

Appendix for “Agency proliferation and the globalization of the regulatory state: Introducing a data set on the institutional features of regulatory agencies”

Jacint Jordana, Xavier Fernández-i-Marín and Andrea C. Bianculli

December 2018

Technical documentation of the article “Agency proliferation and the globalization of the regulatory state: Introducing a data set on the institutional features of regulatory agencies” published at *Regulation & Governance*.

Description of the dataset

Number of agencies by country

Figure 1 shows the number of actual institutions (not number of sectors covered) in every country. There are a total of 17 sectors.

Sector coverage by country

Figure 2 displays the number of sectors covered by agencies in every sector (not the number of institutions). There are a total of 115 countries.

Cluster analysis

Number of clusters

Figure 3 shows the model comparison of several different cluster analysis models based on EM for a parameterized Gaussian mixture models. The model comparison can be performed using the Bayesian Information Criterion (BIC). Specifying 6 clusters for the dataset is aligned with a number beyond which the gains in the Bayesian Information Criterion are almost negligible and the complexity of describing new clusters increases.

Distribution of countries by cluster

Figure 4 presents the proportion of agencies that, within each country, are classified in each of the clusters.

Distribution of sector by cluster

Figure 5 shows the proportion of agencies that, within each sectors, are classified in each of the clusters.

Classification of countries

Figure 6 presents the dendrogram with the classification of countries based on the means of the institutions within country to calculate the country position in each of the dimensions. The dendrogram is based on hierarchical cluster analysis with the Ward agglomerative method and with the distance matrix calculated using Euclidean distances.

Classification of sectors

Figure 7 shows the dendrogram with the classification of sectors based on using means of the institutions within sector to calculate the sector position in each of the dimensions. The dendrogram is based on hierarchical cluster analysis with the Ward agglomerative method and with the distance matrix calculated using euclidean distances.

Mixed Factor analysis

The measurement model is based on mixed factor analysis. This sections presents the JAGS code for the dimension with more variables (namely, Political independence). It is equal for all dimensions, except that for other dimensions not all parts –continuous, binary and ordinal– are necessary.

```
1  model {
2    # Measurement part
3
4    for (o in 1:O) { # Observations
5      # -- Binary variables
6      for (ib in 1:I.binary) {
7        Y.binary[o,ib] ~ dbern(pi[o,ib])
8        logit(pi[o,ib]) <- mu[o,ib]
9        mu[o,ib] <- delta[ib,1] * (xi[o] - delta[ib,2])
10     }
11     # -- Continuous variables
12     for (ic in 1:I.continuous) {
13       Y.continuous[o,ic] ~ dnorm(mu.continuous[o,ic], tau.continuous[ic])
14       mu.continuous[o,ic] <- gamma[ic,1] + (xi[o] * gamma[ic,2])
15     }
16
17     for (io in 1:I.ordinal) {
18       nu[o,io] <- lambda[io,1] + (xi[o] * lambda[io,2])
19     }
20
21     # -----
22     # -- Ordinal variables
23     # -- A loop is not possible, as the number of categories is not always the same
24
25     Y.ordinal[o,1] ~ dcat(p1[o,]) # first categorical variable
26     logit(Q1[o,1]) <- alpha1[1] - nu[o,1]
27     p1[o,1] <- Q1[o,1]
28     for (j in 2:3) {
29       logit(Q1[o,j]) <- alpha1[j] - nu[o,1]
30       p1[o,j] <- Q1[o,j] - Q1[o,(j-1)]
31     }
32     p1[o,4] <- 1 - Q1[o,3]
33
34     # ----- ... until all categorical variables are expressed
35
36   }
37   alpha1 <- sort(alpha1.0)
38   # Priors over thresholds for ordinal variables
39   alpha1.0 [1] ~ dnorm(-3, 4)
40   alpha1.0 [2] ~ dnorm(-1, 4)
41   alpha1.0 [3] ~ dnorm(1, 4)
42   # ----- ... untill all categories are expressed
43
44   # Priors for measurement part - binary
45   for (ib in 1:I.binary) {
46     delta[ib,1] ~ dnorm(0, 1)T(0,)
47     delta[ib,2] ~ dnorm(0, 1)
48   }
49
50   # Priors for measurement part - continuous
51   for (ic in 1:I.continuous) {
52     gamma[ic,1] ~ dnorm(0, 1)
53     gamma[ic,2] ~ dnorm(0, 1)T(0,)
54   }
55
56   for (ic in 1:I.continuous) {
57     tau.continuous[ic] ~ dt(0, pow(0.1, -2), 1)T(0,)
58     sigma.continuous[ic] <- 1/sqrt(tau.continuous[ic])
59   }
60
61   # Priors for measurement part - ordinal
62   for (io in 1:I.ordinal) {
63     lambda[io,1] ~ dnorm(0, 1)
64     lambda[io,2] ~ dnorm(0, 1)T(0,)
65   }
66
67   # Priors for scores of observations
68   for (o in 1:O) {
69     xi[o] ~ dnorm(0, 1)
70   }
71 }
72 }
```

The factor loadings and discrimination parameters are shown in Figure 8

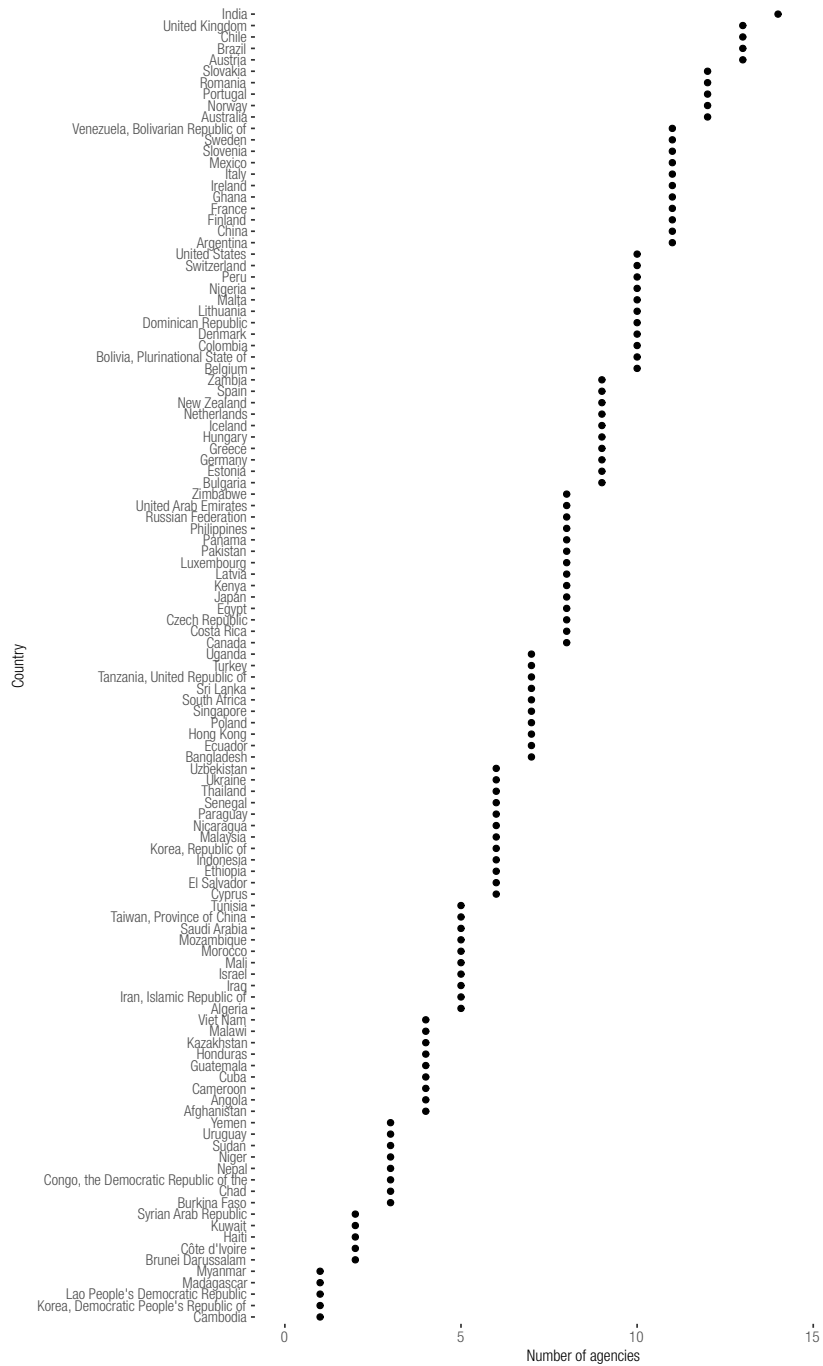


Figure 1: Number of regulatory agencies (institutions) in each country.

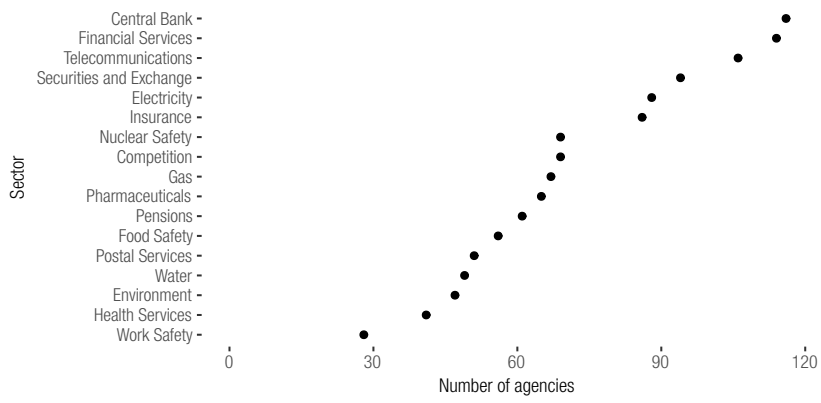


Figure 2: Coverage of regulatory agencies in each sector.

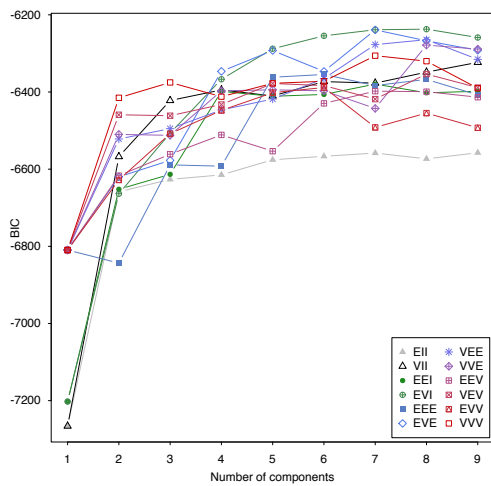


Figure 3: Bayesian Information Criterion (BIC) for several hierarchical cluster analysis models based on EM for a parameterized Gaussian mixture model.



Figure 4: Proportion of agencies that, within each country, are classified in each of the clusters.

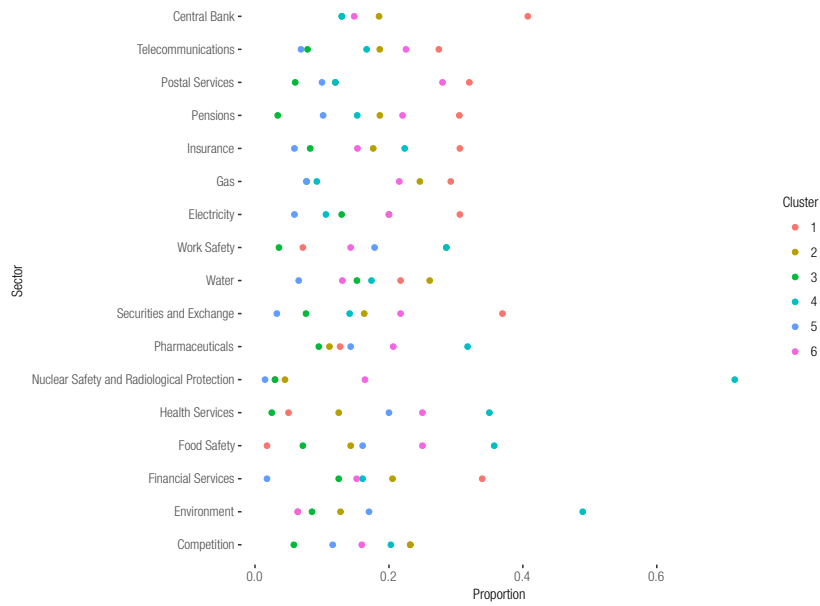


Figure 5: Proportion of agencies that, within each sectors, are classified in each of the clusters.

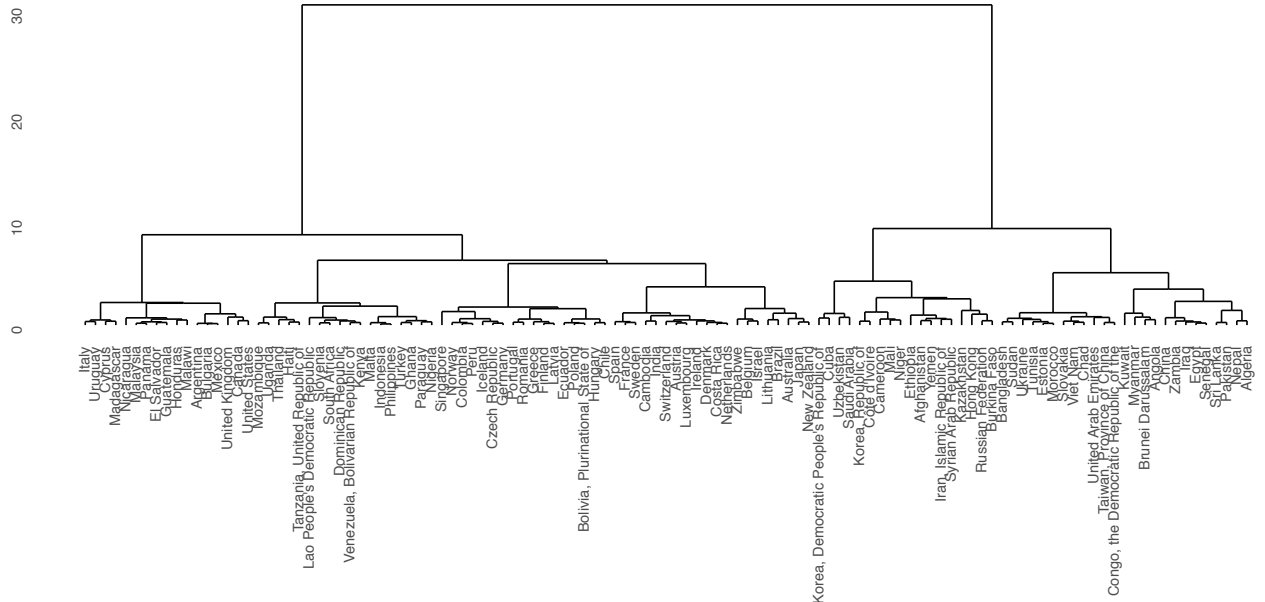


Figure 6: Dendrogram for country means.

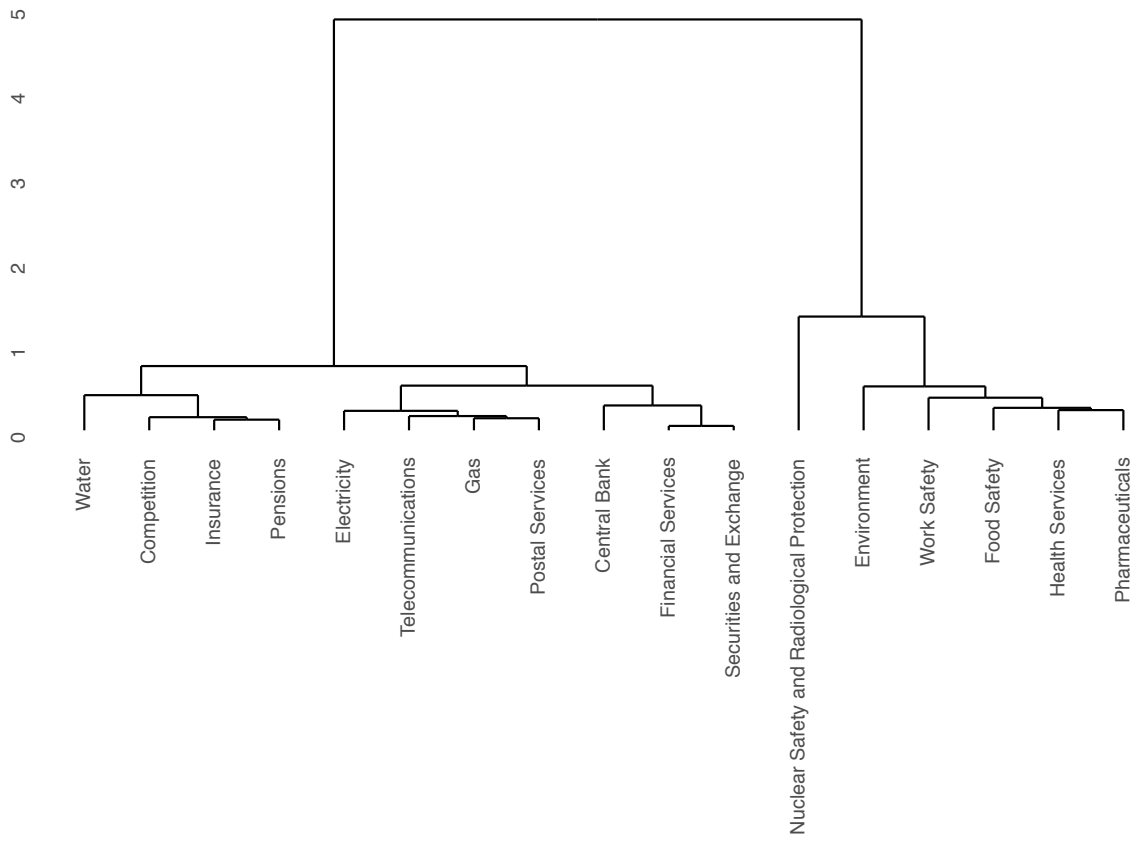


Figure 7: Dendrogram for sector means.

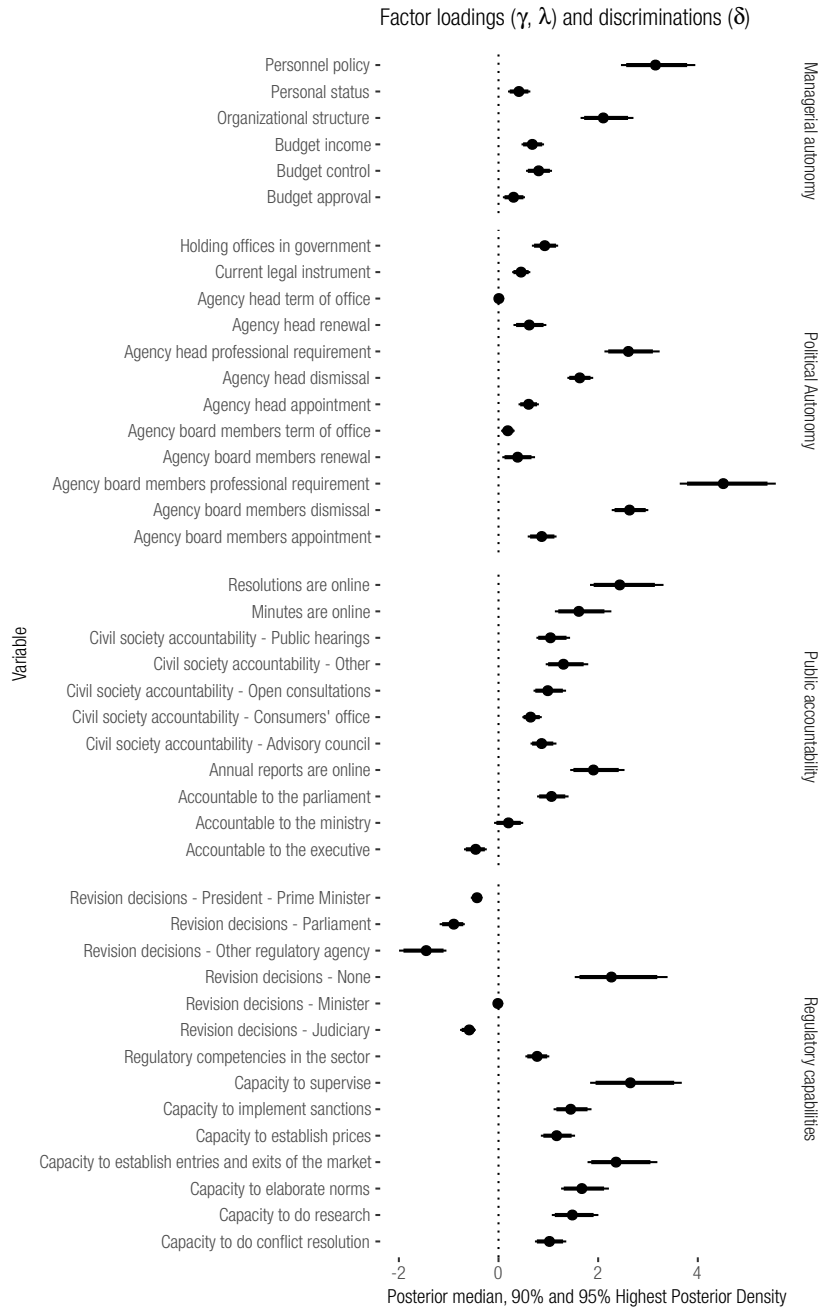


Figure 8: Factor loadings (continuous and ordinal variables) and discriminations (binary variables) for every variable in the measurement model.