A benchmark dataset for detecting frames in multi-topical news content: Online Appendix

A benchmark dataset for detecting frames in multi-topical news content: Online Appendix

Coding scheme (Semetko & Valkenburg)

Attribution of responsibility

- Does the story suggest that some level of government has the ability to alleviate the problem?
- Does the story suggest that some level of the government is responsible for the issue/problem?
- Does the story suggest solution(s) to the problem/issue?
- Does the story suggest that an individual (or group of people in society) is responsible for the issue-problem?
- Does the story suggest the problem requires urgent action?

Human interest frame

- Does the story provide a human example or "human face" on the issue?
- Does the story employ adjectives or personal vignettes that generate feelings of outrage, empathy-caring, sympathy, or compassion?
- Does the story emphasize how individuals and groups are affected by the issue/problem?
- Does the story go into the private or personal lives of the actors?
- ~~Does the story contain visual information that might generate feelings of outrage, empathy-caring, sympathy, or compassion?~~

Conflict frame

- Does the story reflect disagreement between parties/individuals/groups/countries?

- Does one party/individual/group/country reproach another?
- Does the story refer to two sides or to more than two sides of the problem or issue?
- Does the story refer to winners and losers?

Morality frame

- Does the story contain any moral message?
- Does the story make reference to morality, God, and other religious tenets?
- Does the story offer specific social prescriptions about how to behave?

(Economic) consequences frame

- Is there a mention of financial losses or gains now or in the future?
- Is there a mention of the costs/degree of expense involved?
- Is there a reference to economic consequences of pursuing or not pursuing a course of action?

Multiverse analysis of all methods (based on the ground truth)

K-Means + TF-IDF.

Principal Component Analysis + TF-IDF.

LDA.

STM.

ANTMN.

Seeded-LDA.

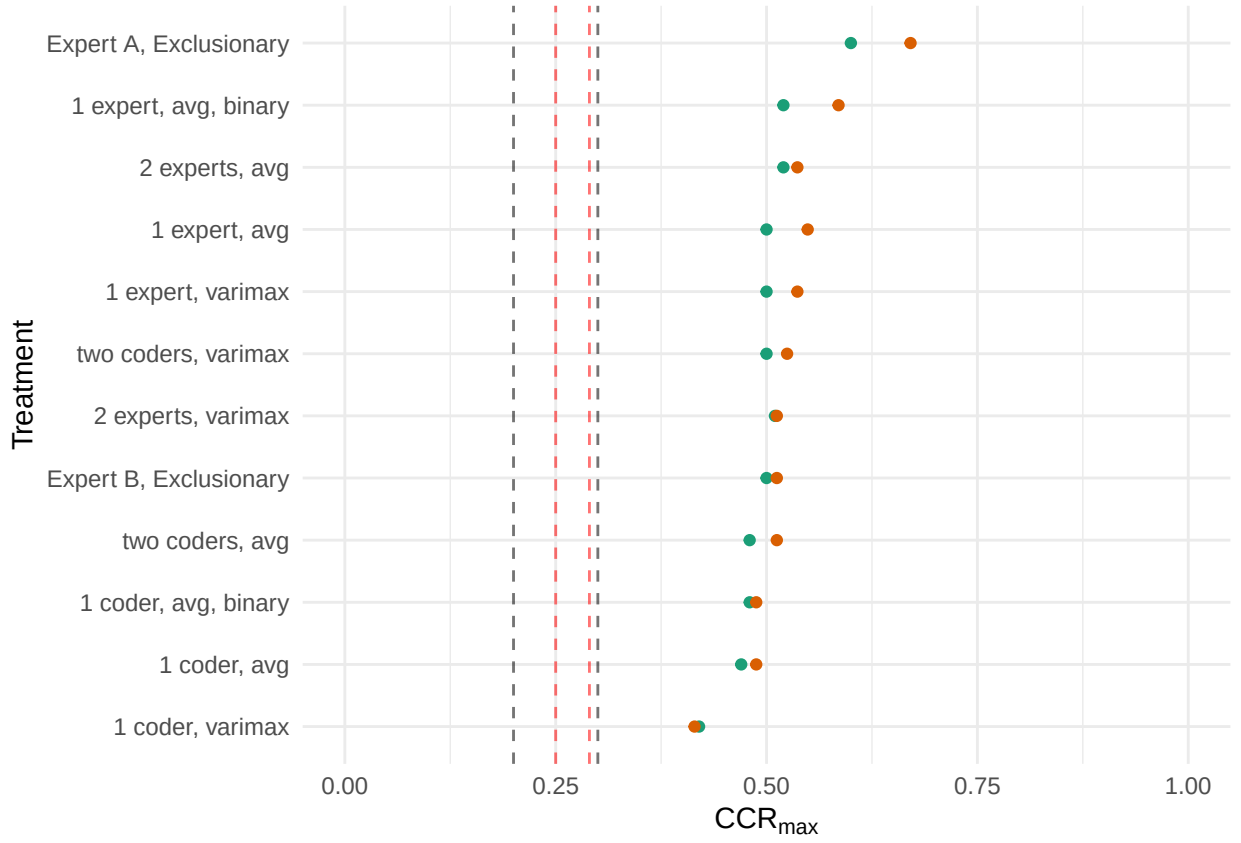
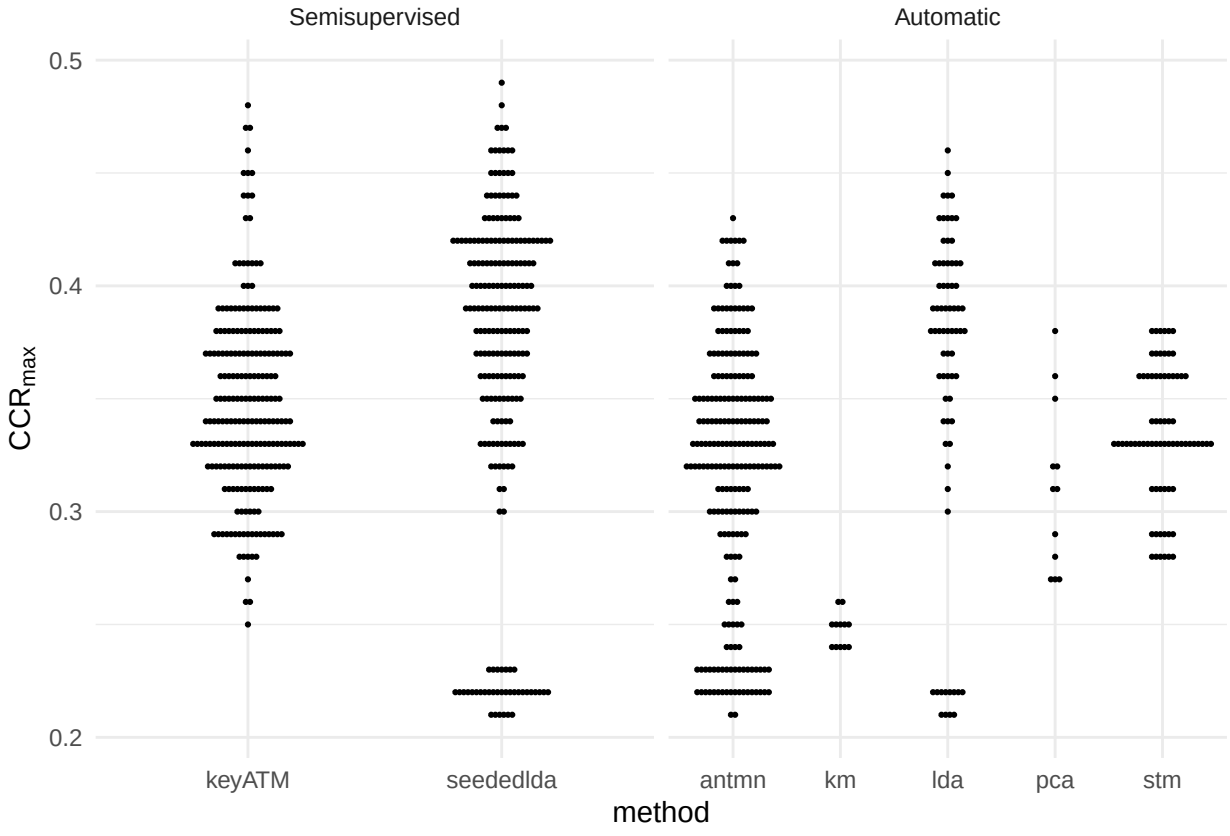


Figure 1. Multiverse analysis of human coding

Visualization of the variance



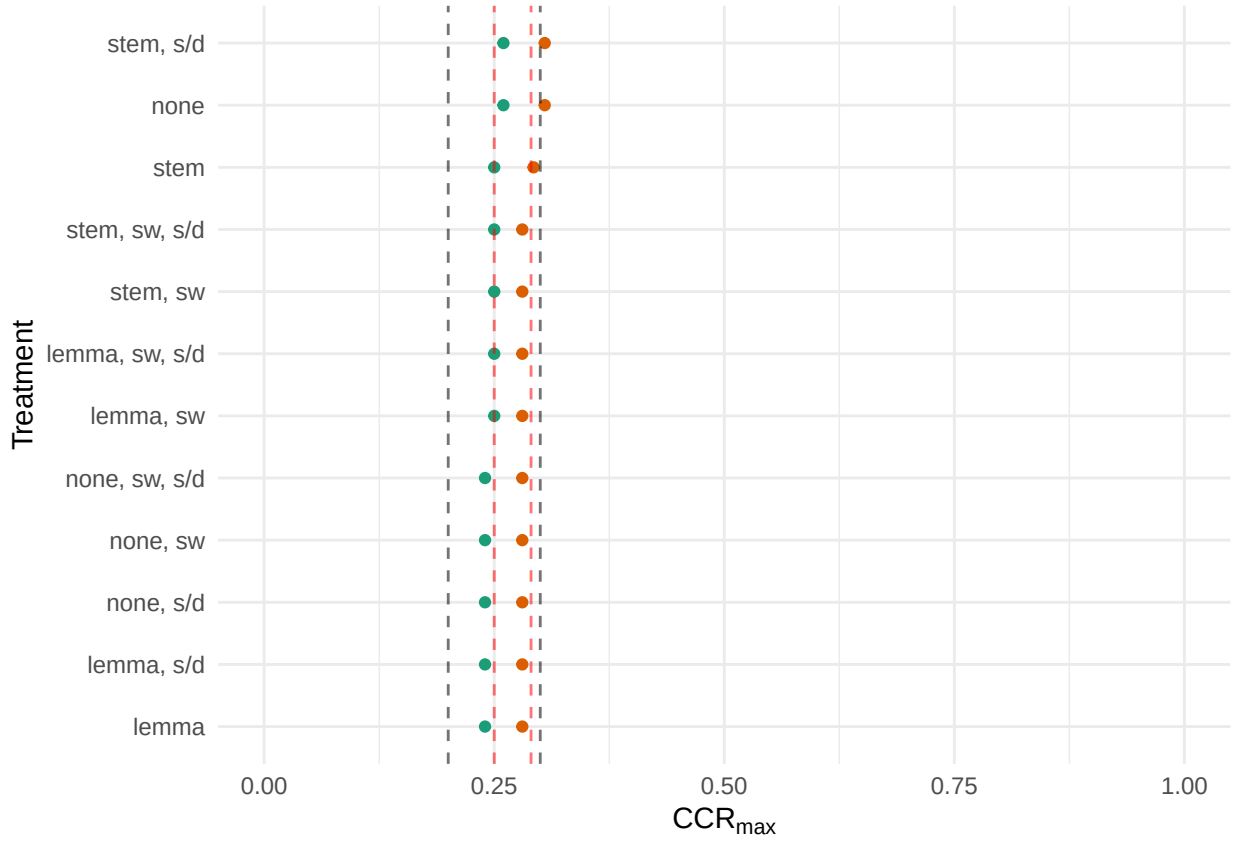


Figure 2. Multiverse analysis of K-Means

Comparing confidence level of correct and incorrect expert coding

We modeled the correctness of expert coding ("F1" is equal to the ground truth) and confidence level ("F2"), while adjusting for individual differences between the two experts using Bayesian multilevel logistic regression analysis. The following is the robust conditional effect plot. There is no evidence to suggest that there is a trend. Therefore, experts can either confidently give correct and incorrect coding.

Simulation of increasing sample size

In this analysis, we simulated the possible outcome of increasing the sample size on the multiverse analysis.

From our 100 articles, we created further synthetic articles following the principle of

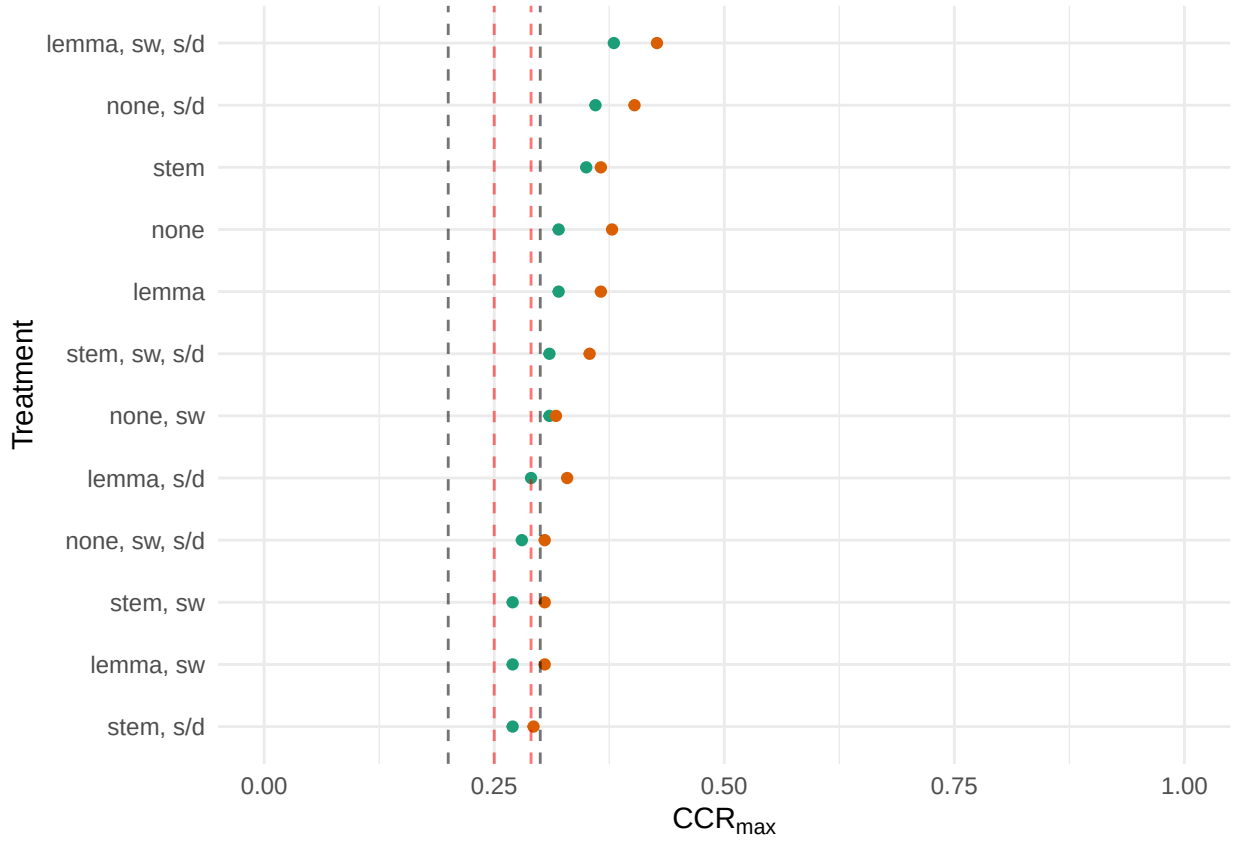


Figure 3. Multiverse analysis of PCA

bootstrapping. We synthesized more articles based on the following algorithm:

1. Randomly select one article
2. Tokenize this article into its n sentences
3. From these n sentences, randomly draw n sentences from these sentences with replacement. Therefore, one sentence can appear more than once.
4. Concatenate these randomly drawn n sentences into a synthetic article, assign this article with the same topic and frame as the original article

We repeat the above process for 500, 1000, and 2000 times to generate 3 different corpora. This approach is compatible with the bag-of-words representation used in all unsupervised and semi-supervised methods because the word order is not considered. Also in step 3, topical and frame clues have the same natural chance of being selected. To put

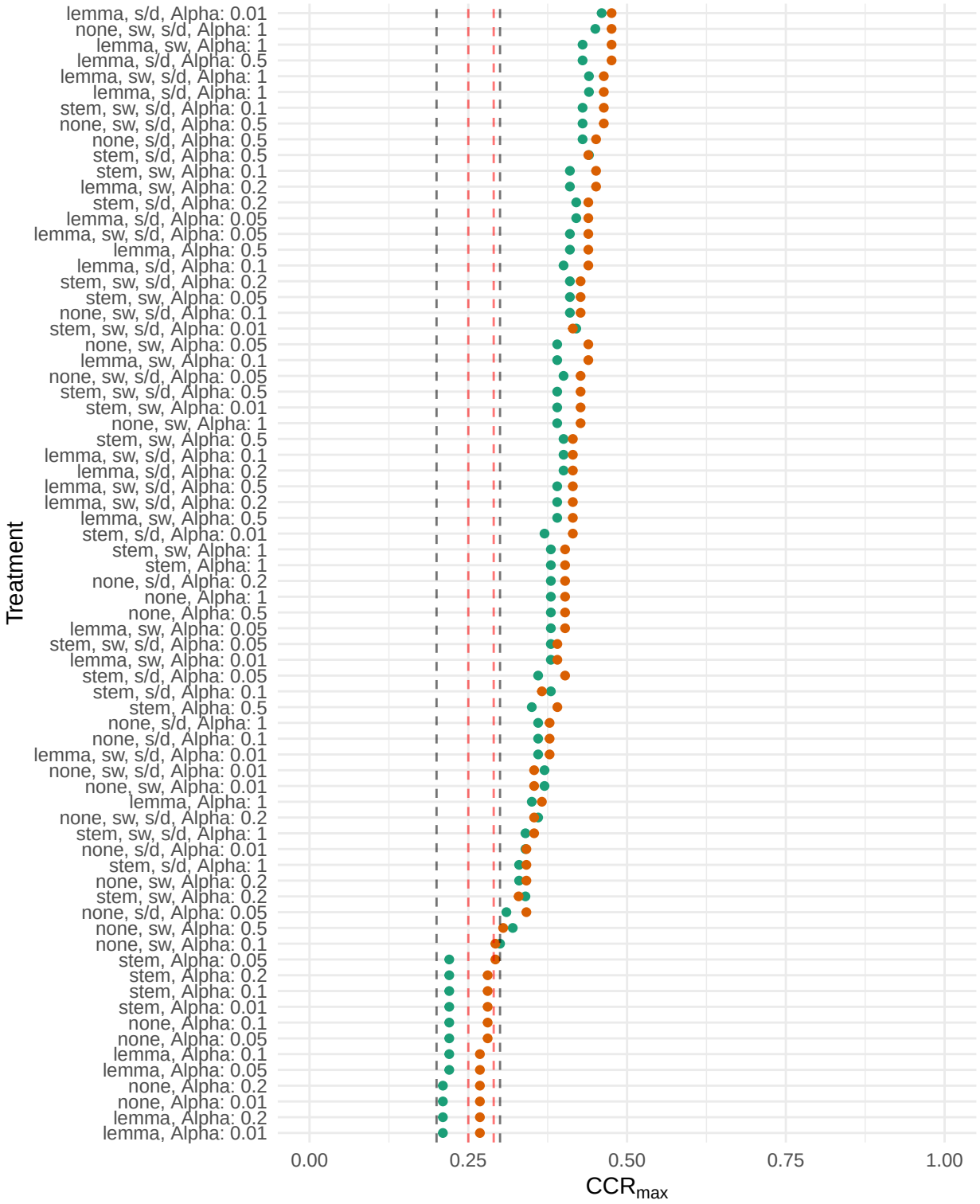


Figure 4. Multiverse analysis of LDA

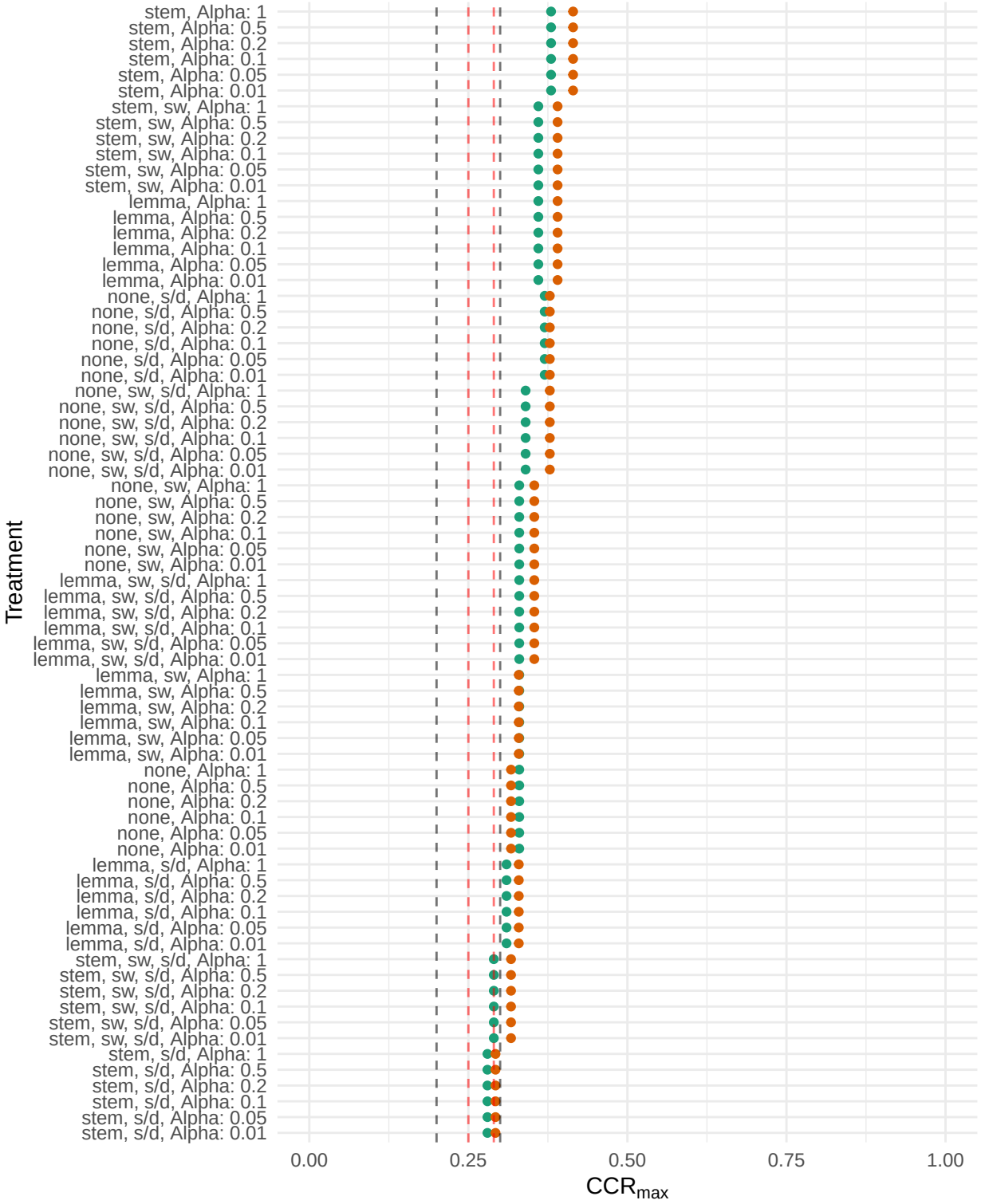


Figure 5. Multiverse analysis of STM

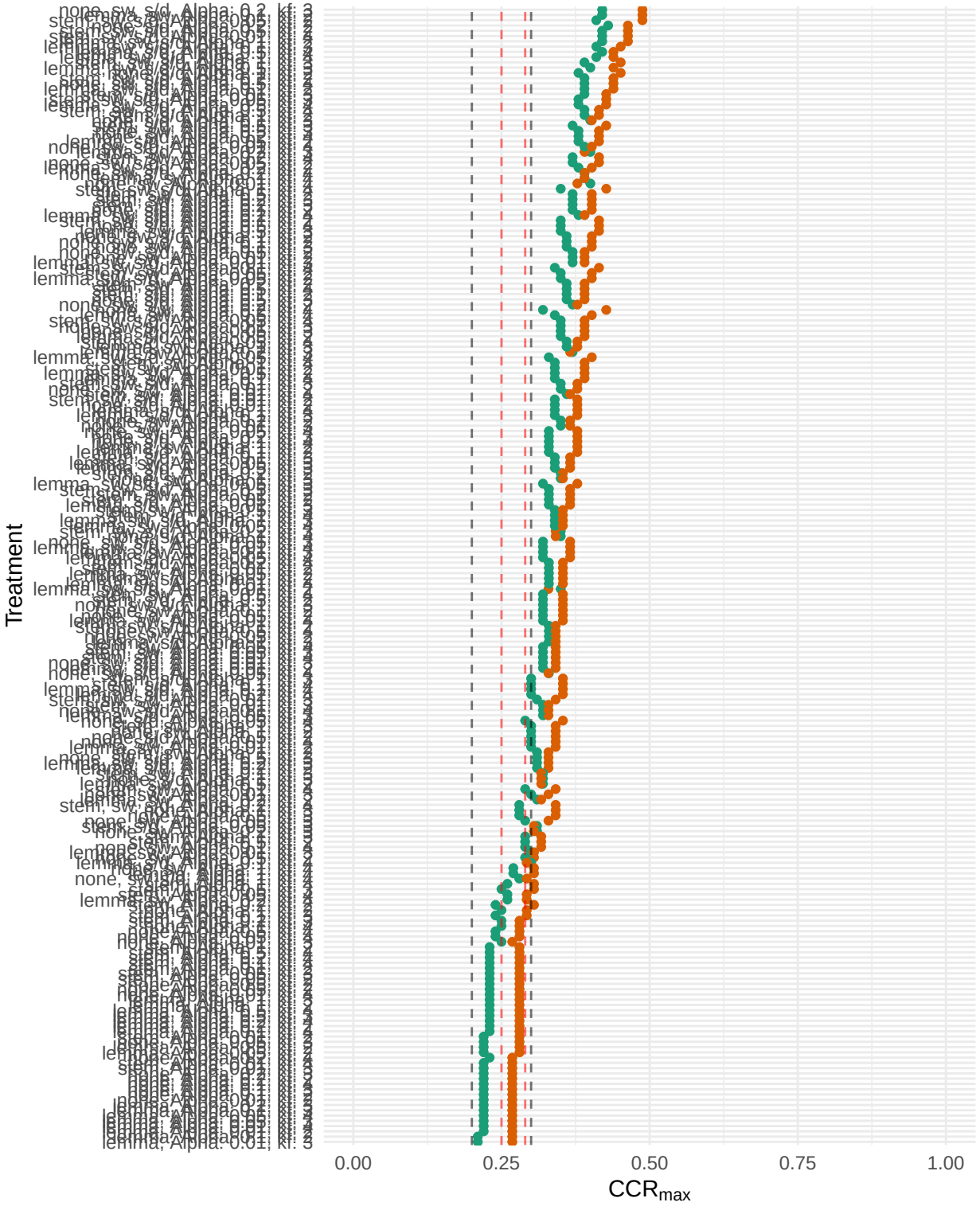


Figure 6. Multiverse analysis of ANTMN

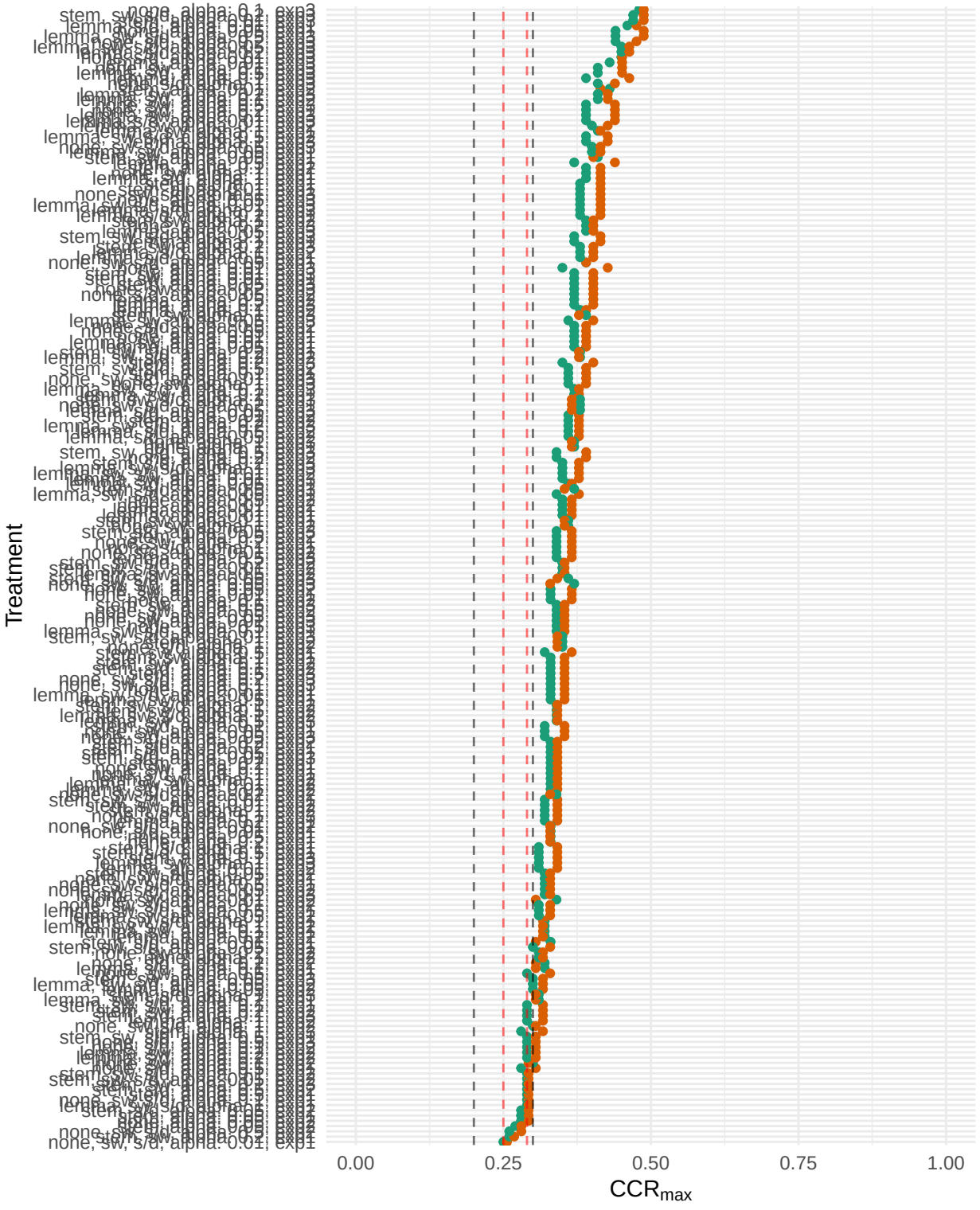


Figure 7. Multiverse analysis of seeded-LDA

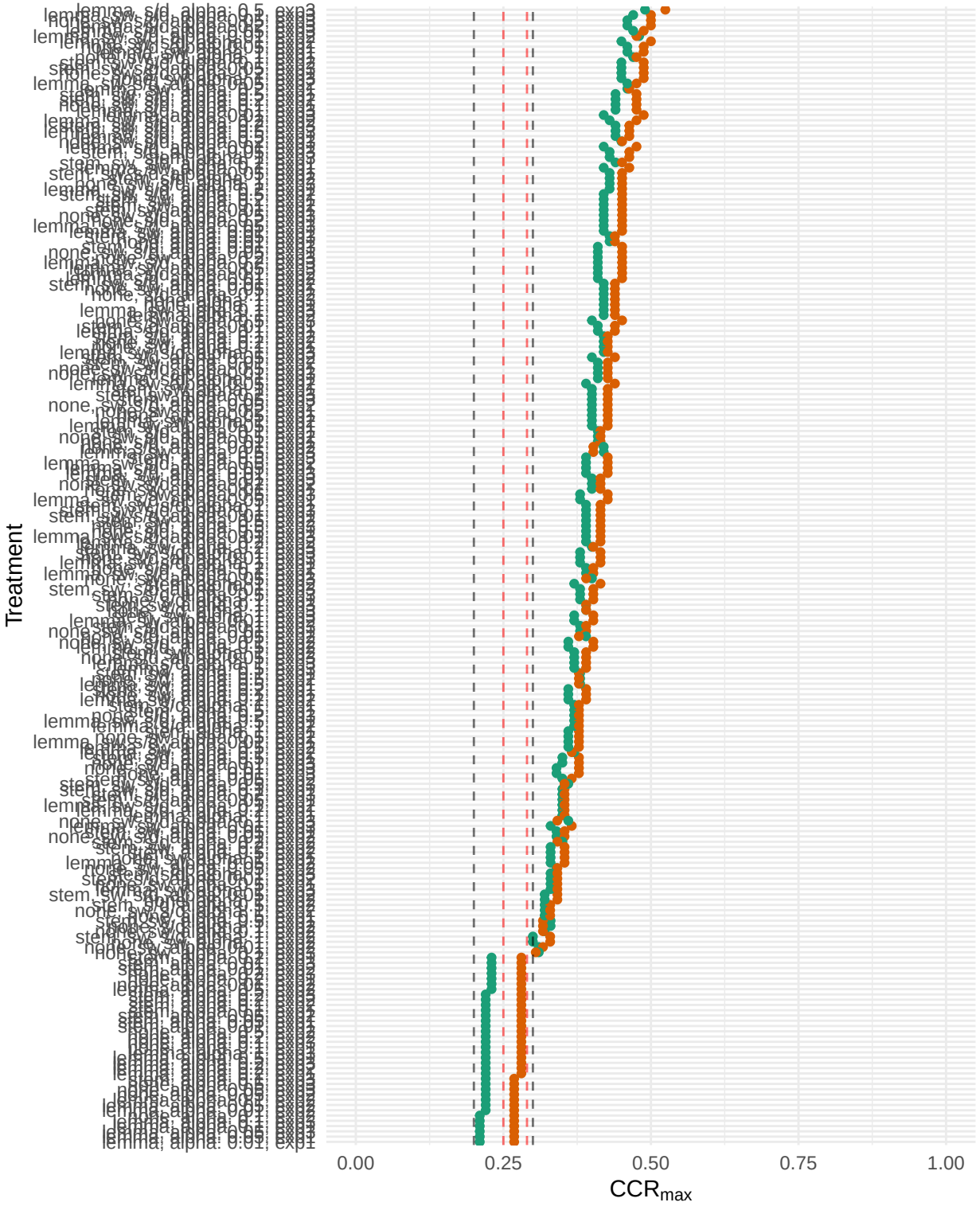


Figure 8. Multiverse analysis of keyATM

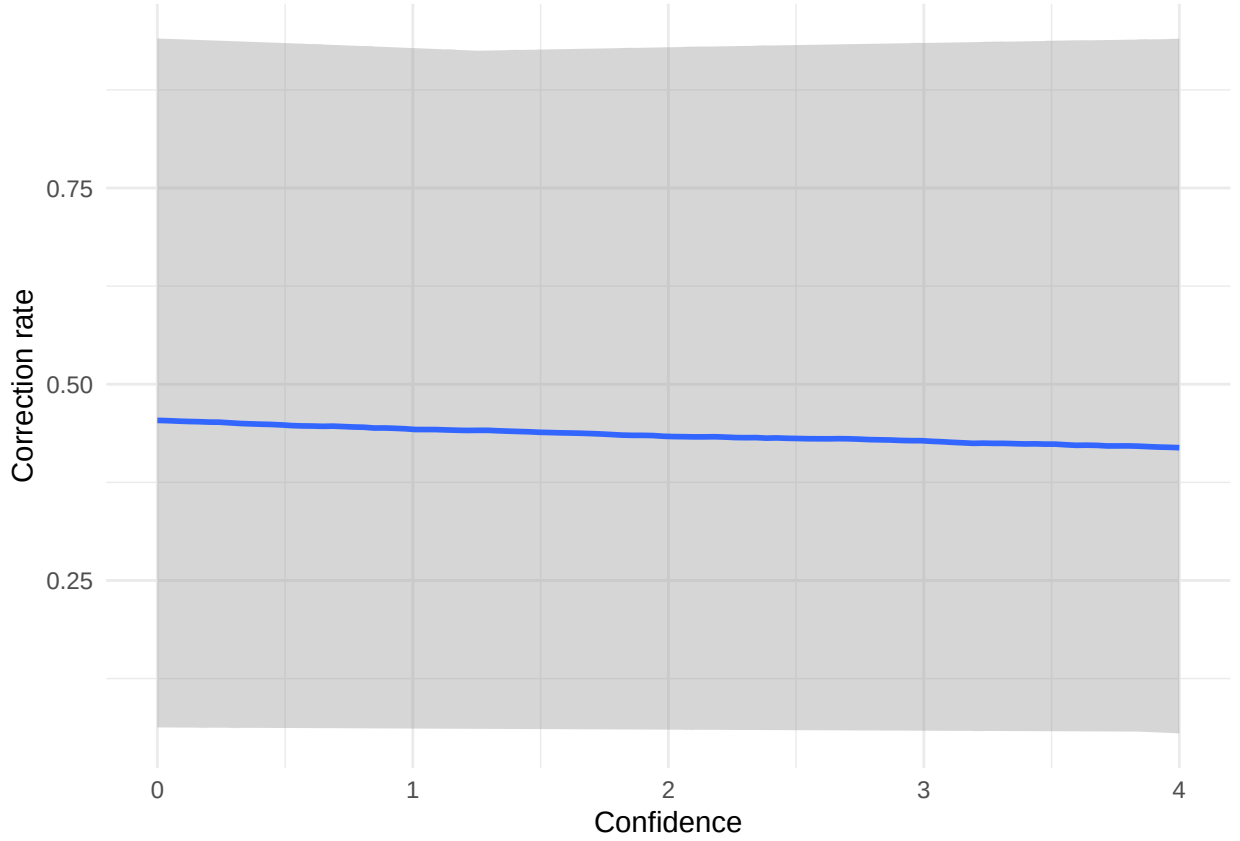


Figure 9. Robust conditional effects from the Bayesian model on the relationship between correction rate of expert coding

this simulation in another perspective, it simulates whether frames, rather than topics, are more likely to be picked up by these unsupervised and semi-supervised methods when the sample size increases.

The above figure shows the sorted CCR_{max} like the visualization of multiverse in the main text. All confidence intervals have been stripped for clarity as we are only interested in the point estimate. The treatments are the same as the main analysis and they are not displayed in the y-axes.

From this simulation, it is extremely unlikely that increasing the sample size would increase the CCR_{max} for all unsupervised methods. Instead, all methods, except KM and STM, perform more like the de-facto null (0.3) with the increasing sample size. Therefore,

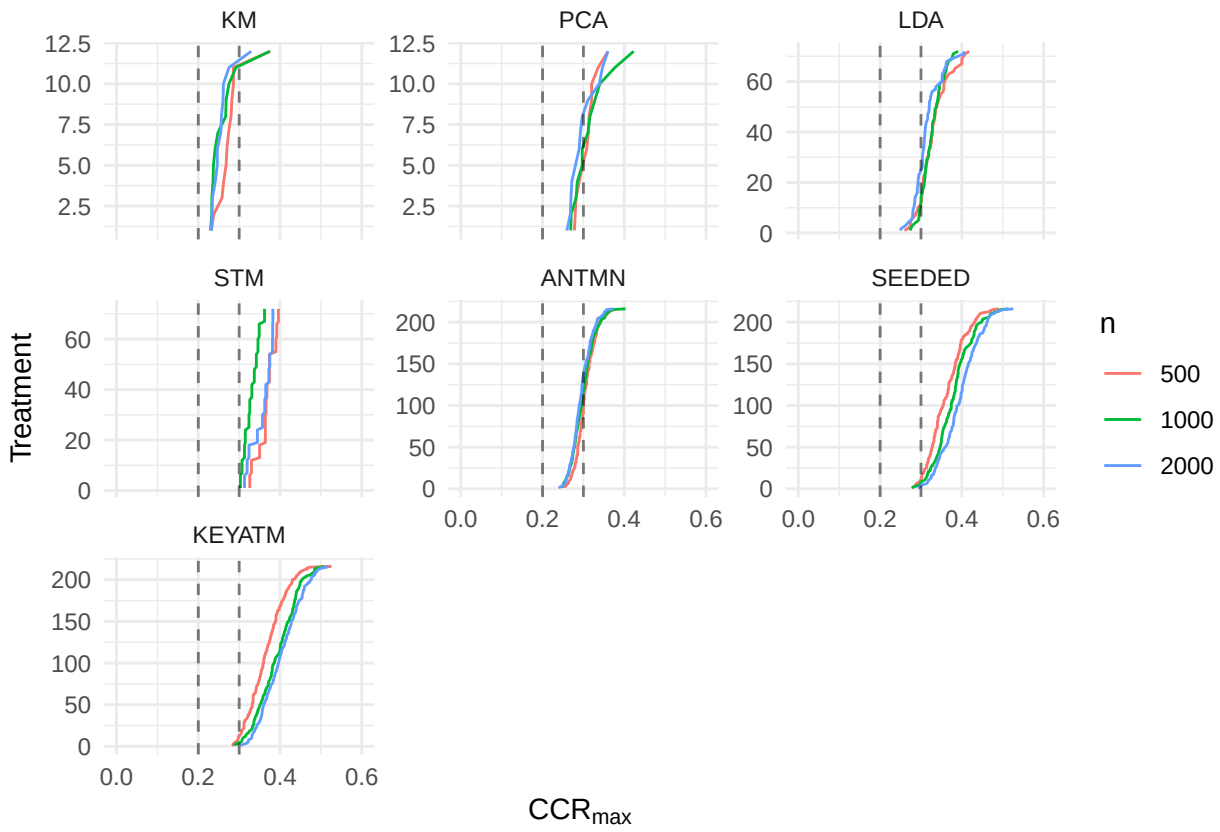


Figure 10. Multiverse analyses with different sample sizes: 500, 1000, 2000

these methods appear to be more keen to pick up **topics**, rather than frames, when sample size increases: the sorted performance curve rescinding towards the null value with the highest sample size. For the two semi-supervised methods, increasing the sample size appears to provide better resistance against the rescinding.

We also performed the same Bayesian analysis using the simulated data. By increasing the sample size, it appears that both H1 and H2 would become more significant, instead of the other way around.

Regression coefficients of the Bayesian model

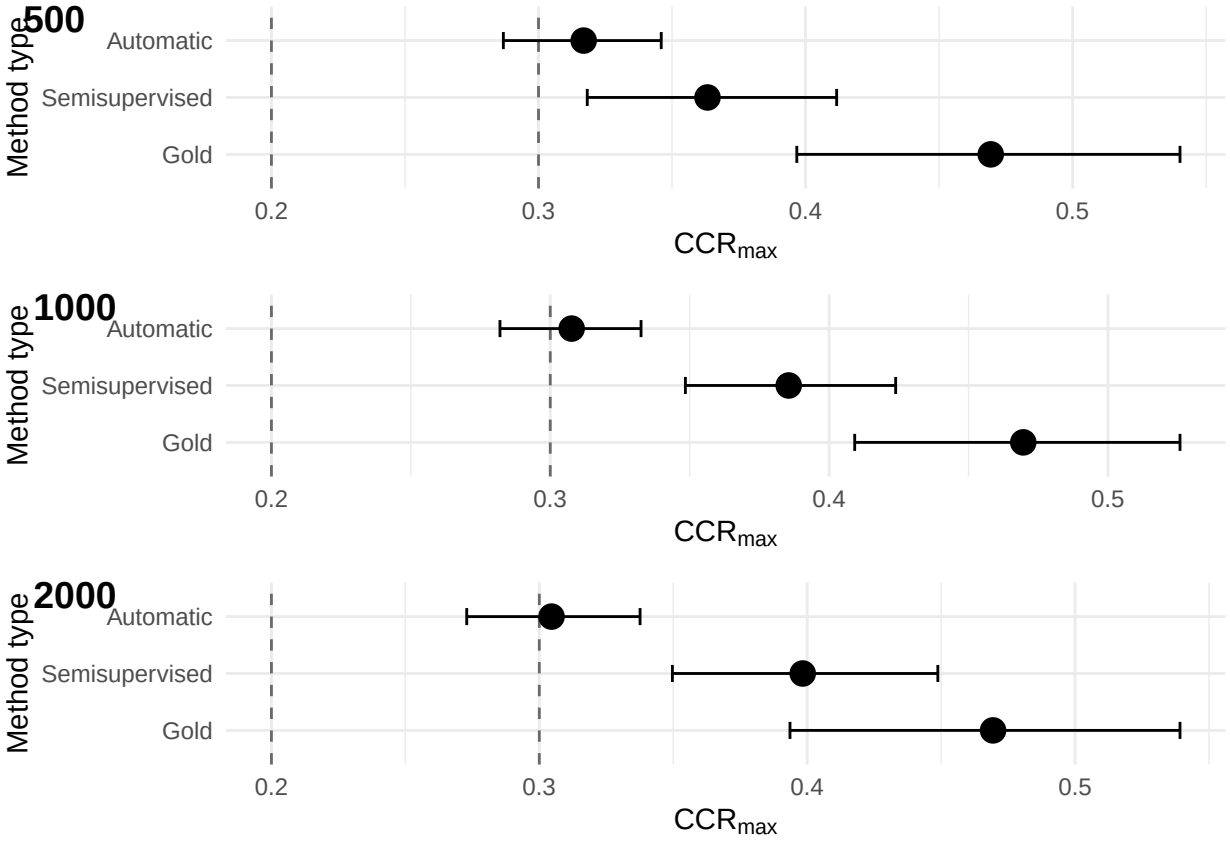


Figure 11. Robust conditional effects from the Bayesian model of the "gold standard", semi-supervised and automatic methods at 89% credibility with different simulated sample sizes

Table 1

 # Fixed Effects

Parameter	Median	89% CI	pd	Rhat	ESS
(Intercept)	0.47	(0.38, 0.56)	99.95%	1.004	1288.00
method_typeSemisupervised	-0.11	(-0.22, -2.93e-03)	94.80%	1.005	899.00
method_typeAutomatic	-0.15	(-0.25, -0.06)	98.05%	1.004	1263.00

Table 2

 # Sigma

Parameter	Median	89% CI	pd	Rhat	ESS
sigma	0.06	(0.06, 0.06)	100%	0.999	3220.00

$R^2 = 0.1446933$