D-efficient or deficient? A robustness analysis of stated choice experimental designs

This paper is motivated by the increasing popularity of efficient designs for stated choice experiments. The objective in efficient designs is to create a stated choice experiment that minimizes the standard errors of the estimated parameters. In order to do so, such designs require specifying prior values for the parameters to be estimated. While there is significant literature demonstrating the efficiency improvements (and cost savings) of employing efficient designs, the bulk of the literature tests conditions where the priors used to generate the efficient design are assumed to be accurate. However, there is substantially less literature that compares how different design types perform under varying degree of error of the prior. The literature that does exist assumes small fractions are used (e.g., under 20 unique choice tasks generated), which is in contrast to computer-aided surveys that readily allow for large fractions. Further, the results in the literature are abstract in that there is no reference point (i.e., meaningful units) to provide clear insight on the magnitude of any issue.

Our objective is to analyze the robustness of different designs within a typical stated choice experiment context of a trade-off between price and quality. We use as an example transportation mode choice, where the key parameter to estimate is the value of time (VOT). Within this context, we test many designs to examine how robust efficient designs are against a misspecification of the prior parameters. The simple mode choice setting allows for insightful visualizations of the designs themselves and also an interpretable reference point (VOT) for the range in which each design is robust. Not surprisingly, the D-efficient design is most efficient in the region where the true population VOT is near the prior used to generate the design: the prior is $20/h and the efficient range is $10–$30/h. However, the D-efficient design quickly becomes the most inefficient outside of this range (under $5/h and above $40/h), and the estimation significantly degrades above $50/h. The orthogonal and random designs are robust for a much larger range of VOT. The robustness of Bayesian efficient designs varies depending on the variance that the prior assumes. Implementing two-stage designs that first use a small sample to estimate priors are also not robust relative to uninformative designs. Arguably, the random design (which is the easiest to generate) performs as well as any design, and it (as well as any design) will perform even better if data cleaning is done to remove choice tasks where one alternative dominates the other.