Referee Report
Southern Economic Journal, Manuscript 2022074

"Gerrymandering in the Laboratory"

1 Summary

This paper examines gerrymandering (strategically drawing boundaries of electoral districts in an attempt to influence election outcomes) and individuals' preference for / willingness to engage in gerrymandering. The basic theoretical framework described is a two-stage model; the electoral map is drawn in the first stage and then, in the second stage, the parties compete via simultaneous Tullock contests (one per district in the electoral map) to try to win over undecided voters.

It is demonstrated that, from the individual's perspective, gerrymandering is optimal; players want to pack rival partisans into one district while dividing their own partisans amongst the remaining districts. The results of a laboratory experiment reveal that participants overwhelmingly engage in gerrymandering (by selecting the map which gives them an advantage in the second stage contest) despite the fact that nearly all report being opposed to the practice of gerrymandering.

2 Evaluation

Overall, the paper is nice and provides some interesting findings on individuals' willingness to engage in gerrymandering–which is a nice contribution to that literature. The model and theoretical results are well presented and provide a clear benchmark against which to compare the experimental data. The experimental design seems well thought out and well suited to test the predictions. However, the sample size ($n = 64$) seems quite small, especially considering that individuals are rematched across rounds, and therefore, there are only eight total independent observations (sessions).

The analysis of the experimental data feels a bit underwhelming at times. While the current analyses all appear correct and provide some interesting results, it seems like there is much more that could be done with the existing data. In particular, the analysis of individual investment behavior in the contest stage seems as though it could be more extensive. More specific comments are provided in the following section.

3 Feedback

3.1 Main Comments

1. Regarding the theoretical prediction in the $Sym_{3,1}$ map, the equilibrium described is symmetric but there is no evidence provided that this equilibrium is unique. This leaves open the question of whether there may be additional, asymmetric equilibria (which would make evaluating the experimental evidence somewhat more complicated). I believe Klumpp and Polborn (2006) would be helpful in resolving this.

2. Regarding experimental evidence of over-investment (discussed on page 12): is this behavior widespread across participants or driven by a handful of participants wildly over-investing? Does the pattern of over-investment differ by map? Some additional, individual level analysis here would provide a clearer picture of how individuals invest in these contests.

3. Looking at percent over-investment (using Table 2 on page 12) and comparing across maps in cases where equilibrium investment is positive, average expenditures are nearly double the predicted value in all cases except $Sym_{1,3}$, where it is less than 50% higher than predicted. It's unclear to me what is driving this, but it seems that subjects may view this map differently in some way. The difference in overbidding also raises questions about overspreading (see, for example, Chowdhury et al. (2014)).

4. Following up on the previous comment, if it is the case that subjects view $Sym_{1,3}$ as fundamentally different in some way, it raises the question of whether they also treat this map differently in the map selection stage. According to Figure 6, it is the second most commonly selected map in Stage 2 (though, unclear if the rate at which it is selected is significantly higher than other maps), and according to Figure 8, it is the most commonly selected map in Stage 3. Can this be explained by differences in investment behavior in the resulting contests?

5. Regression results presented in Table 3 (also, Table 7 in Appendix B) seem to be comparing total expenditures across maps in all districts. This seems a little difficult to interpret since in all maps except $Sym_{3,1}$ the equilibrium involves investing zero in two of the three districts. Are the results similar if restricted to comparing only the districts where positive investment is predicted? What about districts where zero investment is predicted? What about the patterns of over-investment in districts with positive predicted investment compared to those where the predicted investment level is zero?

6. In the analysis of stage 2, the text states that it is assumed that a subject's modal response indicates their preferred map. Were there any cases where a participant indicated three different maps? If so, how are these handled? More generally, could you use all three rounds to differentiate between a strong preference for gerrymandering (selected advantaged map all three times), weak preference (selected advantage two of the three times), etc.?

7. While the primary focus of this study is gerrymandering, employing a Tullock contest in the second stage also relates it to existing contest literature. There is a fair amount of prior work on multi-battle contests that is not mentioned (see, for example, section 5 of Dechenaux et al. (2015) and work cited within). Discussing this literature and describing how the present study is related would help to better tie this study to the existing contest literature.

3.2 Minor Comments

1. (Page 6) Clarification: On page 6, the authors say "...of the three symmetric maps $Sym_{1,1}$ is socially optimal and yields the greatest expected profit to the players." A more precise statement about what it means for an outcome to be socially optimal would be helpful. In this context (elections), it is not quite clear to me what it means for an outcome to be *socially optimal*. It is true that this map yields the greatest expected payoff to the players (political parties), but it seems that *social* optimality should also take into account voters' welfare. Since voters are not modeled directly in this environment, it seems fine to side-step this but some additional discussion here would be nice.

2. (Page 7) Typo: "...Player B has a 25% chance of winning the contest and en expected profit of zero." Should be "an".

3. (Page 8) Typo: "That is, the strategy method was used...preferred choices for all 3 district on all 5 maps each period." Should be "districts".

4. (Page 10) Clarification: The text states that participants "went through several practice contests." It would be helpful to clarify how many such practice rounds took place and whether these were against other participants or bots. If against other participants, is there any reason to believe that this might affect beliefs about how others play a contest, which could in turn affect the contests in the experiment?

References

Chowdhury, S. M., R. M. Sheremeta, and T. L. Turocy (2014). Overbidding and overspreading in rent-seeking experiments: Cost structure and prize allocation rules. *Games and Economic Behavior 87*, 224–238.

Dechenaux, E., D. Kovenock, and R. M. Sheremeta (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics 18*(4), 609–669.

Klumpp, T. and M. K. Polborn (2006). Primaries and the new hampshire effect. *Journal of Public Economics 90*(6-7), 1073–1114.