FGV DIREITO RIO

Ideias, interesses e a efetividade da participação social: uma análise a partir das consultas públicas da Anatel

GT33: Pesquisa Empírica em Direito da Regulação e Política Regulatória (EPED 2025)

Lucas Thevenard Gomes (FGV Direito Rio / Regulação em Números)

FGV DIREITO RIO

Como grupos de interesse influenciam decisões regulatórias em mecanismos de participação social?

- Duas grandes vertentes teóricas:
  - Teoria do triângulo de ferro e captura regulatória
  - Teoria das redes temáticas e coalizões de interesse

Não necessariamente uma dicotomia...

> *No one argues that there are only issue networks or only subgovernments active in policymaking. Rather, the argument is over what is most typical and most descriptive of the policy process. Which should serve as our framework for analyzing how laws and regulations are made (p. 243-44)?*
>
> ***Jeffrey Berry (1989)***

- **1º estudo: "*Who says it, or what is said? Who says it, or what is said? Language, ideas, and influence in public participation*"**
  - Como o conteúdo textual das contribuições públicas afeta seu impacto regulatório nas consultas públicas da Anatel?
- **2º estudo: "*Speaking as One? Discursive Consistency and Fragmentation among Interest Groups*"**
  - Os grupos de interesse se posicionam de forma coesa e coerente nos mecanismos de participação social?
- **3º estudo: "*Ideas in Motion: Co-Participation and Ideational Proximity in Regulatory Issue Networks*"**
  - Qual é o nível de assiduidade dos participantes dos grupos de interesse? Há padrões de co-participação fixos ("panelinhas")?

Hipótese geral da tese

- **As ideias** — o conteúdo substantivo das contribuições públicas — desempenham um papel decisivo em pelo menos uma parte dos processos participativos.
  - Ou seja, **a teoria das redes temáticas tem grande relevância** para a análise dos mecanismos de participação social.

Who says it, or what is said? Who says it, or what is said? Language, ideas, and influence in public participation

Data

- 9,404 contributions received by Anatel **since 2013** that were accompanied by a **classified response**.
- Simplification of impact categories:
  - Accepted, partially accepted, not accepted, and not applicable ---> 0 x 1
- Classification of Interest group affiliation:
  - (1) regulated economic agents, (2) non-regulated economic agents, (3) sector professionals, (4) amateur radio operators, (5) interested individuals, and (6) others.

Three hypothesis of this study

Hypothesis 1: Contributions with similar textual content receive similar regulatory responses.

Hypothesis 2: Contributions that are textually closer to the agency’s discourse are more likely to be accepted.

Hypothesis 3: Once textual content is accounted for, group affiliation loses explanatory power.

Text Representation: TF-IDF

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \text{IDF}(t, D)$$

This score increases with the frequency of a term in a document but is offset by how common the term is across the corpus.

- **Intuition**:
  - Common words (like “the”, “and”) get low scores.
  - Terms that are frequent in a document but rare across the corpus get high scores.
  - Documents (contributions) become vectors in a high-dimensional space, where each dimension corresponds to a term in the vocabulary.

Predictive modelling

- Open code is available for all the modelling stages.

Table 4 – Consolidated predictive performance (test and validation)

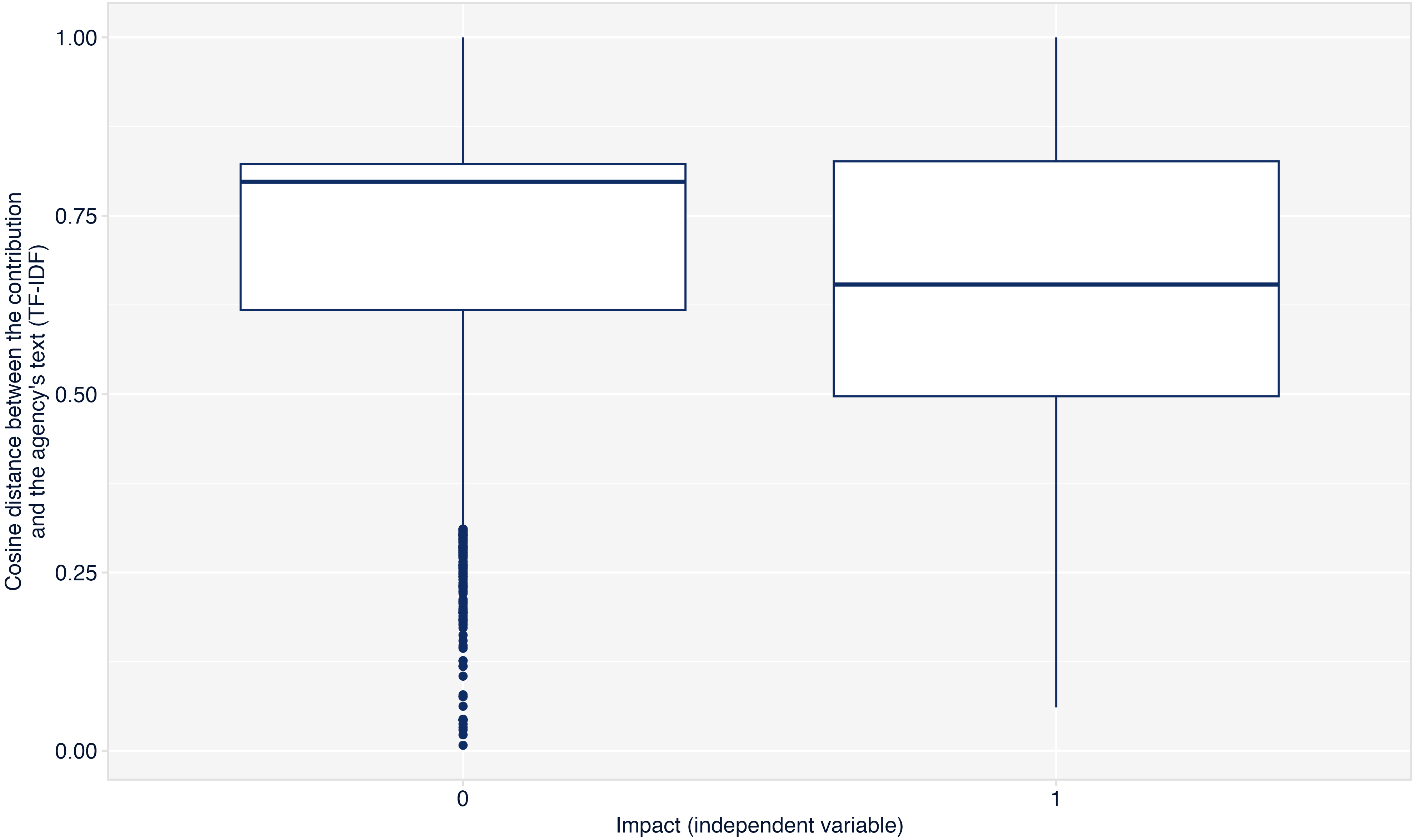
Model	Evaluation	Accuracy	ROC-AUC	Precision	Recall	F1	N
KNN	Test	85.25%	0.91	0.83	0.80	0.81	1,505
	Validation	84.90%	0.91	0.83	0.79	0.81	1,881
RF	Test	86.05%	0.94	0.87	0.78	0.81	1,505
	Validation	86.76%	0.94	0.87	0.80	0.82	1,881
SVC	Test	88.24%	0.94	0.89	0.82	0.84	1,505
	Validation	88.57%	0.94	0.89	0.83	0.85	1,881

Explanatory modeling (Model 1)

- Cosine distance as a textual distance metric

Table 6 – Logistic regression: impact explained by textual distance and year of consultation

	Coefficient	Std. Error	z-value	p-value	
Intercept	182,02	30,02	6,05	1,44E-09	***
Cosine distance	-1,77	0,12	-15,13	2,00E-16	***
Year	-0,09	0,01	-6,04	1,56E-09	***
Significance codes: 0 *** 0,001 ** 0,01 * 0,05 . 0,1					


Cosine distance between the contribution and the agency's text (TF-IDF)
1.00
0.75
0.50
0.25
0.00
0
1
Impact (independent variable)

Explanatory modeling (Model 2)

Table 7 – Logistic regression: impact explained by model-based probabilities and consultation year

	Coefficient	Std. Error	z-value	p-value	
Intercept	54,71	68,99	0,79	0,4277	
Year	-0,03	0,03	-0,87	0,3820	
KNN Probability	-0,71	0,31	-2,32	0,02	*
RF Probability	11,26	0,48	23,51	2,00E-16	***
SVC Probability	2,15	0,33	6,36	2,04E-10	***
	Significance codes: 0 *** 0,001 ** 0,01 * 0,05 . 0,1				

Table 8 –Logistic Regression: Impact explained by year, ML scores, and interest group category

	Coefficient	Std. Error	z-value	p-value	
Intercept	9,70	69,39	0,14	0,8887	
Year	-0,01	0,03	-0,22	0,8282	
KNN Probability	-0,68	0,31	-2,21	0,0269	*
RF Probability	11,62	0,52	22,55	2,00E-16	***
SVC Probability	1,94	0,35	5,55	2,79E-08	***
Participant: Regulated Agent	-0,37	0,13	-2,93	0,0034	**
Participant: Non-Regulated Agent	-0,55	0,26	-2,08	0,0376	*
Participant: Sector Professional	-3,04	0,50	-6,10	1,05E-09	***
Participant: Amateur Radio Operator	-1,67	0,60	-2,75	0,0059	**
Participant: Other	-0,69	0,3	-2,33	0,0197	*
Significance codes: 0 *** 0,001 ** 0,01 * 0,05 . 0,1					

Going back to our hypothesis

Hypothesis 1: Contributions with similar textual content receive similar regulatory responses.

Hypothesis 2: Contributions that are textually closer to the agency’s discourse are more likely to be accepted.

Hypothesis 3: Once textual content is accounted for, group affiliation loses explanatory power.

Some additional analysis

- Is it technical language that matters?
- Content diversity and impact: which groups are more likely to repeat their language?

FGV DIREITO RIO

Things to improve

- Better representation of the textual content
  - Use of word embeddings (e.g., BERT, Word2Vec)
  - Use of topic modeling (e.g., LDA, NMF)
- Separation of stylistic and substantive content
- Use of all the categories of impact

Obrigado!