

# Responses to Reviewer 1 Comments

Thanks to **Reviewer 1** for the valuable feedback that helped us to revise and strengthen our manuscript “Building Political Hashtag Communities: A Multiplex Network Analysis of U.S. Senators on Twitter During the 2022 Midterm Elections.”

Below, we list each reviewer comment and our response. Based on the reviewer’s comments, we have made specific revisions to the study. In the updated manuscript, we have made changes in red corresponding to our comments below.

In each notation below, the first number refers to the number of a reviewer, and the second number refers to the number of task (e.g., “Reviewer 1|3” implies this is the third task of reviewer 1).

**Reviewer 1|1:** How you evaluate the quality of ChatGPT clustering?

- **Author(s):** We included several references that demonstrate the superiority of GPT4 compared to other models in NLP tasks. We explained why GPT4 is one of the best tools available in topic modeling using hashtags under *hashtag topic modeling* subsection. We have also added information about our expert review of clusters created by ChatGPT.

**Reviewer 1|2:** The time span you investigate includes the election. In this regard, one may hypothesize that behavior patterns before and after the election can differ. Please address this point.

- **Author(s):** We thank the reviewer for pointing out this limitation. We definitely understand the reviewer’s concern regarding the potential variations in behavior patterns before and after the election period. The main reason why we did not incorporate a time-based analysis distinguishing pre and post-election periods was the primary focus of our study, which is on understanding how U.S. senators strategically use hashtags within an election environment. Accordingly, our decision is grounded in the concern that broadening the scope to include both pre and post-election periods may compromise the clarity and depth of our analysis, deviating from the core objective. Currently, we do not have any theoretical explanation or prediction for this topic, despite its importance. However, we acknowledge the validity of the reviewer’s point and concur that analyzing behavior patterns before and after the election could be a promising avenue for future research. In the conclusion section of our revised manuscript, we now explicitly address this limitation, highlighting the absence of a theoretical explanation for this temporal distinction in our study and providing suggestions for future research directions that could explore these dynamics in more detail. See lines 551-58.

**Reviewer 1|3:** Table 1. The last column (# of clusters). What are these clusters about? Please give more detailed explanation. In Subsection 4.2, the authors talk about hashtag clustering. If the corresponding column stands for hashtag clusters than I do not fully understand why the number of clusters is of the same order than the number of hashtags. Do we need such a clustering?

- **Author(s):** In response to this critique, we have made significant revisions to eliminate any confusion around the term "cluster" in our text. Specifically, we have removed the repetitive use of "cluster" to clarify its meaning and context within our study.

Regarding the row in Table 1 that previously mentioned clusters in the context of cleaning and merging hashtags, we have now removed this row for greater clarity. Instead, we have integrated this information into the text of our article. This integration aims to provide a more detailed and clear explanation of how we handled the hashtags in our dataset, ensuring that the concept of clustering is well-understood and logically consistent throughout the paper.

**Reviewer 1|4:** Please introduce Jeffrey degree and Pearson degree (if I am correct, the capital letters are needed - see lines 189, 190).

- **Author(s):** We welcome this suggestion. As the reviewer suggested, we now introduced these metrics and clarified how they differ. The revised paragraph (in red) can be tracked in section 3.3.

**Reviewer 1|5:** Figure 3. I suppose that for effective comparing of distributions, one need to extend the figure by adding relative quantities. That is, what is the fraction of the edge Economy among the Democrat party? Further, I suppose that proportion confidence intervals should be included.

- **Author(s):** Thank you to the reviewer for the insightful suggestion. In response, we have incorporated two additional tables in the Appendix (see Table A1 and A2), providing a detailed distribution of edges across layers for each sub-category, accompanied by proportional confidence intervals. To guide the reader to the newly added tables, we have included an additional sentence in the second paragraph of Section 5. Nonetheless, regarding the figure in the main text, we would like to maintain the use of counts. Our decision was not arbitrary. We believe counts offer a more nuanced representation of the varying magnitudes of networks captured by different topic-modelling and sub-group strategies. However, if the reviewer and editors see greater value in using proportions, then we are happy to be overruled.

**Reviewer 1|6:** Figure 4. Please add a color bar and the correlation values.

- **Author(s):** Figure 4 has been reorganized. Values and color bars were added. In the new version, Pearson Degree scores are in the first row and Jeffrey Degree correlations are in the second row. The figure note has been revised accordingly. Note that the new grouping of the order of plots forces us to visualize non-detected networks (see labor column/row for Republicans). That means no reciprocal hashtag involvement is detected for this layer for Republicans. We already discussed these findings in our results section, but an explanation is also added to the Figure 4 note.