The purpose behind conducting experiments is to determine the independent influence of different variables (attributes or factors depending on the literature cited) on some observed outcome. In stated choice (SC) studies which are often used in transportation studies for estimating and forecasting behavior of travelers, road authorities, etc., this translates into the desire to determine the influence of the design attributes upon the choices that are observed to be made by sampled respondents undertaking the experiment. However, an acknowledged limitation of SC studies is that unless the number of person specific observations captured in a survey is extremely large, it is necessary to pool the responses obtained from multiple respondents in order to produce statistically reliable parameter estimates. As such, SC studies typically consist of numerous respondents being asked to complete a number of choice tasks in which they are asked to select one or more alternatives from amongst a finite set of alternatives. In each task, the alternatives, whether labeled or unlabeled, are typically defined on a number of different attribute dimensions, each of which are further described by pre-specified levels drawn from some underlying experimental design. The number of choice tasks each respondent is asked to undertake will generally be up to the total number of choice situations drawn from the experimental design. Consequently, an archetypal SC experiment might require choice data be collected on 200 respondents, each of whom are observed to make eight choices, thus producing a total of 1600 choice observations.

The necessity to pool data has lead several authors to seek ways to reduce the number of choice observations necessary for reliable analysis of choice data (e.g., Huber and Zwerina 1996; Sándor and Wedel 2001; Kanninen 2002; Bliemer and Rose 2006). Primarily, these research efforts have attempted to produce more statistically efficient experimental designs which for a given level of accuracy, allow for either a reduction in the number of choice set profiles shown to individual respondents or alternatively, a reduction in the number of respondents required to complete the experiment. Such designs have been widely studied within the literature. For example, Bunch, et al. (1994) studied statistically efficient main effects designs whilst Anderson and Wiley (1992) and Laziri and Anderson (1994) introduce methods to generate statistically efficient cross-effect designs.

Recently, Bliemer and Rose (2009) demonstrated how it is possible to analytical calculate a theoretical minimum sample size for any given SC study. Using the expected asymptotic variance covariance matrix of a SC design, Bliemer and Rose showed that there exists a
relationship between the expected standard errors of a design and the sample size requirements for that design. They show that this relationship can be manipulated to provide an indication as to what sample size will be required for each parameter estimate related to the design to be found to be statistically significant. Bliemer and Rose then go onto derive a statistical measure, which they term S-error, which can be used to generate a design that will minimise the theoretical minimum sample size required for a SC study using that design.

Whilst the S-error provides an analytical indication of the minimum sample size required for a SC design, Bliemer and Rose found that depending on the assumptions made in generating the design, sample sizes of 10 or less may theoretically provide statically significant parameter estimates in some cases. Such sample sizes are unlikely to work in practice however given that estimation procedures often require substantially more observations in order to yield stable parameter estimates. Thus, whilst a design may theoretically yield low standard error values at a given small finite sample size, the question remains as to whether the parameter estimates themselves are stable in such small samples. If this is not the case, then reliance on S-error as an indicator of minimum sample size may result in biased study results.

In this paper, we explore the issue of sample size further, using Monte Carlo simulations to test the veracity of S-efficient designs to determine the sample sizes required for SC studies. Using simulated data for designs of varying S-efficiency, using bootstrapping we i) test the performance of the S-efficiency criteria and ii) examine at what sample size the parameter estimates obtained from different types of discrete choice models (i.e., MNL and MMNL models) become stable. In doing so, this research seeks to provide guidelines to transport researchers using SC experiments as to the minimum sample sizes that may be used, as well as provide researchers with the tools to calculate these requirements for specific studies.

References


