# Codebook for diversity dataset

* **consulta\_x**  is the title of the consultation
* **year** is the year in which the consultation took place
* **density** is the total number of actors participating in the consultation (including 1) company, 2) public organization, 3) European body or institution, 4) private person, 5) interest group, 6) international organization or 7) other)
* **nums\_s\_ig** is the number of interest groups (in hundreds of groups) participating in the consultation, calculated as the sum of the following variables divided by one hundred; the following substantive interest group types are based on a modified version of the Schlozman et al. 2008 categorization
	+ **ig\_trade** is the number of trade and other business associations
	+ **ig\_occup** is the number of occupational associations
	+ **ig\_unions** is the number of unions
	+ **ig\_public** is the number of groups representing the public interest
	+ **ig\_idents** is the number of identity groups
	+ **ig\_govts** is the number of organizations that include public authorities as members
	+ **ig\_others** is the number of groups representing other interests
* **div\_all** is the diversity (the Herfindahl-Hirschmann Index) of consultations with respect to the general actor categories
* **subst\_ig** is the diversity (the Herfindahl-Hirschmann Index) of consultations with respect to substantive interest group type
* **pro\_bizocc** is the proportion of interest groups from the trade and other business associations and the occupational associations categories (the sum of these categories divided by the number of interest groups participating in each consultation)
* **p\_costs** is a dummy variable indicating zero if the costs of the policy subject to consultation are diffused across multiple sectors or industries or the taxpayers at large and one if the costs are concentrated within a small number of sectors or industries
* **costs** is a categorical variable indicating one for consultations concerning regulatory policies (those that involve costs to be covered mainly by private actors), two for consultations concerning expenditure policies (costs to be covered by the EU or member state budgets), and three for consultations concerning administrative policies (those with no or only minor costs), following Kardasheva 2009
* **wlength** is the duration of a consultation in weeks (the number of weeks between the announcement of a consultation and the deadline for submissions)

# Codebook for Transparency Register dataset

* **organization** is the name of the organization as it appears in the Transparency Register
* **subst\_interest** is a categorical variable indicating the substantive interest group type based on a modified version of the Schlozman et al. 2008 categorization, as follows:
	+ 1 = trade and other business
	+ 2 = occupational
	+ 3 = unions
	+ 4 = public interests
	+ 5 = identity
	+ 6 = associations including public authorities
	+ 7 = other