Sustainability
Reporting Navigator

SFB/Transregio 266
ACCOUNTING FOR
TRANSPARENCY

How does the CSRD change sustainability reporting?

Analyzing the first reports

Victor Wagner, Maximilian A. Müller & Lea Hagemeier

2025-03-10

Data

- We have compiled a (continuously updated) list of CSRD-compliant reports on our homepage.
- This presentation presents some evidence on the content and characteristics **for the reports of 90 firms,** compared to their 2023 counterparts.
- *Text used:* If a firm issued a stand-alone sustainability report in 2023, we use only the sustainability report as the 2023 baseline, if not, we use the pages in the 2023 annual report that are dedicated to sustainability. For the 2024 reports, we use only those pages that contain the CSRD-compliant sustainability statement.
- *Textual analysis method:* A more detailed explanation on how we apply *word2vec* to identify the topics of the sustainability statements, please see the Appendix or our registered report proposal.

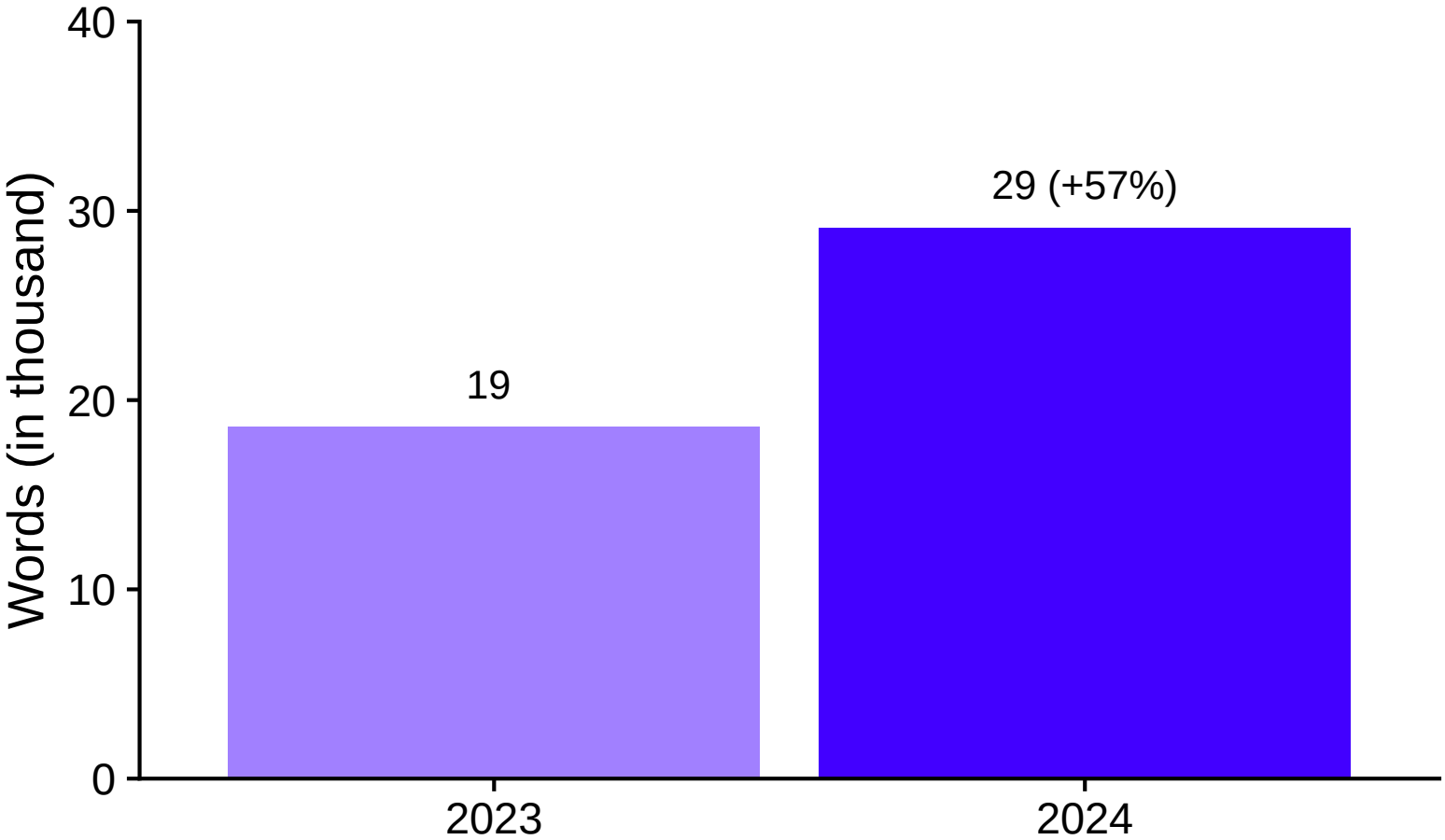

Fig 1. Growth in reporting volume. This figure shows the average number of words in a firm's 2023 and 2024 sustainability reporting.

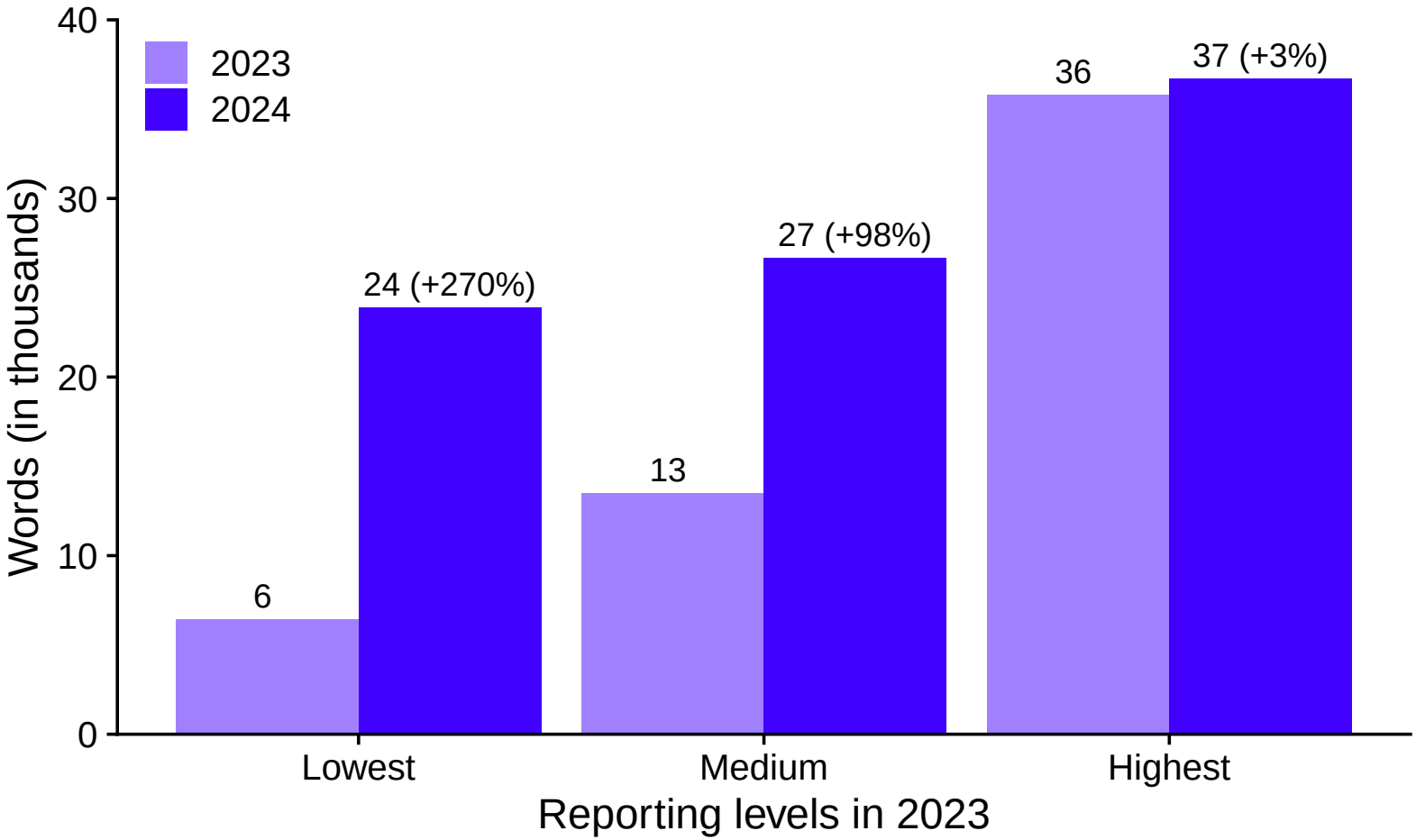

Fig 2. Growth in reporting volume by reporting level in 2023. This figure shows that the firms with the lowest and medium reporting levels in 2023 increased their reporting disproportionally in 2024 (+270% for the lowest and +98% for medium reporting levels). The firms with high reporting levels in 2024 only moderately reported more (+3%).

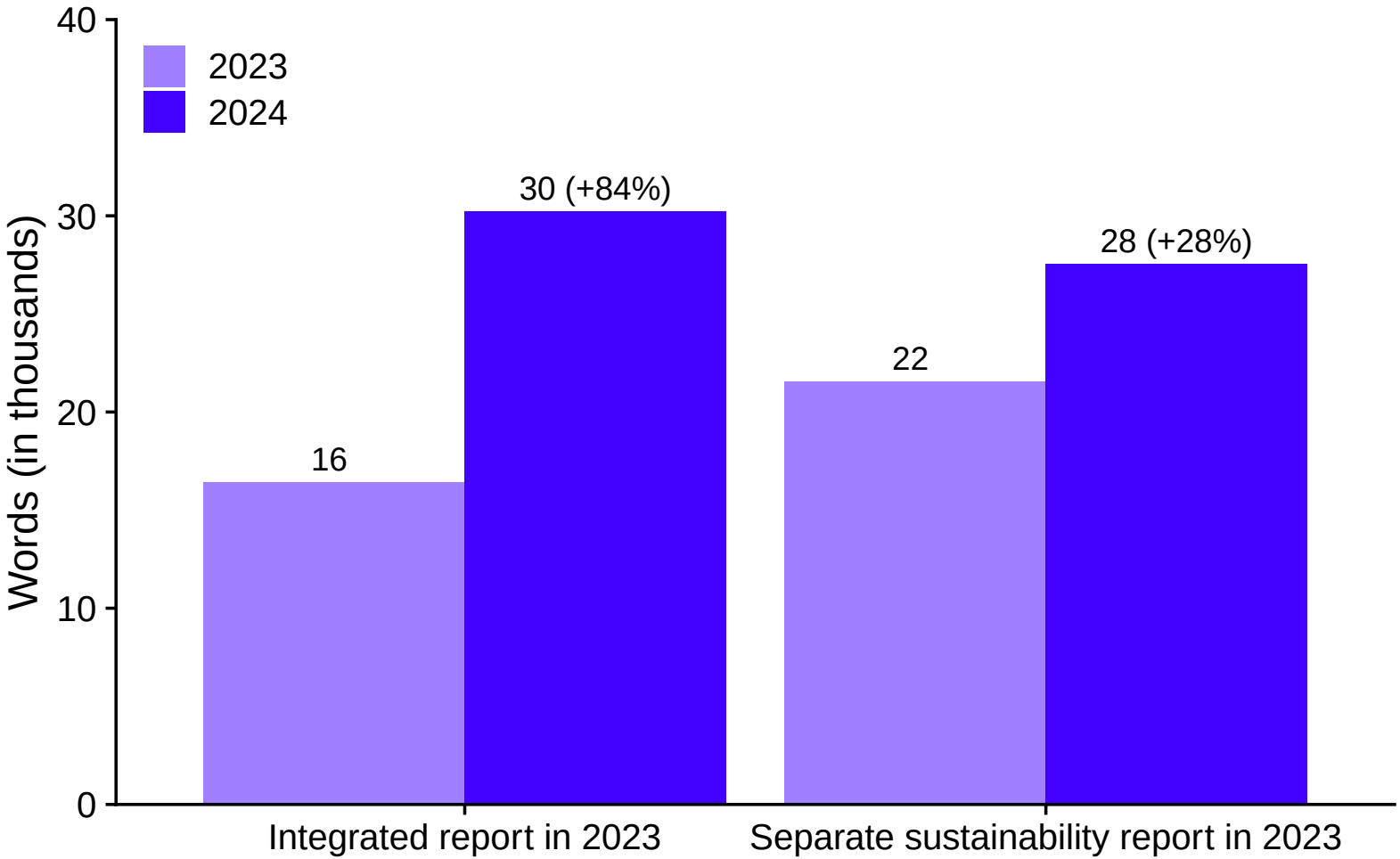

Fig 3. Growth in reporting volume for firms with a separate sustainability report in 2023. This figure shows that firms that published a sustainability report in 2023 increased their reporting

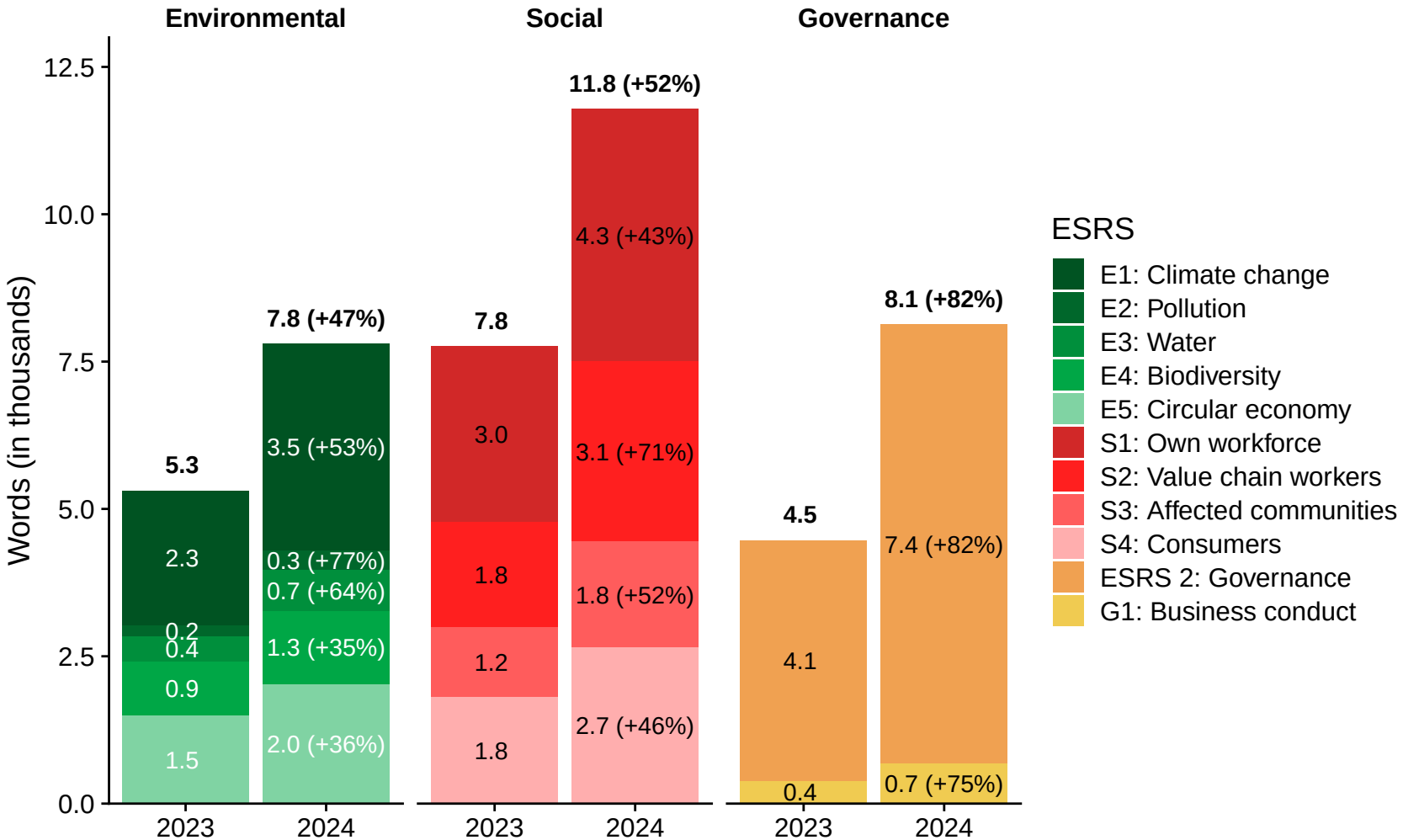

Fig. 4: Reporting volume by ESRS. This figure shows the average word count per topic as well as the relative increases in 2024. To identify topics, we use a *word2vec* algorithm that finds similar words based on a list of seed words. We define ten seed words based on the ESRS and score each sentence from the sustainability report using these keywords. Find more information on the algorithm at the end of this presentation.

Something about materiality

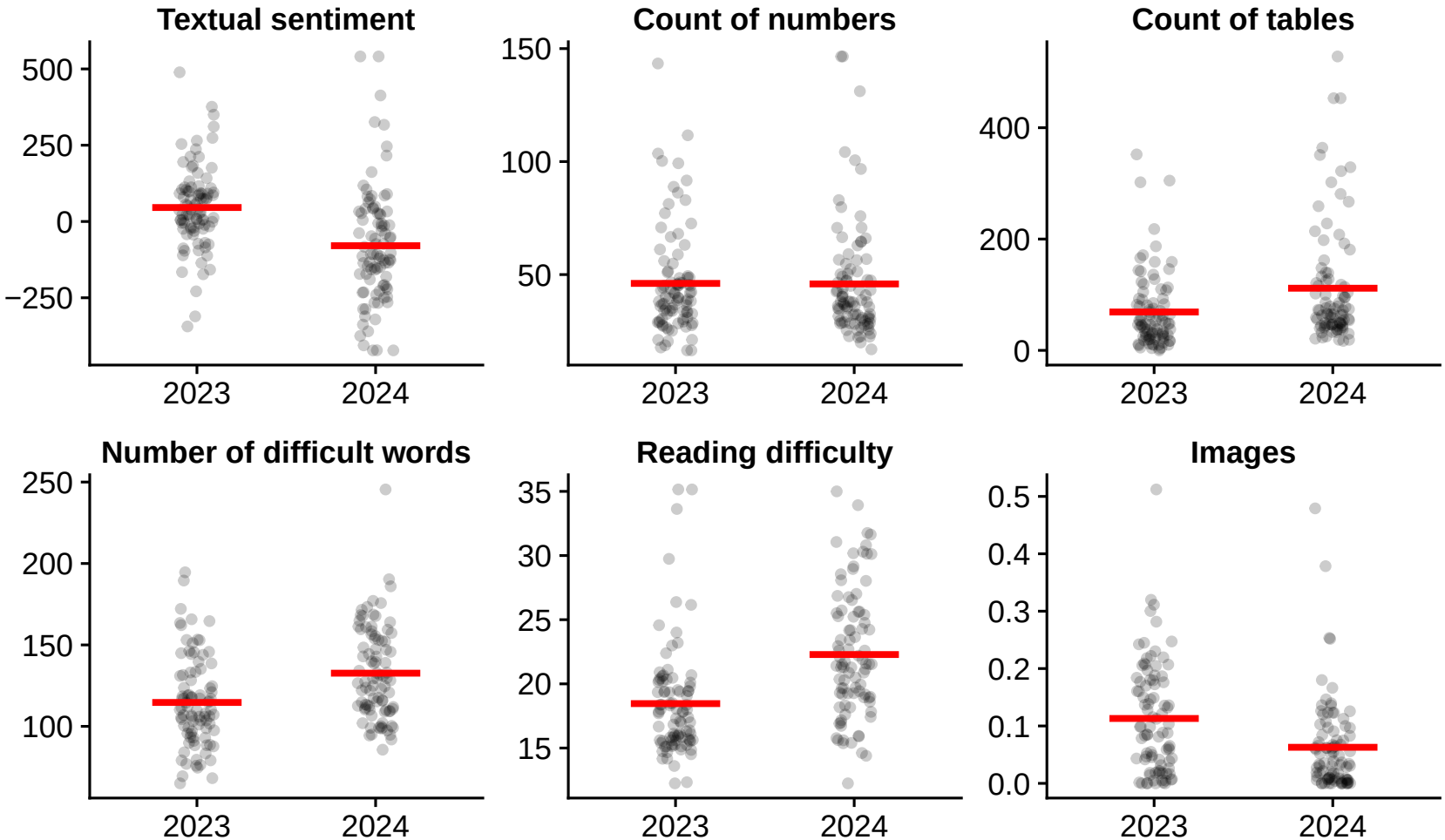

Fig. 6: Change of textual attributes in sustainability statements. This figure shows the development of six textual attributes over time. Every point represents one observation, the red horizontal line is the mean. *Textual sentiment* is the number of positive minus negative words following Loughran and McDonald (2011), *Count of numbers* and *Count of tables* is the count of numerical information and tables in the reports, *Number of difficult words* is the count of words that have more than two syllables, *Reading difficulty* is the Fog−Index, an aggregate measure of readability where higher values indicate higher sophistication, and *Images* is the share of pages that is filled with images.

Appendix: Finding topics in sustainability reports

- We rely on an established method in accounting and finance research (e.g., Li et al. 2021, Lin et al. 2024) to extract topics from unstructured text.
- This method utilizes *word2vec* (Mikolov et al. 2013), an algorithm that "learns" the meaning of words in a text using a neural networks.
- We use the resulting textual embeddings to generate a dictionary of keywords for each ESRS. Based on general "seed words" (e.g., greenhouse gas emissions for E1 climate change), we pick the 500 most similar words based on the embeddings.
- The resulting list of keywords allows us to broadly capture ESG-related discussions in reporting even before ESRS-specific terminology has been introduced.
- The main measure shown in this presentation is the number of words from sentences that contain a keyword from one of the 11 ESRS standards.

Feedback and questions?

We are interested in your feedback and happy to answer any questions regarding our analyses.

Reach out to Victor Wagner (victor.wagner@lmu.de) or visit our homepage!