic in september 2015. But it is not so obvious that media of each countries

has mentionned these countries in relation with migrant *before* or *after* the period of maximal interest. For the same reason, it is not obvious that media of all countries has offered the same coverage of migrant crisis located outside Europe and its neighborhood like the Rohingias crisis or the wall established by D. Trump. It is the reason why it can be interested to introduce our third hypothesis as follow.

- **H3 : Over a sufficiently long period of time the choice of foreign countries associated to geopolitical topic is influenced by national perceptions of newsworthiness. The correlation between the spatial distribution of countrries mentionned in relation with a topic is higher between media of the same country than between media of different country**

Migrant

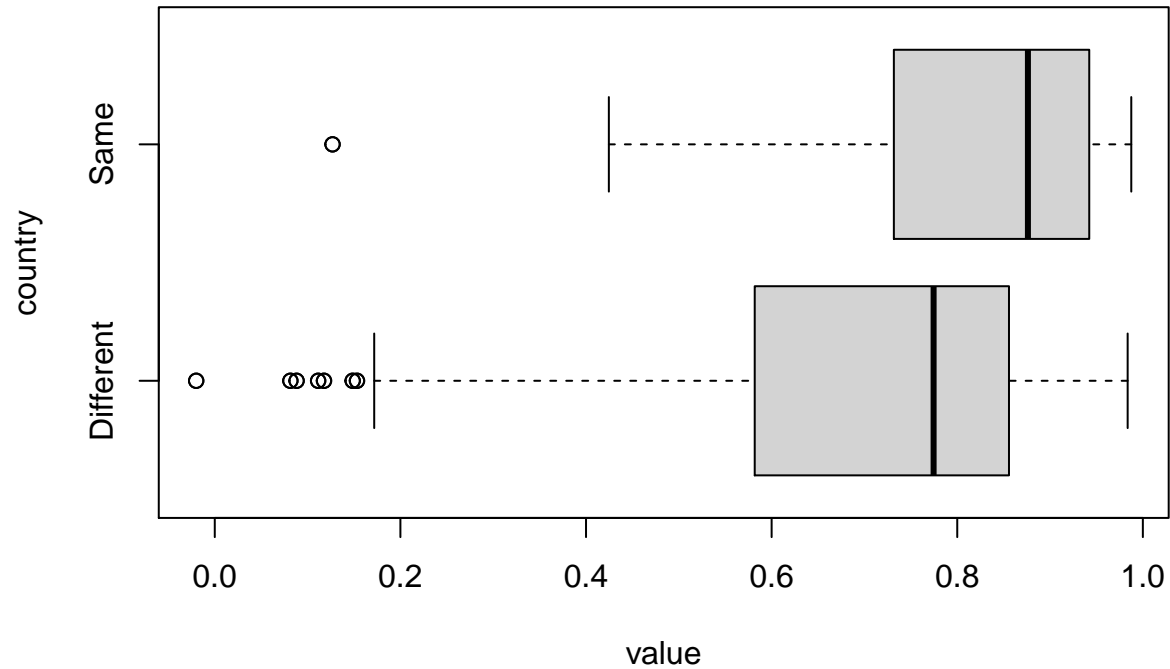


```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.35327 -0.04401  0.00842  0.06061  0.13626
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.821420   0.004682 175.455 < 2e-16 ***
## countrySame  0.093970   0.011782   7.976 1.82e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08375 on 378 degrees of freedom
## Multiple R-squared:  0.144, Adjusted R-squared:  0.1418
## F-statistic: 63.61 on 1 and 378 DF, p-value: 1.82e-14
```

In the case of migrant's topic, we observe a very strong correlation between the geopolitical maps of newspapers. In all media, the countries that has been the most associated to trhe migrant topic in relative terms are the same (Greece, Hungary, Turkey). But some specificities appears and finally the correlation si significantky

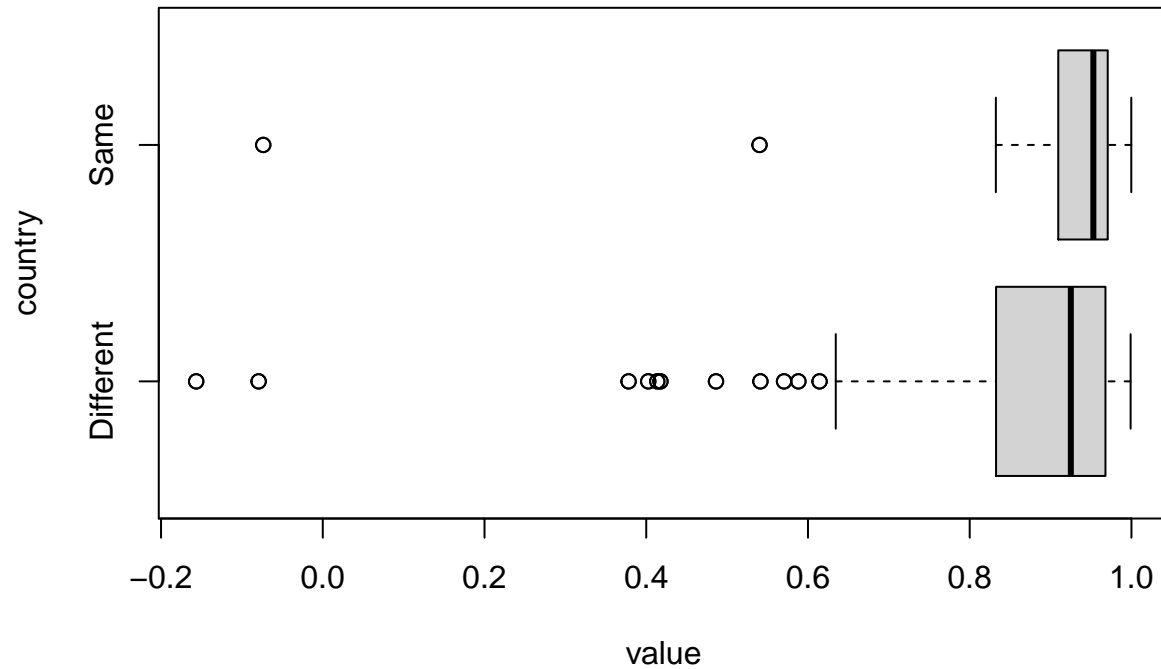
larger between media of the same country (+0.91) and media of different countries ($p < 0.001$)

Borders



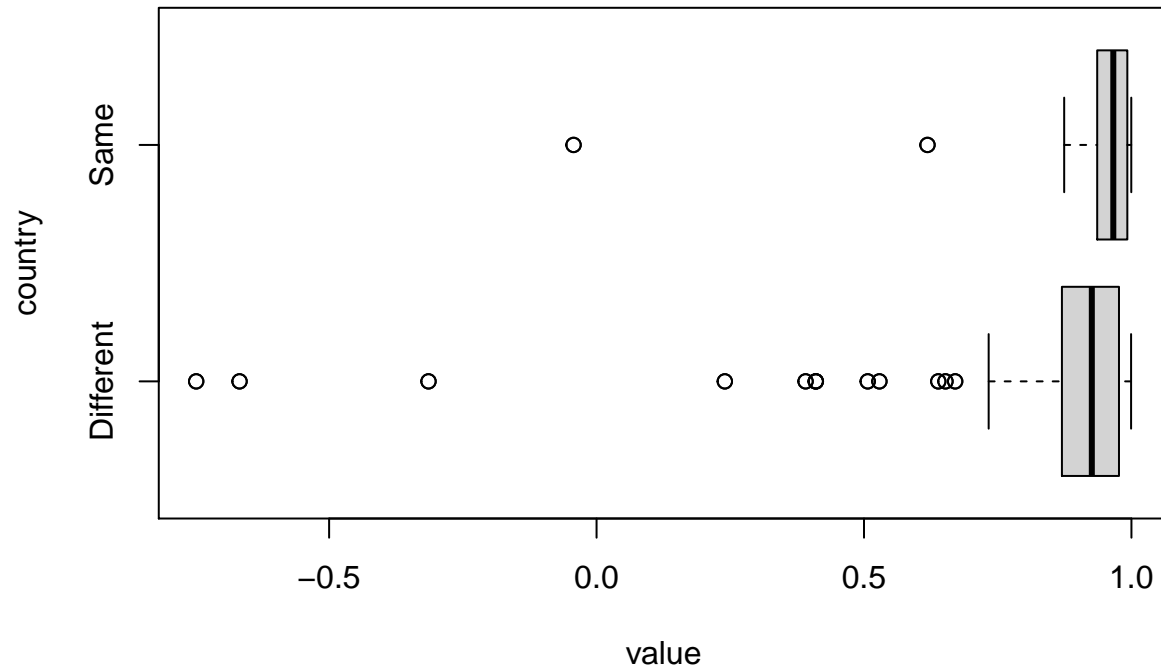
```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.71044 -0.10659  0.08301  0.15886  0.29322
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.69039    0.01292  53.426 < 2e-16 ***
## countrySame  0.10798    0.03252   3.321 0.000986 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2312 on 378 degrees of freedom
## Multiple R-squared:  0.02834,    Adjusted R-squared:  0.02577
## F-statistic: 11.03 on 1 and 378 DF,  p-value: 0.0009859
```

Pandemic (1)



```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.02595 -0.01491  0.05613  0.09530  0.12957
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.869412   0.009669  89.918  <2e-16 ***
## countrySame  0.026584   0.024333   1.092    0.275
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.173 on 378 degrees of freedom
## Multiple R-squared:  0.003148,    Adjusted R-squared:  0.0005104
## F-statistic: 1.194 on 1 and 378 DF,  p-value: 0.2753
```

Pandemic (2)



```
##
## Call:
## lm(formula = value ~ country, data = col)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.62098  0.00303  0.05304  0.09554  0.12691
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.87244    0.01286  67.861  <2e-16 ***
## countrySame  0.04536    0.03235   1.402   0.162
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.23 on 378 degrees of freedom
## Multiple R-squared:  0.005174,    Adjusted R-squared:  0.002542
## F-statistic: 1.966 on 1 and 378 DF,  p-value: 0.1617
```