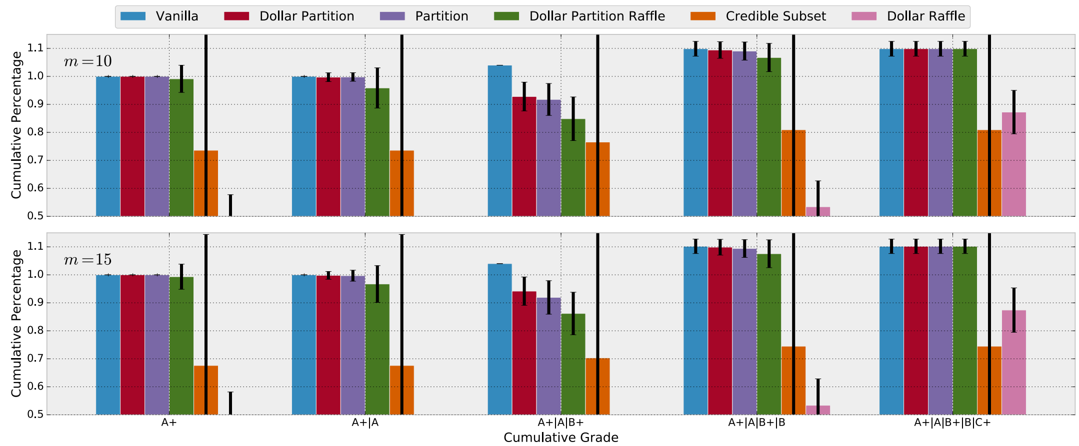

Figure 1: Mean cumulative percentage of each grade of agent selected by the six peer selection algorithms presented in this paper on 1000 random iterations selecting $k = 25$ agents from a population of $n = 130$ agents providing $m = 10$ (top) and $m = 15$ (bottom) reviews divided into $l = 5$ clusters with a Mallows dispersion $\phi = 0.1$. To enable comparisons, every mechanism selects $|W|$ equal to that of Dollar Partition; hence the ≥ 1.0 averages as $k = 25$ is the denominator. Error bars represent one standard deviation from the mean. Dollar Partition selects more agents from a higher grade more often, selects more agents from a higher grade in the worst case, and does so more consistently, than any other strategyproof mechanism. To highlight Partition and Dollar Partition we have cropped results where they are the same (cutting off Dollar Raffle).

General Results: We explored a realistic part of the large parameter space to investigate the mechanisms. The practical upshot, after running hundreds of thousands of instances, is that there are numerous tradeoffs that system designers must consider, critically depending on their target domain. In general, varying other parameters, such as k, ℓ, D and F did not change the ranking of mechanisms shown here. However, increasing the number of clusters improved Dollar Partition's performance in comparison to Partition's, which may stem from the increased chance that Partition will select the bottom candidates of a given cluster instead of better ranked candidates in a different cluster. Accordingly, as it generally selects the top candidates, Partition's performance improves when scoring rules are exponential in comparison to less extreme scoring rules, such as Borda.

Dollar Partition is much better when there is sufficient information, in terms of the number of reviews and the granularity of the grades, to have a chance of recovering the ground truth ordering. Settings like conferences with $n = 2000$ papers and $m = 5$ reviews split into 5–8 grades often have no clear cutoff between accept and reject; the grades contain too many items. In these cases all the mechanisms perform poorly, as selecting a set of winners is akin to randomly selecting agents from the set of possible winners. See, e.g., the NIPS experiment[9] and the recent paper on the limits of noisy rank aggregation using data from the KDD conference (**?**). As the ratio of m to n grows, and the granularity of the grades increases, it becomes possible to recover the ground truth ranking, and Dollar Partition outperforms the other mechanisms.

[9] http://blog.mrtz.org/2014/12/15/the-nips-experiment.html

6 Conclusion

We introduce a novel peer selection mechanism—Dollar Partition. Overall, Dollar Partition's flexibility in setting the number of agents to be selected from each cluster addresses the worst-case instances where partitions may be lopsided, allowing Dollar Partition to reach higher quality, more consistent results than existing mechanisms. Combined with the ability to always return a winning set, it is an improvement over current mechanisms.

Among strategyproof mechanisms, Partition and Dollar Partition may have a certain 'psychological' advantage: they may incentivize agents to report truthfully because an agent's contribution in selecting other agents (with whom he is not competing) is more direct. Moreover, partitioning into groups helps deal with conflict of interest cases, when there is fear of collusion among several agents; putting them in the same cluster prevents them from influencing one another's chance of success. Peer selection is a fundamental problem that has received less attention than voting rules. We envisage the need to develop robust solutions with good incentive properties, as these are widely applicable in large-scale, crowdsourcing settings.

Acknowledgments

Data61 (formerly known as NICTA) is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program. This research has also been partly funded by Microsoft Research through its PhD Scholarship Program, and by Israel Science Foundation grant #1227/12. This work has also been partly supported by COST Action IC1205 on Computational Social Choice.

Aperiam hic odit architecto nisi cumque unde libero, quia suscipit eos quasi unde autem deleniti ea placeat quas ex voluptate, qui sed ex odit repellendus iste fuga voluptas ea laborum voluptates pariatur?