

Ideas, interests and the effectiveness of public participation

2025.2

Lucas Thevenard

Chapter 1: Introduction

Chapter 1: Introduction

- Does commenter identity matter in notice and comment procedures?
 - Balla (2022): bureaucrats say it doesn't
 - Legal framework: participation as information exchange (a dialogue)
- Interest group theory: theoretical focus on who participates
 - Participation as representation
 - Facing the research challenges of measuring ideas: two ways to use IA
- Implications for regulatory policy
 - Concrete policy implications
 - Access is not enough (Regulations.gov, examples from Anvisa and Aneel)

Chapter 1: Introduction

Outline:

- Chapter 2: Theoretical Framework
- Chapter 3: ANATEL's public consultations: methodology and descriptive statistics
- Chapter 4: Who says it, or what is said?
- Chapter 5: Uncovering interest group dynamics
- Chapter 6: Issue networks and the circulation of ideas
- Chapter 7: Conclusion

Chapter 2: Theoretical Framework

Chapter 2: Theoretical Framework

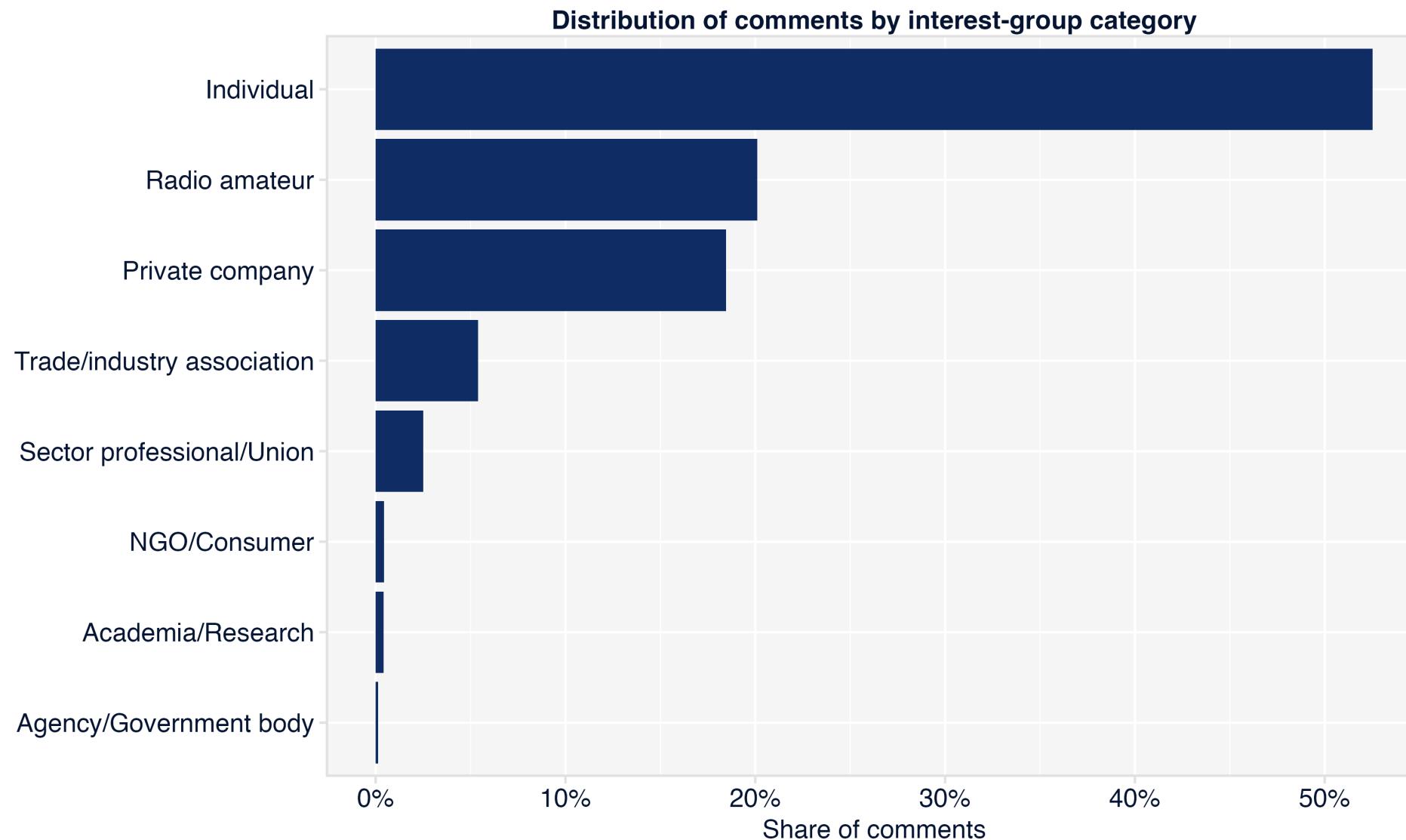
-  UNDER CONSTRUCTION 
- "Knowledge" vs. "Interests"
- More importantly: applied research

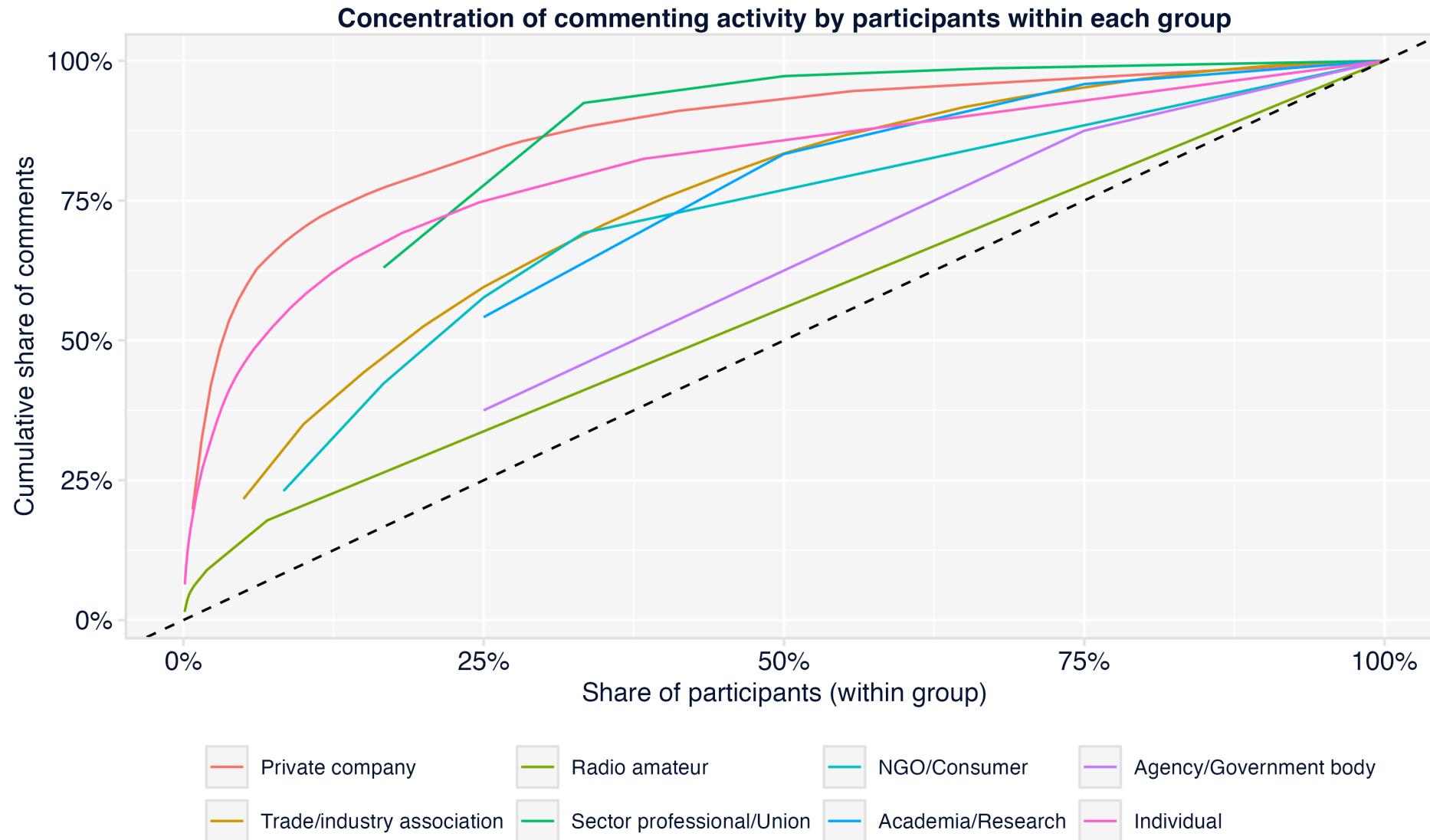
Chapter 3: Methodology and descriptive statistics

Chapter 3: Methodology and descriptive statistics

- **Methodology**
 - Data collection
 - Text preprocessing and representation (TF-IDF, BERT, GPT-3 embeddings)
 - LLM-assisted dataset enrichment
 - Interest group classification
 - Regulatory direction classification
 - Types of arguments
- **Descriptive statistics:** ...

Composition of participants





Regulatory direction of comments

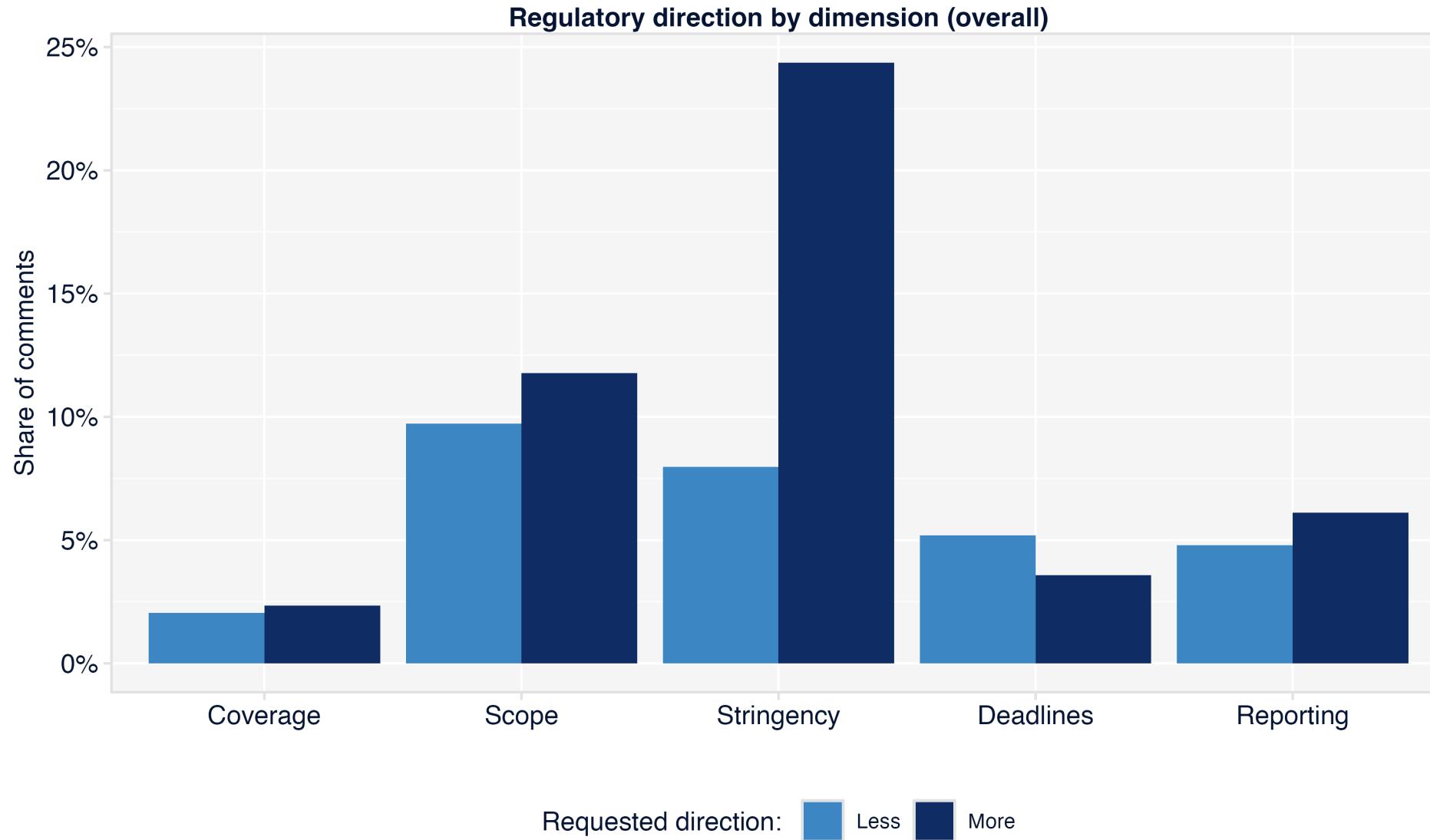
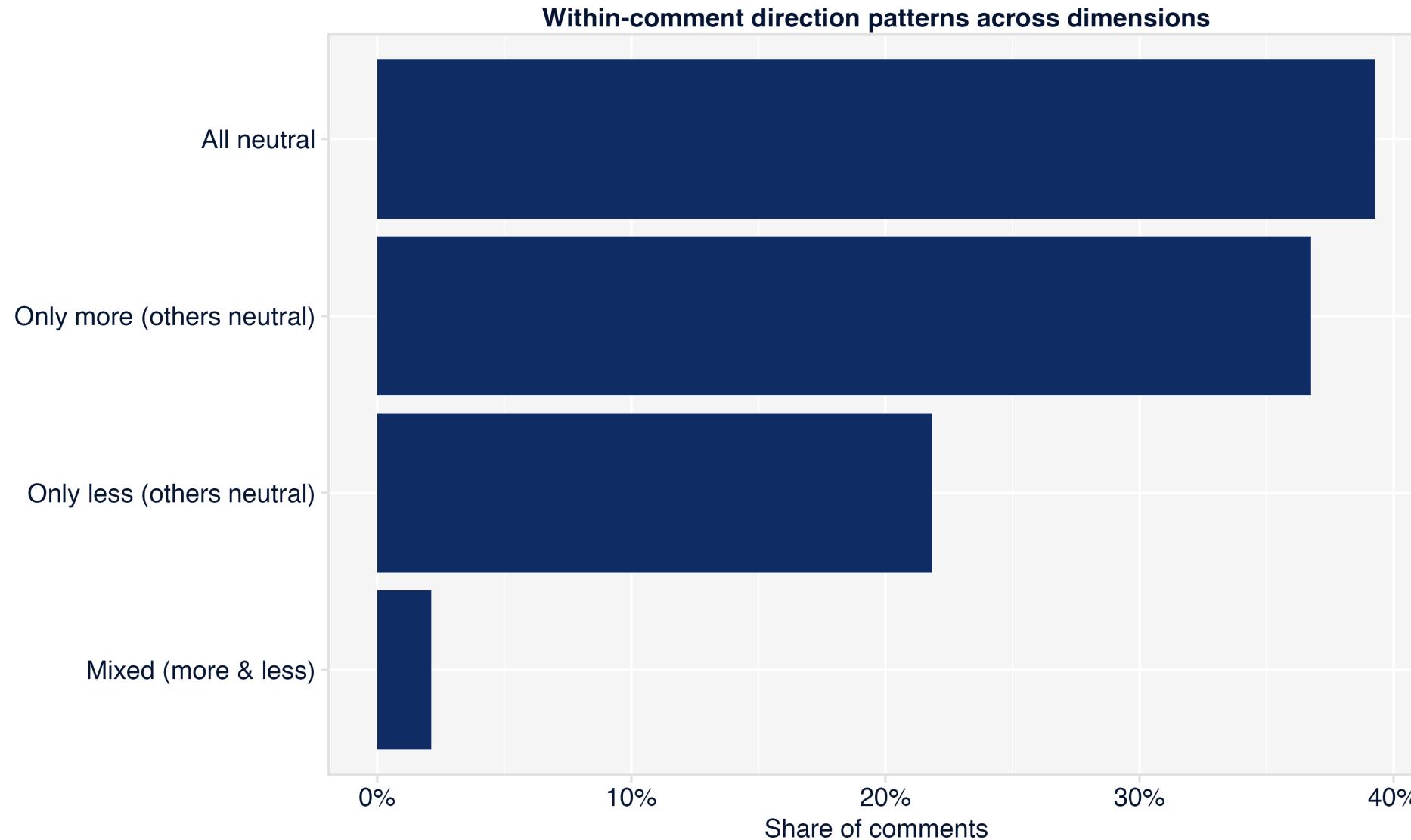
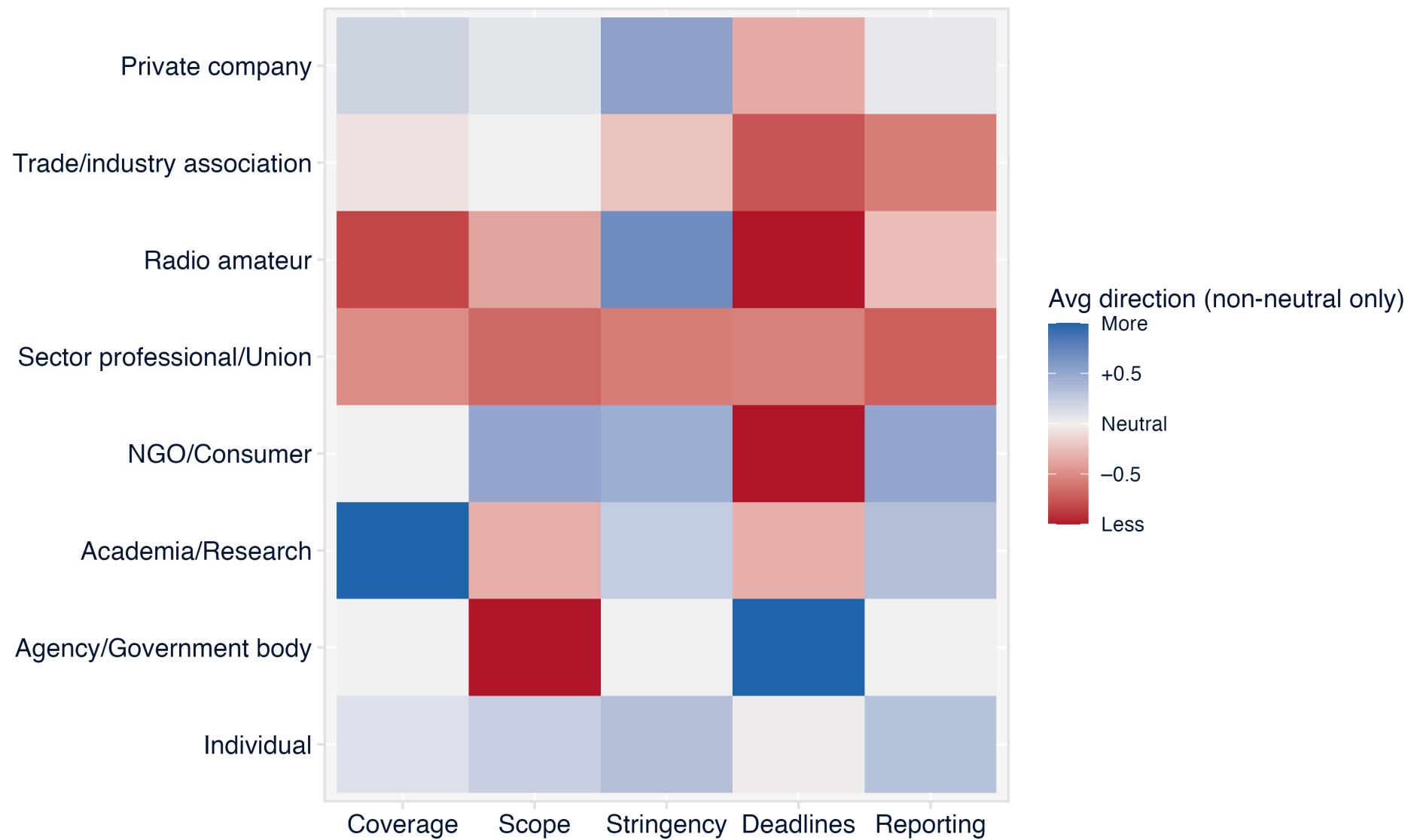


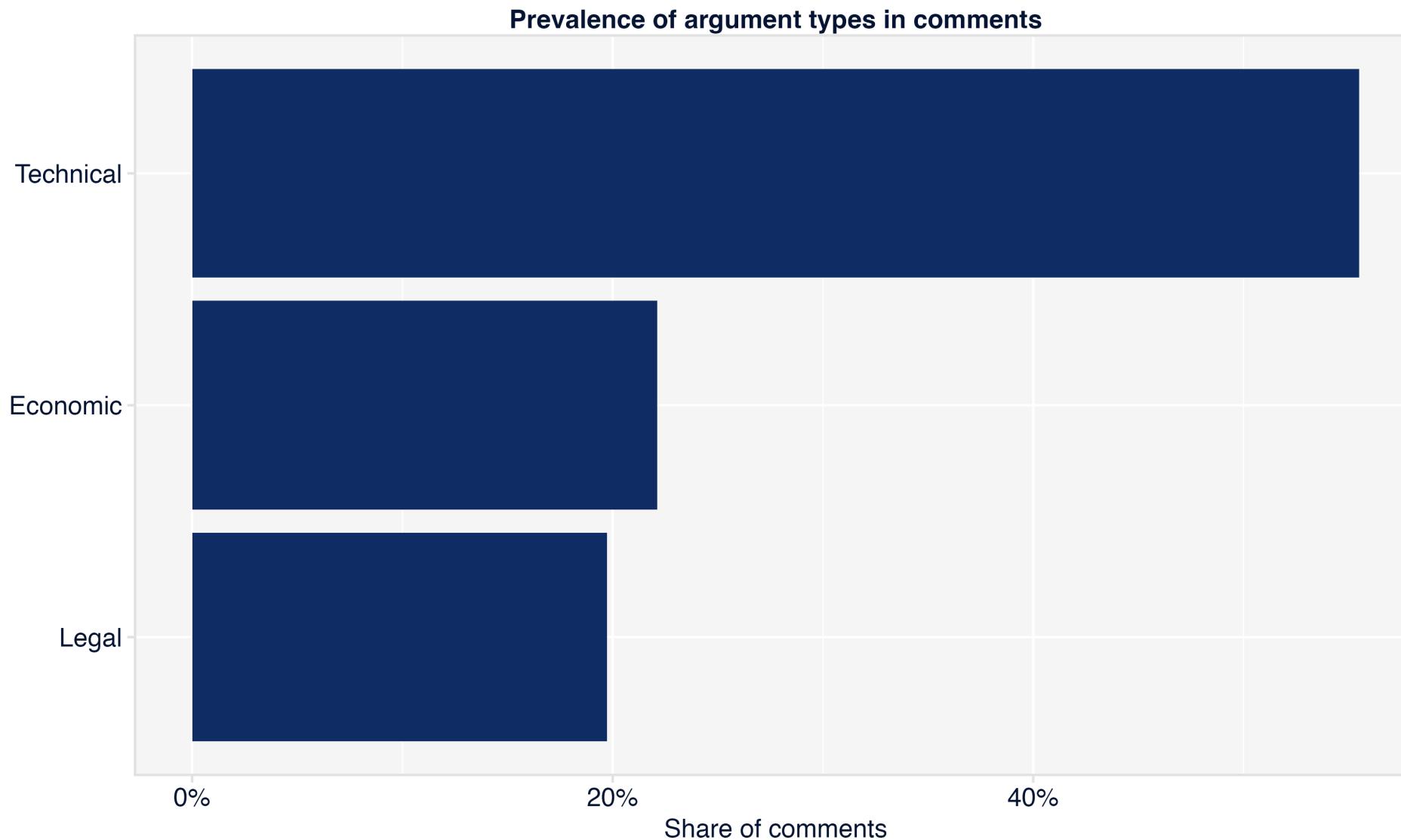
Table X – Regulatory direction dimensions addressed in the comments

Number of dimensions addressed in each comment	Number of comments	Percentage of Comments
0 dimensions	2284	39.3%
1 dimension	2726	46.9%
2 dimensions	631	10.9%
3 dimensions	151	2.6%
4 dimensions	22	0.4%
5 dimensions	0	0.0%





Argument types in comments



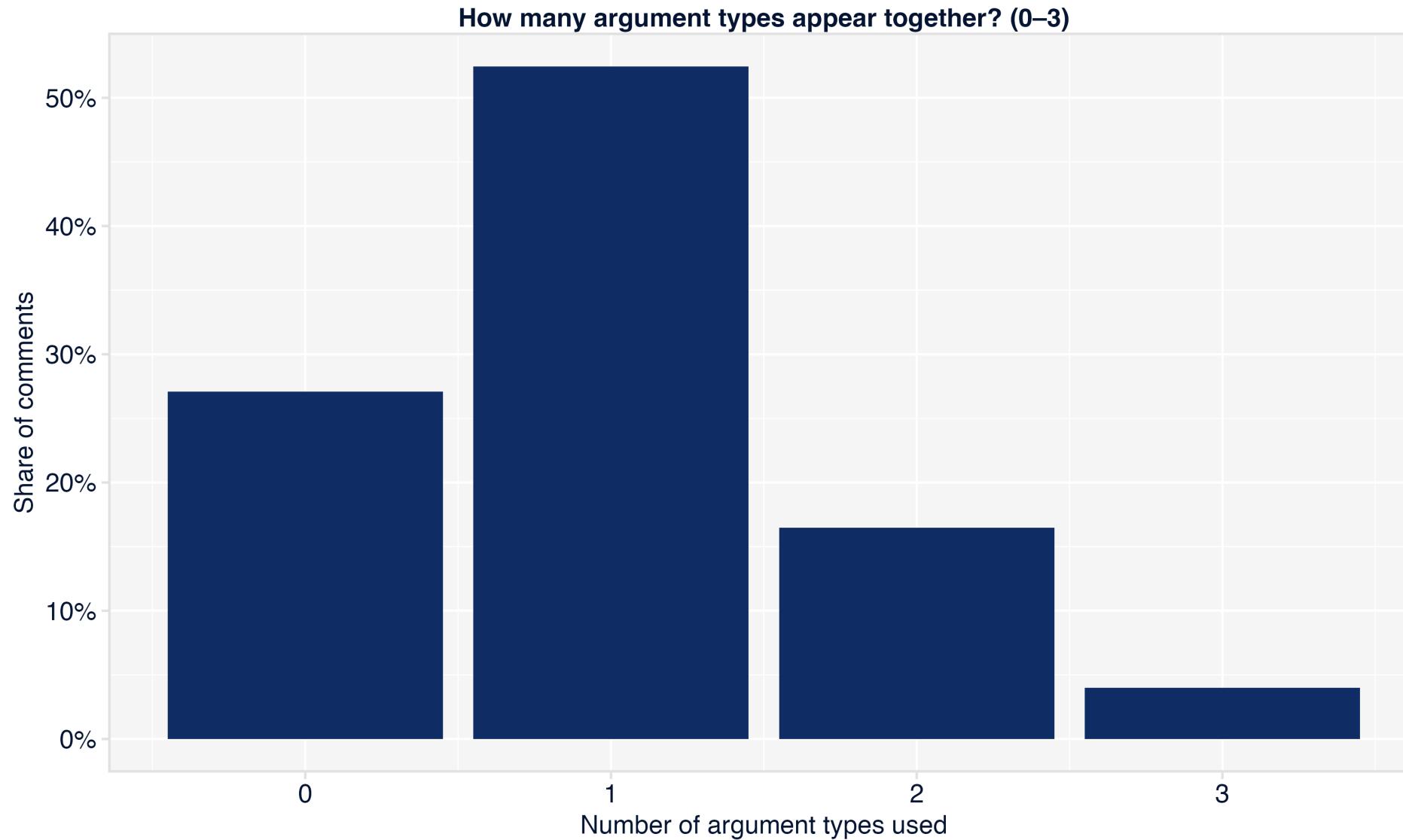
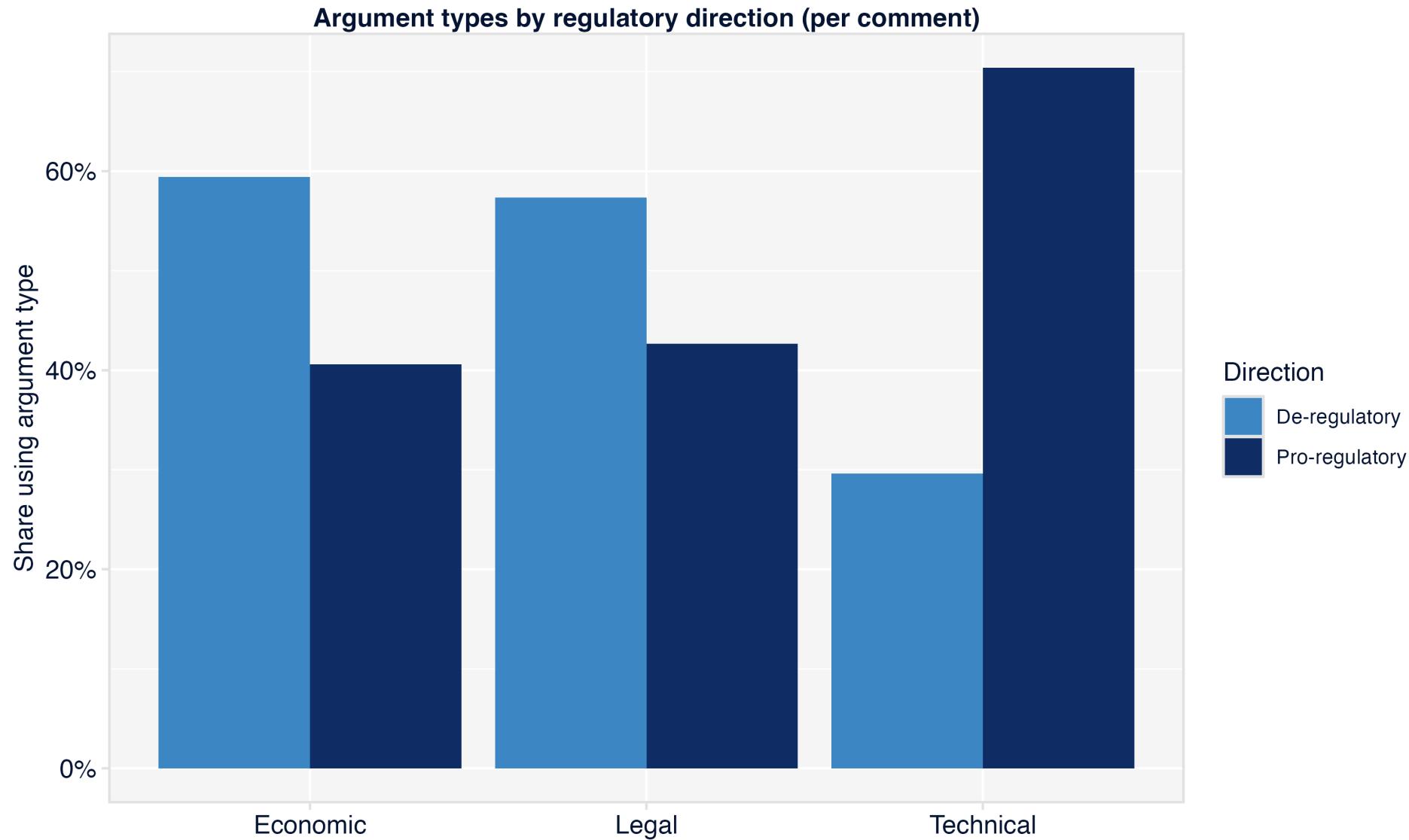
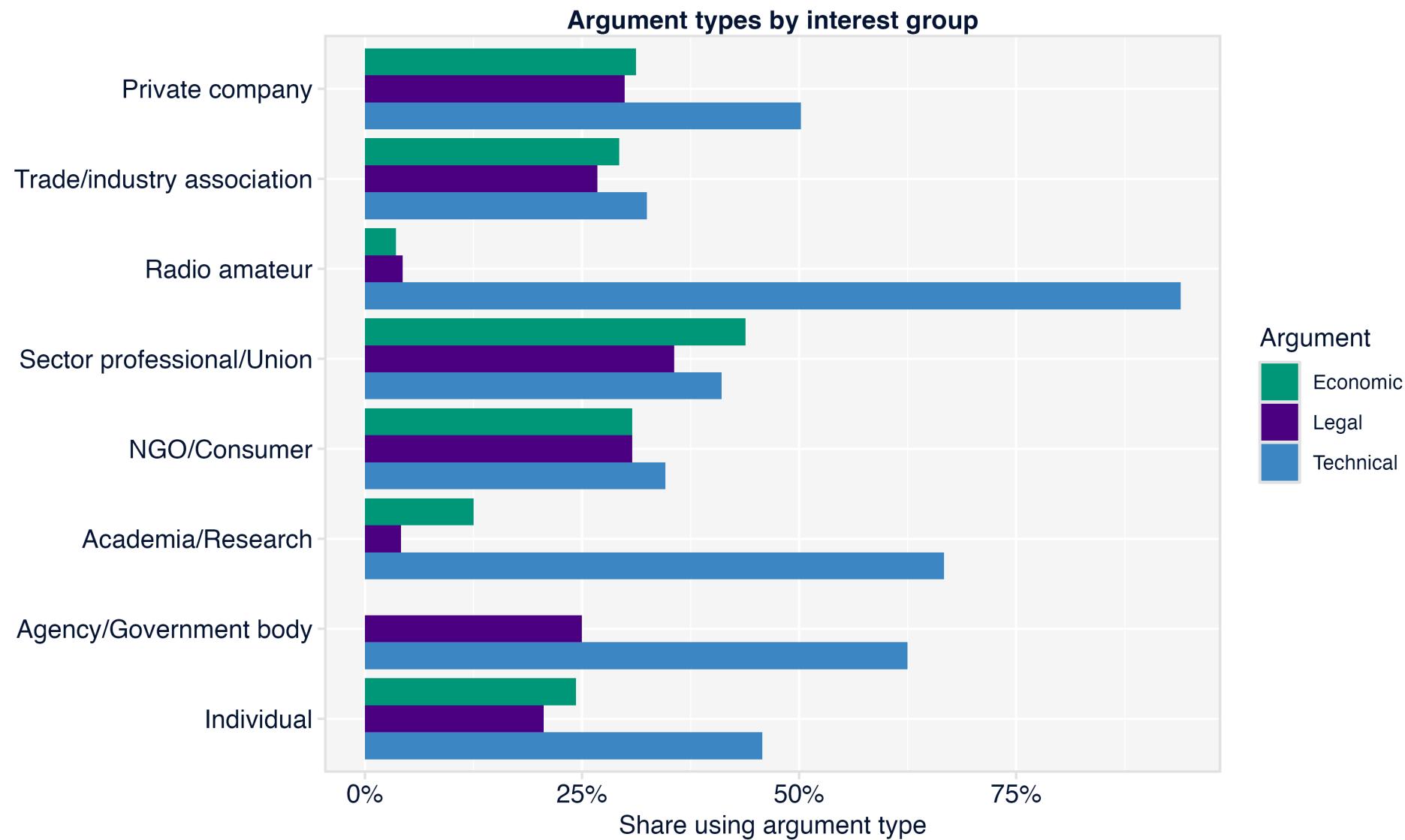


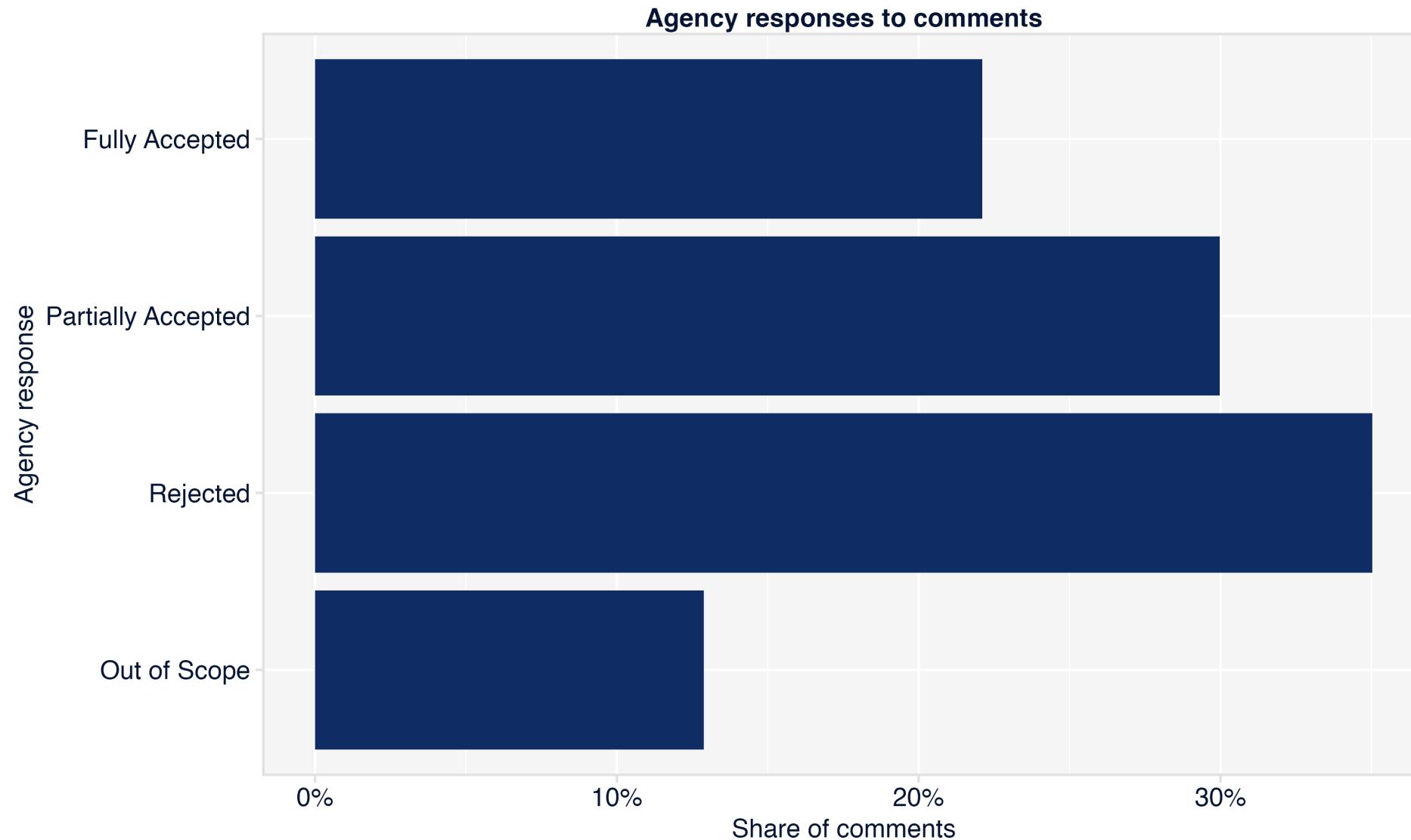
Table X – Concurrence of types of arguments

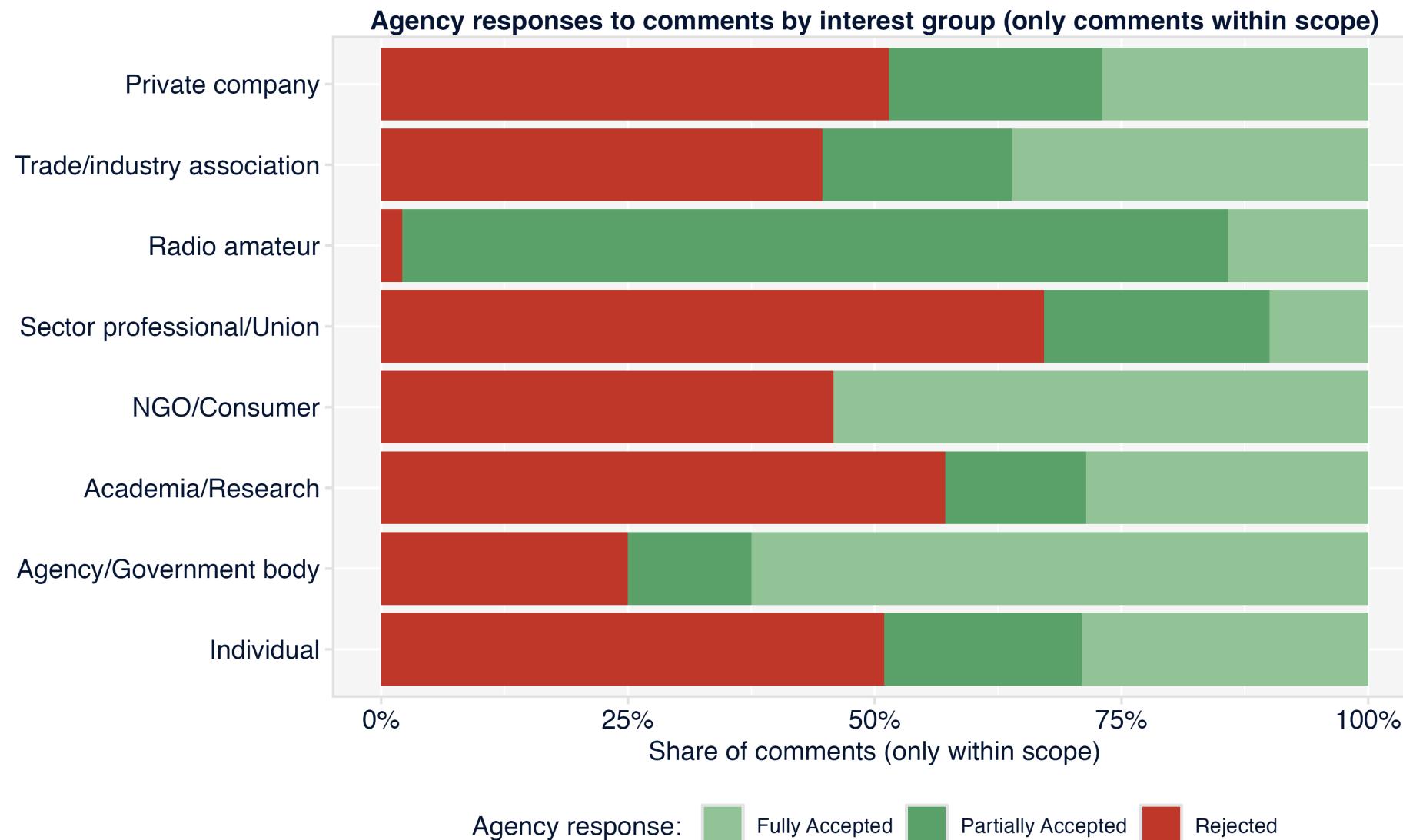
A	B	N Both	Jaccard (A,B)	P(B A)	P(A B)
Economic	Legal	451	0.228	0.351	0.393
Economic	Technical	688	0.18	0.535	0.213
Legal	Technical	514	0.133	0.448	0.159

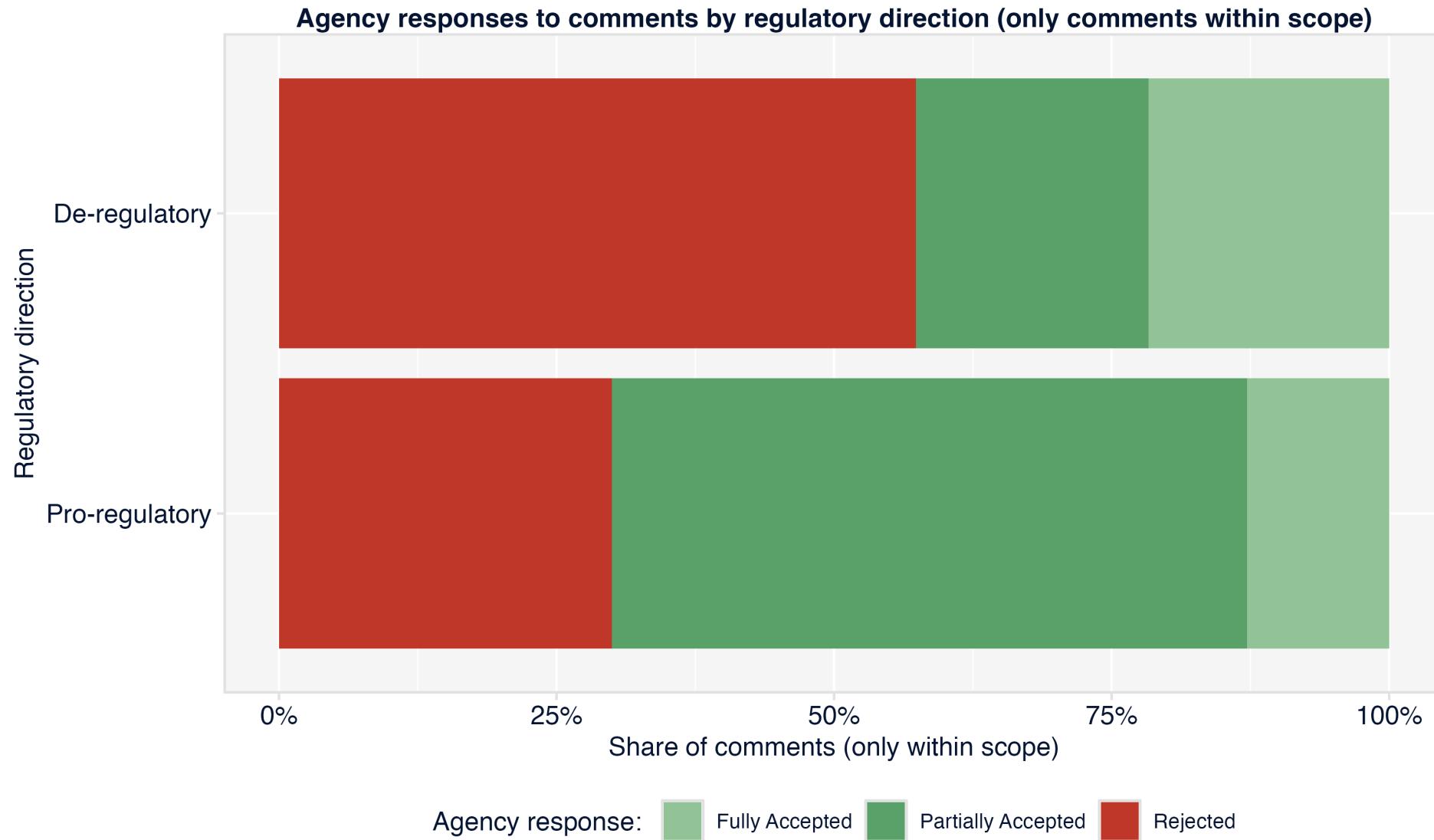


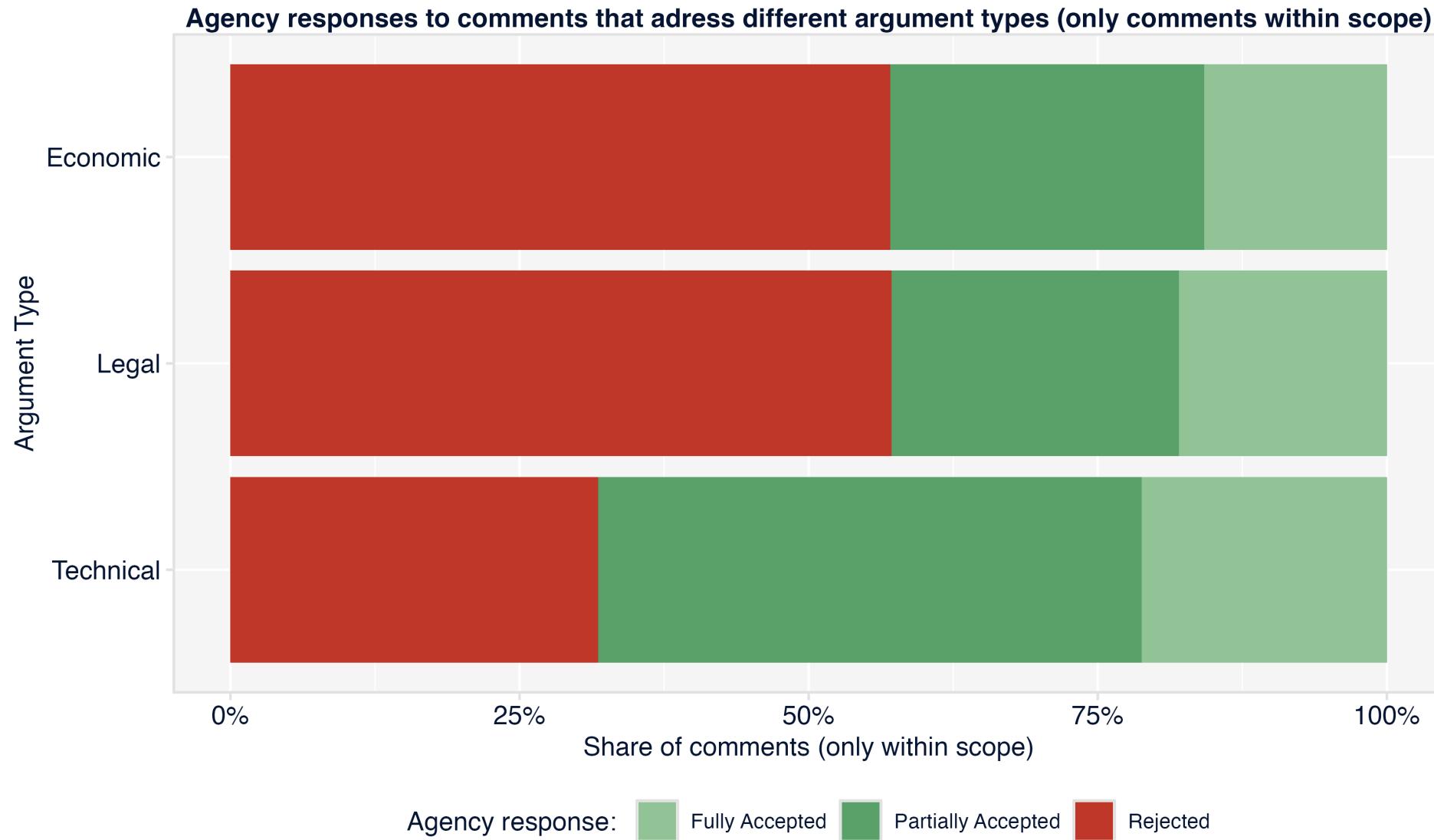


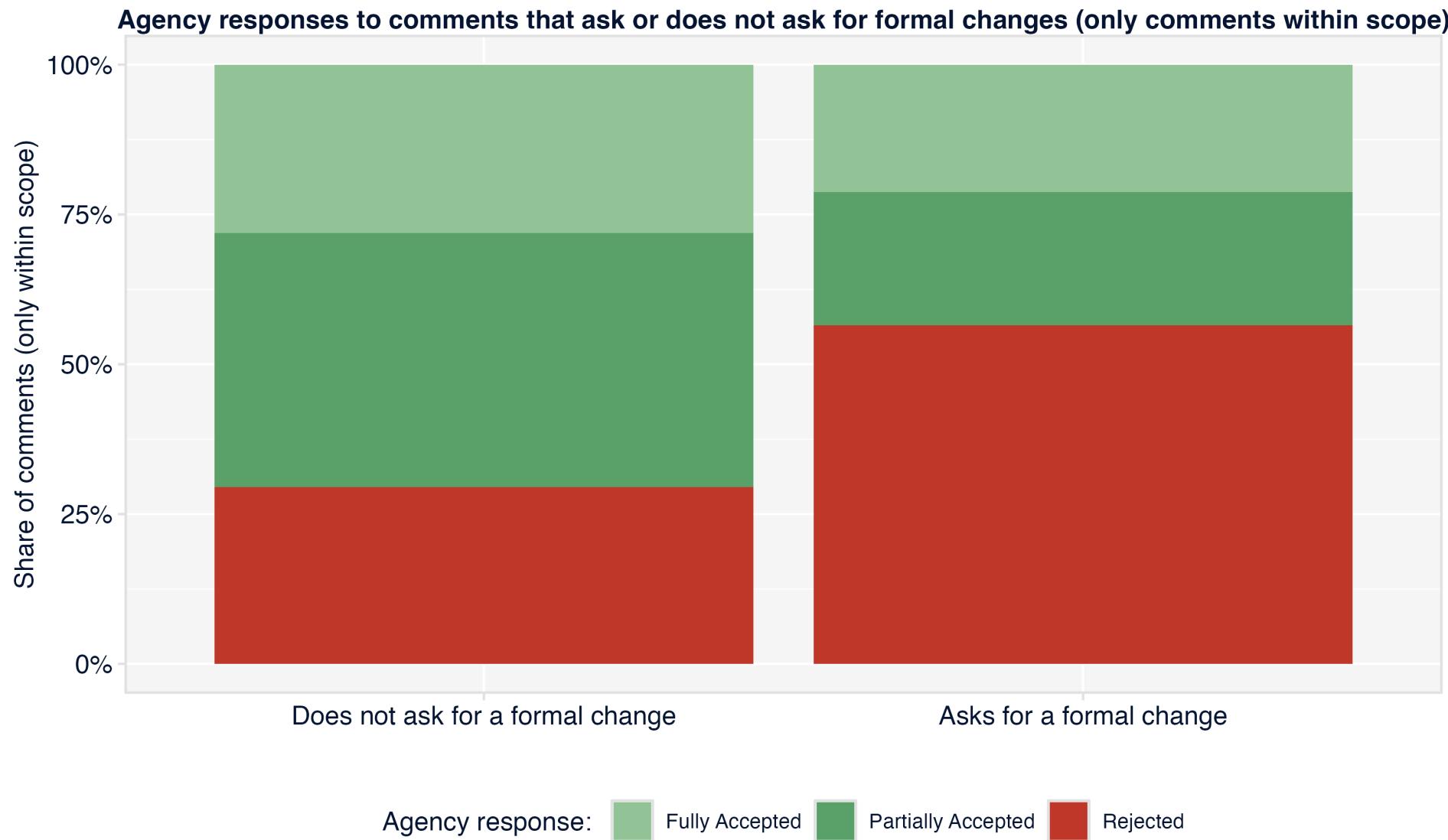
Agency's responses to comments











Chapter 4: Who says it, or what is said?

Chapter 4: Who says it, or what is said?

- **Research design and hypotheses:** what matters most, ideas or group affiliation?
 - **Predictive logic:** measuring the informational regularity of responsiveness
 - **Explanatory logic:** assessing the relative weight of ideas and interests

Chapter 4: Who says it, or what is said?

H1 – Predictability of scope:

The linguistic and semantic content of comments is sufficient to predict, with high accuracy and balanced precision-recall, whether ANATEL deems them within scope or out of scope.

Rationale: The agency's first procedural filter is expected to follow consistent legal and technical criteria; therefore, textual cues such as topic relevance and terminology should strongly correlate with scope classification.

Chapter 4: Who says it, or what is said?

H2 – Predictability of impact:

Among comments deemed within scope, textual content predicts acceptance or rejection with performance significantly above the random baseline.

Rationale: If the agency's evaluation of relevant comments is systematic and reason-based, the linguistic and argumentative features of the text should align with the agency's final decisions.

Chapter 4: Who says it, or what is said?

H3 – Explanatory weight of content:

In explanatory regressions, content-based variables (textual and argumentative) substantially increase the model's explanatory power (R^2) relative to identity-only models.

Rationale: If ideas influence outcomes independently of interests, adding content variables should capture a large portion of the variance in acceptance that identity alone cannot explain.

Chapter 4: Who says it, or what is said?

H4 – Diminishing role of identity:

Once content-based variables are introduced, the coefficients associated with interest-group identity should lose magnitude or statistical significance.

Rationale: If responsiveness reflects the informational value of comments rather than the power or status of actors, identity effects will attenuate when controlling for content.

How predictable is the agency identification of 'out-of-scope' cases?

- **Scope classification**
 - Best results (final validation stage, N=150): **SVM** and **KNN**
 - SVM: Accuracy ≈ **89.3%**, F1 ≈ **0.89333**
 - KNN: Accuracy ≈ **88.7%**, F1 ≈ **0.88654**
 - Comments judged “out of scope” can be predicted with high reliability

How predictable are agency's acceptance/rejection responses?

- **Impact classification**
 - Best results (final validation stage, N=815): **SVM** and **KNN**
 - SVM: Accuracy ≈ **83.4%**, F1 ≈ **0.83435**
 - KNN: Accuracy ≈ **81.9%**, F1 ≈ **0.81867**
 - Agency acceptance/rejection is also highly predictable from comment text

Explanatory models (logistic regressions)

- M0 - only identity
- M1 - identity + embeddings
- M2 - only embeddings
- M3 - only content variables
- M4 - embeddings + content variables
- M5 - identity + embeddings + content variables
- M6 - identity + predicted probabilities
- M7 - only predicted probabilities
- M8 - predicted probabilities + content variables
- M9 - identity + predicted probabilities + content variables

In conclusion

Next steps

- Finish explanatory models for chapter 4 (almost done)
- Next empirical chapters (5 and 6): all the data is there, now is just a matter of doing the analysis and writing the results.
- Chapter 2 and Conclusion: for most of next year, refining the theoretical chapter, revision and improving discussions of the results.

Thank you