

cover-memo

Suyeol Yun

Hi everyone, first of all, thank you for taking the time to read this draft. I know how busy everyone is, so I genuinely appreciate your efforts to review this piece and share your comments.

As you know, this paper aims to identify whether congressional investment occurs. My first reader is Prof. Insong Kim, and my second reader is Prof. Teppei Yamamoto.

- My first reader has high expectations and wants to ensure I'm a valid political science PhD candidate. They have clearly stated the conditions for passing this SYP, which I must meet to avoid failure:

1. A clearly stated hypothesis that you are testing
2. Presenting evidence for/against the hypothesis.
 - a. Make sure to draw an appropriate **statistical** inference rather than descriptive/predictive inference.
3. Motivation, Framing Theory - fully engaging with the literature, especially engaging with every American politics piece if related.
 - a. Do not use any presumed knowledge without citation while introducing motivation and developing your empirical strategy.
 - b. Motivation, question, hypothesis, and empirical design should all be well-positioned within the previous literature.

I am specifically asking you to evaluate these points with a critical eye and provide comments and advice on which literature I should engage with and at which points the paper might fail to satisfy these requirements.

- Currently, the paper covers Motivation, Framing, and Empirical Strategy but has not proceeded to compute the causal quantity of interest. It took a long time to merge all

the relevant data, but I am confident that the data portion is now complete.

- At the moment, I plan to estimate CATE on how committee assignment affects the similarity between a senator and committee. This task is exciting because controlling for possible confounders based on learned representation is something I have always wanted to try. As far as I can tell, this has not been attempted before, making it novel but also potentially risky.
- Therefore, I would like to get some advice on what kind of "other" empirical strategies can be adopted to prove congressional investment before moving into this complex method. I tried to estimate ATT using DiD; however, only a small number of senators change their committee pre/post-treatment, so I gave up this approach.
- My first reader also wants to start with a causal quantity that resides in the shared understanding within the range of methodology we learn from Quant 1, 2, and 3 - then move on to more complex designs.
- My primary motivation for engaging with Graph Neural Networks (GNN) stems from my previous experience leveraging the power of GNNs while working at eazel Inc. There, I successfully utilized GNNs to learn effective representations for each artist, enabling the classification of emerging artists and subsequently guiding the company's investment in artworks. Similarly, I want to emphasize the importance of using graph-structured data as opposed to tabular data, which tends to overlook relationships between variables, while graph-structured data is capable of encapsulating such relationships. I hope this point is effectively conveyed to readers. If you have suggestions for a better way to present this argument or thoughts on the gap between tabular data and graph-structured data utilization in political science, I would appreciate your input. Additionally, any relevant citations are welcome. I contend that the majority of social science data is inherently graphical, as it encodes interactions between different types of entities on a large scale.

As you all know, I appreciate more straightforward comments rather than indirect and polite ones. Also, please feel free to ask for clarification on unfamiliar concepts, and I will be happy to explain my understanding of them.