S2 Text. Details of cluster analysis for identifying poll answering patterns

We used the elbow method as an initial approximation of the appropriate number of clusters and then examined the results of several options to make a final decision. We used the Hartigan-Wong algorithm (PlotHartigan within the FitKMean package in R) to plot a chart of Hartigan numbers. This was partially inconclusive as the score did not drop below 10, but did produced a clear elbow at 6 clusters.

We began with 6 clusters which produced the following mean points:

Table A

```
Cluster means:
       BanS     LabS     LevyS     KidS     SubS
1 8.7057569 9.063966 9.106610 8.989339 8.240938
2 6.8708241 6.565702 6.665924 6.792873 5.982183
3 4.2284946 8.330645 7.486559 6.975806 7.037634
4 4.1452991 4.951567 5.367521 5.045584 5.131054
5 0.7118644 1.107345 1.141243 1.395480 2.039548
6 1.7933884 5.842975 2.512397 4.107438 5.851240
```

With the following number of respondents in each cluster:

```
[1] 469 449 372 351 177 121
```

We then compared this with the results using 5 clusters. The key difference is that the 'neutral' cluster gets divided in the 6-cluster scenario, to give a small cluster who are neutral on all except a levy and a ban – which they are opposed to.

We decided to proceed using 5 clusters, as each of these clusters has an intuitive explanation. The additional 6th cluster did not yield as intuitive an explanation and was also the smallest of the clusters. Separating this group out from the broader 'neutral' group did not add significant nuance to the findings, as it is already clear that the ban was least popular. Finally, adding the 6th cluster only increased the cluster sum-of-squares/ total sum-of squared score by >3%.

For every implementation of the clustering, we used a random seed with a high number of starts (40). This reliably produced the same 5 clusters. These clusters appeared well defined (between cluster sum-of-squares / total sum-of-squares = 65.9 %). They also allowed for an intuitive interpretation in light of the underlying data, as described in the paper.