Journal of Statistical Theory and Practice

ISSN: 1559-8608 (Print) 1559-8616 (Online) Journal homepage: http://www.tandfonline.com/loi/ujsp20

Possible design-induced artifacts associated with designs for discrete choice experiments

Tiago Ribeiro, Richard Carson, Jordan J. Louviere & John M. Rose

JOURNAL OF STATISTICAL THEORY AND PRACTICE
http://dx.doi.org/10.1080/15598608.2016.1258680

Possible design-induced artifacts associated with designs for discrete choice experiments

Tiago Ribeiro[a], Richard Carson[b], Jordan J. Louviere[c] and John M. Rose[a]

[a]Institute for Choice, University of South Australia, North Sydney, Australia; [b]Department of Economics, University of California San Diego, La Jolla, California, USA; [c]School of Marketing, University of South Australia, Adelaide, Australia

ABSTRACT

Discrete choice experiments (DCEs) are widely used in many areas of applied social science research. The results of DCEs depend on the particular experimental design for the identification of the key parameters of interest and the statistical efficiency with which those parameters are estimated. Work on experimental designs for DCEs has almost always assumed that the particular design one uses does not influence the nature of the responses to the choice tasks other than via the precision with which parameters are estimated. We examine this assumption by testing whether particular experimental designs influence the probability that a separating hyperplane exists that perfectly predicts the observed choices at the individual level in four DCE data sets. Our empirical results suggest that the particular statistical design used can influence the nature of the choice responses obtained.

ARTICLE HISTORY
Received 8 June 2015
Accepted 5 November 2016

KEYWORDS
Behavioral response; choice model; experimental design

AMS SUBJECT CLASSIFICATION
Primary 62K05; secondary 62K99; 62P25; 91B42

1. Introduction

Discrete choice experiments (DCEs) are widely used in many applied fields in the social sciences, including economics, health policy, marketing, political science, psychology, and transport (Louviere, Hensher, and Swait 2000). They are most often used in a survey context where participants provide stated preference information in the form of a discrete choice between available options that differ on the levels of one or more attributes that are assigned according to an experimental design. Revealed preference versions of DCEs are often used in field experiments and test markets. DCEs have proven especially popular in computer-assisted interviewing, in Internet surveys, and in test markets run online, as they allow quite complex experimental designs to be easily used.

Experimental designs used in a DCE are important because they determine what parameters are statistically identified and the efficiency with which the statistics of interest can be estimated. Two jumping-off points for thinking about experimental design for DCEs are (1) dose-response experiments in the biometrics literature (Finney 1978), because the simplest DCEs vary only one treatment variable, and (2) the literature on industrial experiments (Box, Hunter, and Hunter 2005), as most DCEs involve varying multiple factors. Because many readers will be unfamiliar with experimental design

CONTACT Jordan J. Louviere Jordan.Louviere@unisa.edu.au 10700 NE 4th St., Unit 2102, Bellevue, WA, 98004, USA.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ujsp.

applied to DCEs, we provide an extensive literature review. We pay particular attention to how the features of a DCE can influence potential responses to it.

Statistics of interest from a DCE often involve ratios of estimated model parameters because these represent marginal trade-offs, such as marginal willingness to pay (WTP) for an increase in a factor, which can be estimated if one of the model parameters represents the cost of an option. Efficiency is important for the usual reasons, that is, obtaining a tight confidence interval for a given sample size or minimizing the number of observations needed to obtain a given confidence interval. However, this article focuses on a different issue, namely, the possibility that the particular experimental design used influences the estimated statistics of interest in ways unrelated to the usual statistical concerns of identification and efficiency. Indeed, such design-related behavioral effects have been observed for different aspects of DCEs, such as the number of attributes or the number of choice options. We focus on an indicator of individual behavior: Is it possible to identify a separating hyperplane that perfectly predicts observed responses? Such a hyperplane is usually regarded as undesirable at the aggregate level as it suggests the response from the population of interest is perfectly predictable and therefore lacks the random component that underlies the random utility models (Hensher, Rose, and Greene 2005; McFadden 1974) typically fitted to choice data from DCEs.

The separation measure we propose has a more complex interpretation at the individual level. The standard economic formulation of a random utility model assumes that behavior at the individual level is deterministic. Our indicator provides an upper bound estimate on the fraction of the sample for whom this assumption cannot be ruled out, assuming away decisions under uncertainty. Our indicator does not depend on a particular distributional assumption for the random component, but it does depend on the particular experimental design used, and this dependence can be used to explore the ways in which different experimental designs interact with behavior.

2. DCE experimental design literature review

Choice scenarios (known as "choice sets") presented in DCEs represent a finite set of options defined on a number of attribute dimensions, the levels of which are systematically varied by the associated design. Experimental participants are asked to indicate their preferred option in each choice set. Conceptually, an experimental design is simply a matrix of values used to determine what goes where in a DCE. These values can be numbers or labels, depending on how the analyst chooses to communicate the information in the experiment to participants. Figure 1, from Vincent et al. (2014), displays one choice task an individual was asked to complete about policy options for protecting a large tract of tropical rainforest in Malaysia. A review of the literature (Louviere, Pihlens, and Carson 2011) suggests little consensus about specific experimental design approaches appropriate for DCE studies; hence, multiple approaches appear in the literature.

The primary focus of research into experimental design theory related to DCEs has tended to be on producing designs deemed more statistically efficient, where statistical efficiency is related to the expected standard errors that a design will produce. All else being equal, designs expected to produce smaller standard errors are said to be more statistically efficient (Rose and Bliemer 2009). Thus, a direct link exists between the statistical efficiency of a design and the sample size requirements of DCE studies; that

	Policy A	Policy B	No protection
Logging	150,000 ha	0 ha	300,000 ha
	Half these species go extinct	None of these species go extinct	All these species go extinct
Poaching	0 ha	150,000 ha	300,000 ha
	None of these species go extinct	Half these species go extinct	All these species go extinct
Floods in Perak	3 per year	1 per year	5 per year
Jobs created in Perak	5,000	7,500	7,500
Cost to you	RM6 per month	RM6 per month	No cost

Figure 1. A single choice task from Vincent et al. (2014).

is, more efficient designs should produce the same t-ratios as less efficient designs, but with smaller samples. Alternatively, they should produce larger t-ratios than less efficient designs given the same sample size.

DCE data typically are analyzed using nonlinear models such as the conditional multinomial logit and mixed multinomial logit models (Hensher, Rose, and Greene 2005; McFadden 1974). In turn, this implies that the efficiency of a design depends on the unknown parameter vector (Atkinson and Haines 1996). Given that the true parameter vector is unknown at the stage at which the design is generated, analysts must make assumptions about specific values for the parameters. By assuming specific parameter values associated with a given design matrix X, it is possible to calculate the expected utilities for each of the choice options. Once known, these expected utilities can be used to calculate the expected choice probabilities. Next, given knowledge of the attribute levels (the design), expected parameter values, and the resultant choice probabilities, it is a straightforward exercise to calculate the Fisher information matrix, $I_{N'}$ which is computed as the Hessian matrix of negative expected second derivatives of the log-likelihood function of the model to be estimated, where N is the number of observed choices (Train 2009). The asymptotic variance–covariance (AVC) matrix, Ω_N, which is the inverse of the Fisher information matrix, then can be determined and the expected standard errors derived. By manipulating the attribute levels of the options for the known (assumed) parameter values, analysts can minimize some meaningful function of the elements in the AVC matrix (such as the trace or determinant), which generally implies lower standard errors, and hence greater reliability in the estimates for a fixed sample size.

Three different approaches to the problem of having to assume some prior information about parameter values have been developed. The first is to assume a priori precise knowledge of the parameter estimates, leading to what are termed *locally optimal designs* (so called because they are optimized for these specific prior parameter values). They quickly can lose efficiency if the true parameter values differ from those assumed in design generation, such as a common assumption of zero prior parameter values. In this case, linear experimental design theory can be used to generate designs that will be orthogonal in the attributes (e.g., Anderson and Wiley 1992; Kuhfeld, Tobias, and Garratt 1994; Lazari and Anderson 1994; Street, Bunch, and Moore 2001; Street and Burgess 2007). Alternatively, some researchers generated locally optimal designs under non-zero prior parameter values (e.g., Carlsson and Martinsson 2003; Huber and Zwerina 1996); in this case, nonorthogonal designs tend to be more statistically efficient than orthogonal designs.

A second, more recent, approach tries to integrate uncertainty about assumed parameter values with Bayesian design methods (e.g., Chaloner and Verdinelli 1995). First applied to DCEs by Sándor and Wedel (2001), the Bayesian approach involves assuming prior parameter distributions instead of specific fixed values, and examining the AVC matrix generated over draws from these distributions. This design approach has been shown to produce Bayesian optimal designs that are less efficient than correctly specified locally optimal designs but more robust to prior parameter misspecification (e.g., Sándor and Wedel 2001). As with locally optimal designs assuming nonzero prior parameter values, nonorthogonal designs tend to be more statistically efficient for this case (e.g., Kessels et al. 2009). Ongoing research for this class of designs studies how to best represent Bayesian prior parameter distributions (e.g., Bliemer, Rose, and Hess 2008; Goegebeur, Goos, and Vandebroek 2007; Yu, Goos, and Vandebroek 2008, 2010).

A third approach assumes priors can be sequentially updated by estimating the parameters on progressively larger samples during the data collection process. For each participant (or subsample of participants), a new design can be generated based on the currently set local or Bayesian priors. This process was proposed by Kanninen (2002), while Bliemer and Rose (2010) and Yu, Goos, and Vandebroek (2011) showed that such sequentially optimal designs can improve the efficiency of the design significantly but require more complex data collection methods and analysis.

Aside from prior parameter estimates, different assumptions are required to generate an experimental design. Each experimental design is optimized based on the AVC matrix of a specific model, and different models produce different AVC matrices, even when the design is fixed. For example, assuming different functional forms (e.g., main effects versus main effects and first-order interactions), different assumptions about the distribution of parameters (e.g., fixed versus random parameters), or the distribution of the error component all produce different optimal designs (e.g., Tudela and Rebolledo 2006). There also are different statistical criteria (e.g., C-optimality, D-optimality, G-optimality) that effectively trade off a tight confidence interval around one or a small number of statistics of interest versus an ability to look at a wide range of effects with reasonable precision. This implies that analysts must specify the particular model they intended to estimate and the statistic(s) of interest before generating a design. The AVC matrix also is influenced by the type of coding used when generating a design. For example, dummy or effects codes produce different AVC matrices than linear or other coding structures. Finally, analysts sometimes impose constraints on attributes, attribute levels, and combinations of levels that may influence a design, such as whether a design is balanced in the levels, orthogonal, or restrictions in the appearance of certain attribute-level combinations such that choice options with more desirable attribute levels do not appear at prices lower than competing options with less desirable levels. All of these considerations can impact the final design generated.

While the DCE literature often tends to describe different design paradigms, we view this as one overarching design methodology with analysts making different assumptions that take them in different directions as to what they consider optimal designs. For example, designs generated using the Street and Burgess (2007) approach are often viewed as a different design paradigm. However, it is straightforward to show that Street and Burgess designs are locally optimal designs generated under the assumptions of zero priors, a conditional multinomial logit model, and orthonormal orthogonal attribute-level coding. Thus, we view the Street and Burgess approach as a way to produce designs optimized for a very specific set of assumptions, not a separate class of designs.

The purpose of this article is to examine whether designs generated under different assumptions can influence behavioral outcomes from DCEs. It is well known and long established in psychology that features of experimental tasks can influence outcomes (a phenomenon known as demand characteristics or demand-induced effects; e.g., Orne 1962, 1969). However, this has only received attention in the DCE literature recently, with most attention focused on DCE design characteristics, such as the number of options, the number of attributes, or the number of choice sets administered to each person (e.g., Beck, Kjaer, and Lauridsen 2011; Caussade et al. 2005; Hensher, Stopher, and Louviere 2001). The latter work (examined in more detail in the following) suggests the possibility of behavioral effects in response to choice tasks; thus, we focus on another major aspect of DCE design generation, namely, whether designs generated under different assumptions can influence the outcomes of DCEs beyond simply influencing parameter estimate precision. We study this by looking at a measure of perfect predictability of choices at the individual level. This measure couples the notion of how tightly a particular experimental design identifies the space in which the choice model's parameter vector lies and the influence of the particular experimental design on the estimated error component.

3. Experiments looking at influence of DCE features on choices

There is a small literature on the influence of various features of DCEs on choice behavior relevant to our work on the role of particular DCE design features on choice behavior. We begin by noting that if people make choices consistent with the basic rational economic model typically assumed in DCE applications and analysts know the correct model of choice behavior up to a small vector of parameters that can be estimated from the data, neither the features of a DCE nor the particular experimental design would influence choice behavior. Furthermore, work in economic applications of DCEs often adopts the strong assumption (Manski 1977) that error components reflect fixed factors not observed by analysts and lacks a true random component. It is unlikely that these assumptions are true in practice. For example, people face time constraints that make effort costly, suggesting satisficing behavior (Krosnick 1991), where reasonably good options are chosen but the best choice is occasionally overlooked, and a specific parametric functional form for modeling choices is unknown. At the very least, the features of a particular choice task seem to robustly influence estimates of the error components of choice models once observations on more than one choice are available.

Features of DCEs that have been studied regarding their influence on choice behavior are (a) the number of choice sets observed from the same individual, (b) the number of choice options in a choice set, (c) including a status quo (i.e., an option with fixed attribute levels that describes the only current option) and/or no action/purchase option, (d) the number of attributes used to describe each option, (e) the order in which attributes appear, and (f) the number of levels of each attribute. Empirical tests suggest no or small effects for some of these design features, but there is fairly consistent evidence that other features can have a substantial influence on how choice questions are answered, as we now discuss.

The number of choice sets each individual answers is important from a number of perspectives, including survey cost and analysts' ability to consider different sources of variability between parameters at the individual level. Work on numbers of choice sets has found no effect or small (precisely estimated) effects; also, effects found so far tend to impact error variances instead of the preference parameters in typical choice models (e.g., Hensher 2004). Sometimes these error variance effects suggest some type of learning and fatigue process (Brazel and Louviere 1998), but the evidence is mixed: Caussade et al. (2005) systematically varied several design features and found that numbers of choice tasks (choice sets) had the least influence of any design dimension, while Rose et al. (2009) found no significant influence on the estimate of willingness to pay for a good in Australia, limited impact in Taiwan, but a reasonably large impact in Chile.

In contrast to numbers of choice sets, numbers of options per choice set presented to participants often influence choices, particularly if one option is the status quo (SQ). Some of the observed differences previously reported are predictable from underlying economic theory on incentive and information structures of different types of choice sets (Carson and Groves 2007, 2011) that can be present in a sequence of choice sets with only two options if people do not treat the choice sets as being independent (Carson and Groves 2007; Day et al. 2012). Adamowicz, Dupont, and Krupnick (2006) and Rolfe and Bennett (2009) found participants more likely to choose a SQ option than would be predicted from preference parameters estimated from a two-option version. DeShazo and Fermo (2002) found a quadratic relationship between numbers of options and the error variance, suggesting error variance first decreases, then increases with numbers of options, but this finding is not universal. For example, Arentze et al. (2003) found no significant error variance differences between DCEs for two versus three options; Caussade et al. (2005) found numbers of options had the second largest influence on error variances of all design criteria tested. Hensher (2006) and Rose et al. (2009) found some influence of the number of options on estimates of WTP for changes in particular attributes.

SQ options are different from simply being an additional option because they ensure individuals can maintain their current level of utility (Bateman et al. 2003). A complication for the SQ option with private and quasi-public goods is that it can differ between individuals (Rose and Hess 2009). For private goods, instead of an SQ option, the fixed option often is the possibility of not choosing or waiting to choose until later (e.g., Dhar 1997). For public goods, an SQ option may have special status from an informational standpoint if public projects have a chance of failure or cost overruns. Moreover, as choice options become more complex, SQ options tend to increase in relative attractiveness (Boxall, Adamowicz, and Moon 2009). The special role of SQ options often is hypothesized to be reflected in violation of the independence of irrelevant alternatives assumption

(IIA) underlying simple conditional logit models. However, the empirical evidence is mixed: Dhar and Simonson (2003) found IIA violations, while Brazell et al. (2006) and Rose and Hess (2009) did not.

In contrast to the previously noted mixed evidence, increasing numbers of attributes almost always seem to influence choices. The effect is particularly large as the number of attributes becomes fairly large, as some people seem to use only a subset of attributes to make choices. In an influential paper, Green and Srinivasan (1990) argued that participants cannot process many attributes simultaneously; they become tired and consequently ignore or process attributes in random and uncontrolled ways or tend to use heuristics leading to biased preference measures. Caussade et al. (2005) found the number of attributes had the largest influence on error variances, of all the design criteria they examined. Others (e.g., Arentze et al. 2003) found similar results for error variances but also reported differences in parameter estimates, while Hensher (2006) and Rose et al. (2009) found statistically significant differences in WTP measures as numbers of attributes increased.

There is some evidence that the order in which attributes are presented matters. For example, Kjaer et al. (2006) presented the price attribute as first or last and found order produced statistically significant differences in the estimate of sensitivity to price. While this is not surprising because firms often intentionally manipulate the order in which price appears, such effects do not always occur. For example, Farrar and Ryan (1999) and Boyle and Özdemir (2009) suggested attribute order effects likely depend on the number of attributes and specific application contexts.

There also is mixed evidence that numbers of attribute levels and ranges of attribute levels influence how people respond to choice tasks. Wittink, Krishnamurthi, and Reibstein (1989) and Wittink, Krishnamurthi, and Reibstein (1992) found adding one or more intermediate levels to a two-level attribute increased its impact, while Hensher (2006) found the number of attribute levels influenced the probability of participants ignoring some but not all DCE attributes. Caussade et al. (2005) found the number of attribute levels had a statistically significant impact on the estimated error variance but this effect was marginal and less important than effects associated with most other design criteria studied; Rose et al. (2009) found impacts on WTP estimates depended on country.

For continuous variables, ranges of attribute levels also may matter. For example, Meyer and Eagle (1982) found attributes with larger ranges had larger effects than those with smaller relative ranges (all else being equal). A difficulty with such effects is that increasing the range of an attribute also influences the nature of models that can be fitted as one moves away from local linearity to being able to capture curvature. For example, Ohler, Louviere, and Swait (2000) showed that changing the range of one attribute did not influence estimates of parameter values when a more complex functional form was fitted to their data. Caussade et al. (2005) concluded that attribute range had a relatively large impact on error variances; Hensher (2004) found that increasing the range of levels influenced WTP estimates.

Most relevant to this article are studies explicitly exploring choice task complexity because different approaches to deriving statistical experimental designs for DCEs can influence relationship between different choice options presented to individuals. Of course, there are many definitions of complexity, none of which is ideal. An early study defined complexity as the degree to which attribute levels differ across two options, and showed that variation in the correlational structure of attributes in choice tasks influenced parameter estimates (Mazzota and Opaluch 1995). That result is consistent with the result of Mellers and Biagini (1994), who found that when attribute levels overlap across options, sensitivities to changes in other attributes increase. Dellaert, Brazell, and Louviere (1999) showed that greater choice difficulty defined by price differences across options led to larger estimated random error variances. Swait and Adamowicz (2001) used entropy to measure task complexity and found that the degree of task complexity systematically impacted estimated choice model parameters. Finally, DeShazo and Fermo (2002) found that WTP estimates varied significantly as they varied correlations between DCE options.

Complexity also can be implicitly defined in terms of the statistical efficiency of a particular DCE experimental design. For example, Louviere et al. (2008) studied 44 different experimental designs that varied different design criteria and levels of statistical design efficiency. They found that the more statistically efficient a design, the greater is the error variance in the data, a relationship independent of differences in design criteria. Viney, Savage, and Louviere (2005) found that an orthogonal main effects design, a utility-balanced design, and a random design did not impact the underlying parameter estimates. However, the utility-balanced design (where participants should not prefer either option) exhibited more random response variability. Bliemer and Rose (2011) compared results from orthogonal and efficient designs under the assumption of nonzero parameter values and found that efficient designs produce lower standard errors. They conjectured without proof that efficient designs may produce higher error variance than orthogonal designs because orthogonal designs have more dominated options that are easier to answer, resulting in lower error variance. Louviere (2011) reported that substantial numbers of experimental participants consistently and systematically chose based on one level of one attribute (e.g., always choosing an option with the lowest price). More recently, Louviere (2013) noted that this phenomenon seemed associated with optimal DCE designs in general, not just Street and Burgess (2007) designs. Louviere (2013) also showed some individuals never chose options with a particular attribute level, such as not choosing options with the highest price level regardless of its other attributes. He found this behavior to be most concentrated in DCEs using Street and Burgess (2007) designs but it also seemed to be associated with other designs deemed efficient under particular assumptions.

Such results suggest interesting trade-offs if a design matrix affects error variances and with how people respond to choice tasks. For example, this may indicate types of satisficing behavior that make choices look more deterministic and are independent of particular random utility models used, as it does not rely on a distributional assumption for the error component. In turn, this suggests that it may be useful to have an indicator that is sensitive to such a state of affairs.

4. Research approach

We study rates of perfect separation among individuals induced by a given design. “Perfect separation” means there is a perfect predictor of a given individual’s choices; that is, there exist no detectable violations to a possible complete preference ordering of the products. Perfect separation occurs if the error a participant made (if any) is smaller than the maximum error detectable by a given design.

We use a simple example to begin our research approach discussion that illustrates how design decisions can impact expected perfect separation rates and how they relate to model error terms. Let there be four products, $i = 1, \ldots, 4$, defined by attributes X_i^a and X_i^b; both assume values in the set {-0.5, 0.5} and utility is defined as $U_i = X_i^a\beta^a + X_i^b\beta^b$. Products described by attribute levels $(X^a, X^b)$are $\{(-0.5,\ -0.5),\ (-0.5,\ 0.5),\ (0.5,\ -0.5),\ (0.5,\ 0.5)\}$. In what follows we label the products 1–4. Participants in a sequence of M choice sets each offering K products indicate their preferred product. If K = 2 and M = 4 with choice sets (1,2), (1,3), (1,4), and (2,3) there are $2^4 = 16$ possible choices. If a participant chooses product 1 in choice sets 1 to 3 and product 2 in set 4, this implies that $U_1 > U_2$, $U_1 > U_3$, $U_1 > U_4$, and $U_2 > U_3$ (we exclude the possibility of ties).

Any pair of coefficients (β^a, β^b) that satisfies the set of inequalities

$$(X_2^a - X_1^a)\beta^a + \left(X_2^b - X_1^b\right)\beta^b \langle 0 \leq \rangle \beta^b < 0$$

$$(X_3^a - X_1^a)\beta^a + \left(X_3^b - X_1^b\right)\beta^b \langle 0 \leq \rangle \beta^a < 0$$

$$(X_4^a - X_1^a)\beta^a + \left(X_4^b - X_1^b\right)\beta^b \langle 0 \leq \rangle \beta^a + \beta^b < 0$$

$$(X_3^a - X_2^a)\beta^a + \left(X_3^b - X_2^b\right)\beta^b \langle 0 \leq \rangle \beta^a - \beta^b < 0$$

could represent these preferences using the above utility function. The pair of coefficients defines a separating hyperplane passing through the origin such that all points defined by the difference in attributes lie to one side of the hyperplane. Columns three and four in Table 1 calculate the difference in attributes for this case, and an example of a separating hyperplane is depicted in the left panel of Figure 2: Any line through the origin with all the points above it is a separating hyperplane.

Had a participant chosen products 1, 3, 1, and 2 as the preferred products in the four choice sets in Table 1, it would be the case that $U_1 > U_2$, $U_3 > U_1$, $U_1 > U_4$, and $U_2 > U_3$, producing a preference reversal because the second and fourth choices imply that $U_2 > U_3 > U_1$, but the first choice gives $U_1 > U_2$. In this case no separating hyperplane exists. The attribute differences for this case are calculated in columns six and seven of Table 1, and are depicted in the right panel of Figure 2.

Table 1. Differences in attributes.

	Separating hyperplane exists			Separating hyperplane does not exist		
Choice set	Preferred	ΔX^a	ΔX^b	Preferred	ΔX^a	ΔX^b
(1,2)	1	0	1	1	0	1
(1,3)	1	1	0	3	−1	0
(1,4)	1	1	1	1	1	1
(2,3)	2	1	−1	2	1	−1

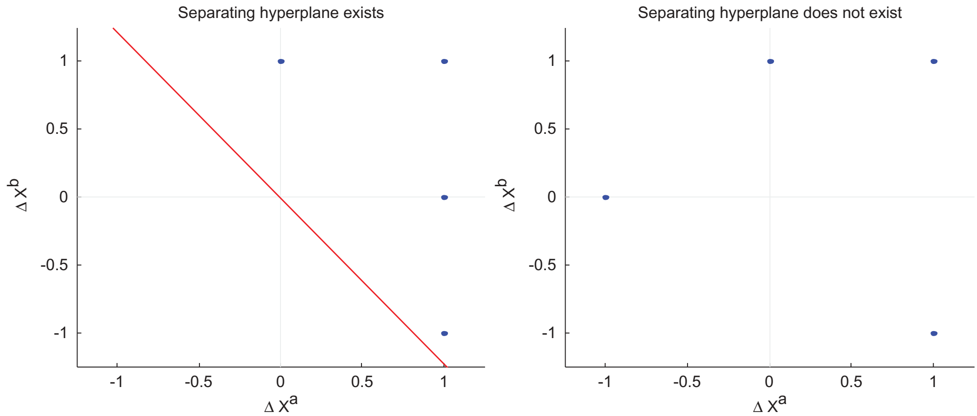

Figure 2. Examples of existence of separating hyperplane.

In the preceding example with 4 binary sets there are $2^4 = 16$ possible choice sequences. Of these 16 possible choice sequences, there is a separating hyperplane for only 8 (50%); that is, by inspection it can be seen that only 8 of the possible choice sequences are consistent with a preference ordering. The remaining 8 sequences have preference reversals, as illustrated in the preceding. Also note that this example with 4 products has 4! = 24 possible preference orderings. Only 8 choice sets are compatible with a preference order, implying several preference orderings are observationally equivalent (e.g., the ordering $U_4 > U_3 > U_2 > U_1$ yields the same pattern of answers as the ordering $U_3 > U_4 > U_2 > U_1$).

If, instead of the four choice sets just described, we consider all pairwise comparisons ($M = 6$, $K = 2$), the percentage of cases where a separating hyperplane exists is 12.5% (out of 64: 6 binary choices have $2^6 = 64$ possible outcomes). In the case of choice sets formed by all triplets ($M = 4$, $K = 3$) the percentage of cases where a separating hyperplane exists is 9.9%. Alternatively, if there are 9 products (e.g., let the attributes $X_i^a$and X_i^b assume values in the set {−1, 0, 1}), and we expand the utility function to allow both linear and quadratic terms in both attributes and maintain the same choice sets illustrated in the preceding, percentages of possible answers consistent with separation are 75.0%, 37.5%, and 14.8%, respectively.

The preceding discussion illustrates variations in proportions of perfect separation that should be expected from different experimental designs if randomly generated choices are analyzed with each option in every choice set having an equal probability of participant choice. This benchmark captures statistical properties of designs without regard to preference and choice processes. Generally, increasing the number of choice sets per participant (all else equal) decreases separation, increasing the number of options per choice set decreases separation, and increasing the number of options increases separation. Now consider the case of participants who behave as if their underlying decision process is given by the conditional logit model (McFadden 1974). In the conditional logit model the utility of a given option is given by the expression: $U_i = X_i^a\beta^a + X_i^b\beta^b + \sigma\varepsilon_i$, where ε_i is an extreme value type I random term. Clearly, if σissmall, choices generated by this model will approach 100% separation and there will be no preference reversals; as σ increases, choices will appear increasingly random. To illustrate the relationship between separation and σ we simulate choices generated by this model: We let X_i^a and X_i^b assume values in the set {–0.5, 0.5} to define four products, coefficients $(\beta^a, \beta^b) = (1, -1)$, the four choice sets defined earlier ((1,2), (1,3), (1,4), and (2,3)), all pairwise combinations and all triplets, and a fixed value of σ. We simulate 1000 sets of choices and calculate the perfect separation percentages in the 1000 simulations[1] for each combination of choice sets and σ. Results for different values of σ are in Table 2.

Perfect separation yields nonidentification of preference parameters at the individual level. At a population level, although perfect separation is rare, the ability to estimate models with heterogeneous individuals relies solely on restrictions imposed across individuals (usually under some form of regularization) for identification.

5. Methodology

For each participant, we determine whether there exists a separating hyperplane that perfectly classifies the observed responses (i.e., a linear utility index that perfectly classifies the data without error). This is done by solving a linear program (LP) for finding separating hyperplanes that we describe in some detail in the following. Thus, each participant is classified with respect to the existence (or not) of a separating hyperplane that perfectly classifies that participant's responses. We then relate this binary classification of each participant to the specific design to which that participant was assigned. We then estimate a logistic regression that describes the effects of design characteristics on the likelihood of obtaining perfect separation.

We start by introducing notation and then describe a method that can be used to determine whether there is perfect separation. Throughout we use the following notation: n indexes participants, c indexes choice sets, i indexes options, X_{nci} is a row vector of attributes of option i in choice set c faced by participant n, β_n is a column vector of coefficients, and $U_{ci}^n = X_{ci}^n \beta_n$ is the participant n's utility of option i in choice set c. The choice options are relabeled according to the rank they have been given by participants; consequently, the most preferred option by participant n in choice set c has utility U_{c1}^n, the second most preferred has utility U_{c2}^n, and so forth. All the analysis is done at the participant level; hence, whenever possible to avoid confusion we drop the superscript n.

If a participant states her most preferred option from several choice sets of four options, the following inequalities should be satisfied for all choice sets c: $U_{c1} > U_{c2}$, $U_{c1} > U_{c3}$, and $U_{c1} > U_{c4}$. Thus, we search for the existence of a coefficient vector β such

Table 2. Separation and choice sets and variance.

σ	0.1	0.2	0.5	1.0	2.0	5.0	10.0
Four pairs	100%	99%	87%	67%	57%	53%	51%
All pairs	100%	97%	55%	27%	16%	13%	13%
All triplets	100%	97%	58%	24%	13%	11%	10%

[1] Even though the systematic component of the utility is the same for products 1 and 4 in this numeric example, the error component will always ensure that there are no ties in the preferences between these two products.

that these inequalities are satisfied for all c. We do this by solving the following LP problem (see, e.g., Mangasarian 1965):

$$\min_{\beta,\varepsilon_{cr}} \sum_{c} \sum_{r=1}^{3} \varepsilon_{cr}$$

$$s.t.$$

$$(X_{c2} - X_{c1})\beta \leq -1 + \varepsilon_{c1}, \forall c$$

$$(X_{c3} - X_{c1})\beta \leq -1 + \varepsilon_{c2}, \forall c$$

$$(X_{c4} - X_{c1})\beta \leq -1 + \varepsilon_{c3}, \forall c$$

$$\varepsilon_{cr} \geq 0, \forall c, r$$

If a solution $\left(\varepsilon^*_{cr}, \beta^*\right)$ to the preceding problem has all $\varepsilon^*_{cr} = 0$, that is, the objective function at the optimum is 0, then β^* defines a hyperplane that perfectly classifies all the participant's choices.[2] Whenever a solution is found that is 0 at the optimum, we classify a participant as perfectly separable.

In the analysis that follows all attributes are qualitative and levels are indicator variables, ruling out functional form issues (except attribute interactions). Because we restrict ourselves to perfect separation, we also rule out issues relating to distributions of error terms.

6. DCE data sets and experimental designs

We begin by describing four DCE data sets used to look at the impact of survey design on participant behavior in choice tasks produced by discrete choice experiments. In each case a design of designs approach[3] was used. The data sets used can be summarized as follows:

[2]Informally, correspondence between a zero at the optimum and existence of a separating hyperplane can be seen as follows: (i) A solution to the problem always exists since for any β, ε_{cr} can be set as large as desired thus satisfying all constraints; (ii) the optimum is always ≥ 0 given the constraints on the ε_{cr}; (iii) if the optimum is 0, then due to constraints on ε_{cr} it must be that all ε_{cr} are 0 at the optimum; (iv) in this case inequalities of the type $\left(X_{cj} - X_{c1}\right)\beta \leq -1$ hold; (v) multiplying both sides of the inequalities by any positive number leaves the inequalities and LP program optimum unchanged, so −1 is just a convenient normalization; (vi) the inequalities hold with 1 replaced by any arbitrary positive number, so one can choose any arbitrary small positive number to conclude a β exists such that inequalities $\left(X_{cj} - X_{c1}\right)\beta<0$ hold; and (vii) these inequalities are the definition of a separating hyperplane. The converse is also true: If a solution is not 0 at the optimum, no separating hyperplane exists. Mangasarian (1965) has a detailed derivation showing the stated LP can identify separating hyperplanes.

[3]The expression "design of designs" refers to different design characteristics being varied by a design (originally proposed by D. A. Hensher 2004).

(1) Carbon tax DCE: a DCE designed to elicit preferences over different possible mechanisms to tax carbon emissions. The different attributes used to define a given taxation policy option are detailed in the Appendix in Table A1. There are 32 different designs that vary in the method that generated the experimental design, the number of choice tasks per participant, the number of attributes with two levels, and the number of attributes with four levels. All choice tasks have four options. The structure of each of the 32 designs is detailed in Table A5.
(2) Solar panels DCE: a DCE designed to elicit willingness to pay for solar panels. The different attributes used to define a given tax policy option are detailed in the Appendix in Table A2. There also are 32 different designs with the same structure as the carbon tax data set, as detailed in Table A6.
(3) Flights DCE: a DCE designed to elicit preferences for different air travel options. The attributes used to define different flight options are detailed in the Appendix in Table A3. There are 33 different designs that vary in the method for generating the experimental design, the number of choice tasks per participant, the number of options per choice task, the number of attributes with two levels, and the number of attributes with four levels. The structure of each of the 33 designs is detailed in Table A7.
(4) Pizza DCE: a DCE designed to elicit preferences for different pizzas. The attributes used to define different pizza options are detailed in the Appendix in Table A4. There are also 33 designs used with the same structure as the flights data set, as detailed in Table A8.

Hereafter, we call the four data sets the carbon, solar, flights, and pizza data sets, respectively.

We sampled all participants from the Pureprofile Web panel that has recruited approximately 600,000 unique households and reasonably represents the general Australian population. Participants were randomly assigned to one experimental condition; sample sizes vary by product or service type but each has approximately 50 people. We classified all participants in each data set by whether a perfect separation existed or not using methods outlined in the preceding. This classification is the dependent variable in our analyses; explanatory variables are design characteristics in Table 3. We computed

Table 3. Variable descriptions.

		Data sets			
Variable	Description	Carbon	Solar	Flights	Pizza
d_sb	S&B design	✓	✓	✓	✓
d_sas	SAS design	✓	✓	✓	✓
d_bibd	BIBD design	✓	✓	✗	✗
d_rd	Random design	✓	✓	✗	✗
d_saw	Sawtooth design	✗	✗	✓	✓
d_16	16 choice tasks per participant	✓	✓	✓	✓
d_24	24 choice tasks per participant	✓	✓	✗	✗
d_32	32 choice tasks per participant	✗	✗	✓	✓
Nprod	Number of unique products	✓	✓	✓	✓
Natt	Number of attributes	✓	✓	✓	✓
Nalt	Number of options per choice set	✗	✗	✓	✓
D max	D_h^{max}	✓	✓	✓	✓
D dif	D_h^{dif}	✓	✓	✓	✓

three measures of disparity of options allocated to choice sets for each design. Let X_{hci} be a vector of attributes of option i in choice set c of design h and $d^H(x, y)$ be the Hamming distance between vectors x and$y$ (i.e., number of elements where vectors x and y differ). The three measures are the average maximum distance between options in a choice set D_h^{max}, the average minimum distance between options in a choice set D_h^{min}, and the difference between the two $D_h^{dif} = D_h^{max} - D_h^{min}$. Specifically, we have:

$$D_h^{max} = \frac{1}{C}\sum_{c=1}^{C} \max_{i,j,i \neq j} d^H\left(X_{hcj}, X_{hci}\right)$$

$$D_h^{min} = \frac{1}{C}\sum_{c=1}^{C} \min_{i,j,i \neq j} d^H\left(X_{hcj}, X_{hci}\right)$$

We also computed a baseline separation measure for each data set and each design: We drew 1000 choice sequences uniformly at random for each experiment and calculated the percentage of separated cases to measure separation induced by a given design. We "effects coded" indicator variables and normalized continuous variables by dividing by twice their standard deviation (a 1-unit increase = an increase of 2 standard deviations in original units).

7. Results

Results of a regression analysis relating separation to design characteristics are in Table 4; Figures 3 and 4 are corresponding graphical results for easy comparison. Results are consistent across data sets: More (a) choice sets per person, (b) options per choice set, and (c) attributes per option all decrease separation; the number of unique design options increases separation. Note the large disparity in unique option numbers across designs: BIBD designs have the least (16); some SAS and S&B designs have over 100. Also, in all cases higher maximum disparity between options in a given choice set (D_h^{max}) gives more separation, and more similarity between options (D_h^{max}) gives less separation. After controlling for design features, S&B designs produced less separation in both carbon and solar data sets, but more in flights and pizza.

Including a baseline measure and controlling for separation expected from a random response pattern maintains all results, except dummy variables for each design are nonsignificant. These dummy variables capture separation induced merely by statistical properties of designs, suggesting separation results are associated with induced participant behavior due to designs. Carbon and solar data sets yield the same results based on most preferred option choices after accounting for differences in constant terms. The same results obtain for flights and pizza. All results are in Table 5.

Table 4. Results for separation based on most preferred alternative.

	Results for separation based on most preferred alternative							
	Carbon	Solar	Flights	Pizza	Carbon	Solar	Flights	Pizza
	Without baseline				With baseline			
Constant	−2.494	−2.154	5.181	3.793	−2.158	−1.685	4.314	3.528
t Statistic	−5.07	−4.34	5.61	4.40	−4.26	−3.14	3.42	3.44
p Value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0017	0.0006	0.0006
S&B design	−0.556	−0.703	0.461	0.349	−0.420	−0.531	0.363	0.336
	−2.19	−2.69	2.65	2.05	−1.62	−1.95	1.79	1.94
	0.0284	0.0072	0.0080	0.0407	0.1049	0.0507	0.0729	0.0523
SAS design	0.063	−0.298	−0.622	−0.788	0.145	−0.238	−0.477	−0.742
	0.33	−1.52	−3.01	−3.93	0.76	−1.21	−1.88	−3.32
	0.7405	0.1281	0.0027	0.0001	0.4460	0.2253	0.0598	0.0009
BIBD design	−0.023	0.562			0.004	0.555		
	−0.05	1.30			0.01	1.30		
	0.9563	0.1923			0.9923	0.1941		
RND design	0.517	0.439			0.271	0.214		
	1.76	1.45			0.87	0.67		
	0.0777	0.1482			0.3845	0.5026		
SAW design			0.161	0.439			0.114	0.406
			0.55	1.52			0.38	1.36
			0.5831	0.1281			0.7033	0.1753
16 choice sets	0.676	0.745	1.035	1.439	0.579	0.534	0.935	1.398
	5.09	5.42	3.30	4.60	4.20	3.20	2.85	4.30
	0.0000	0.0000	0.0010	0.0000	0.0000	0.0014	0.0043	0.0000
24 choice sets	−0.676	−0.745			−0.579	−0.534		
	−5.09	−5.42			−4.20	−3.20		
	0.0000	0.0000			0.0000	0.0014		
32 choice sets			−1.035	−1.439			−0.935	−1.398
			−3.30	−4.60			−2.85	−4.30
			0.0010	0.0000			0.0043	0.0000
Nprod	0.742	1.534	1.592	2.229	0.857	1.372	1.537	2.190
	1.06	2.12	2.88	3.89	1.22	1.90	2.77	3.78
	0.2894	0.0336	0.0040	0.0001	0.2206	0.0569	0.0056	0.0002
Natt	−1.915	−1.871	−2.205	−0.358	−1.432	−1.589	−1.629	−0.226
	−2.70	−2.54	−1.83	−0.30	−1.95	−2.13	−1.22	−0.18
	0.0070	0.0110	0.0676	0.7666	0.0517	0.0335	0.2220	0.8551
Nalt			−1.277	−1.482			−1.094	−1.419
			−6.74	−7.56			−4.17	−5.98
			0.0000	0.0000			0.0000	0.0000
D max	4.017	3.724	2.056	0.810	2.886	2.766	1.377	0.652
	3.55	3.16	1.59	0.62	2.35	2.20	0.94	0.49
	0.0004	0.0016	0.1120	0.5333	0.0187	0.0279	0.3454	0.6266
D dif	−1.057	−1.442	0.059	0.112	−0.693	−0.937	0.079	0.118
	−2.16	−2.78	0.24	0.48	−1.34	−1.65	0.32	0.50
	0.0309	0.0054	0.8113	0.6334	0.1788	0.0994	0.7489	0.6168
Baseline					3.185	1.649	0.786	0.430
					2.28	2.17	0.98	0.47
					0.0225	0.0302	0.3253	0.6410

8. Discussion and conclusions

A considerable amount of research effort has been invested in trying to answer the question of how to design DCEs to extract preference information efficiently. Recent work focused on the key role of the nature of the explicit or implicit prior information analysts are willing to assume about the model to be fit to the data and the unknown parameters of that model. We ask a different question in this article that to our knowledge has not been previously addressed. That is, can the particular experimental

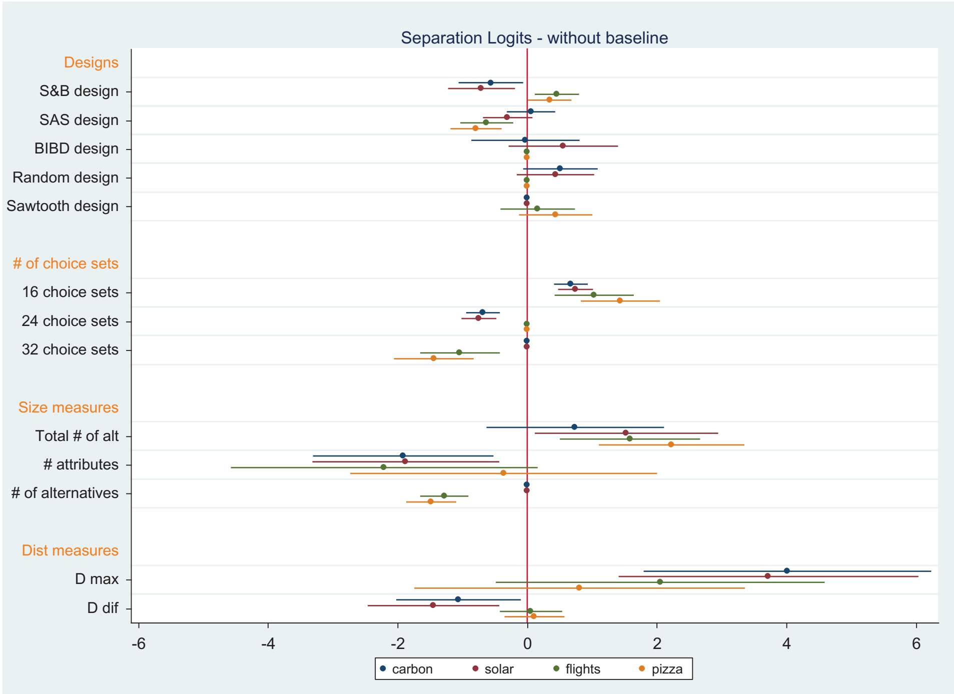

Figure 3. Logit results without baseline.

design used to construct a DCE induce potential demand artifacts in a range of real DCE data sets? The data sets we studied spanned a reasonably large range in terms of types of goods, numbers of choice options, numbers of attributes, and numbers of attribute levels. In particular, we studied four DCE data sets that exhibit large differences in (a) topics of study (e.g., emissions trading schemes, delivered pizza services, and airline flights), (b) numbers of attributes, (c) numbers of attribute levels, and (d) numbers of choice options. These DCE design features have been shown (to greater or lesser degrees) to influence how people respond to choice tasks (e.g., Caussade et al. 2005). We crossed these factors with a range of commonly used experimental designs to allow us to investigate and quantify systematic differences in outcomes associated with types of designs.

Our results indicate that different experimental designs produce choice sets that make it easier or harder to employ simple decision rules, which in turn allow choices to be made on one or a small number of attribute levels. The range of optimally efficient designs (generated under different assumptions) that we studied tend to produce options in choice sets where every level of every attribute appears once, with Street and Burgess (2007) designs being the most pronounced in this regard. Empirically, there is a difficulty in any real data set of interpreting the behavioral outcomes because some people may express their preferences in way(s) that involve putting zero weight on some or possibly all but one attribute. However, one should not expect to observe different degrees of such behavior across different experimental designs unless particular designs induce greater/lesser tendency to engage in such behavior. Because such preferences are generally

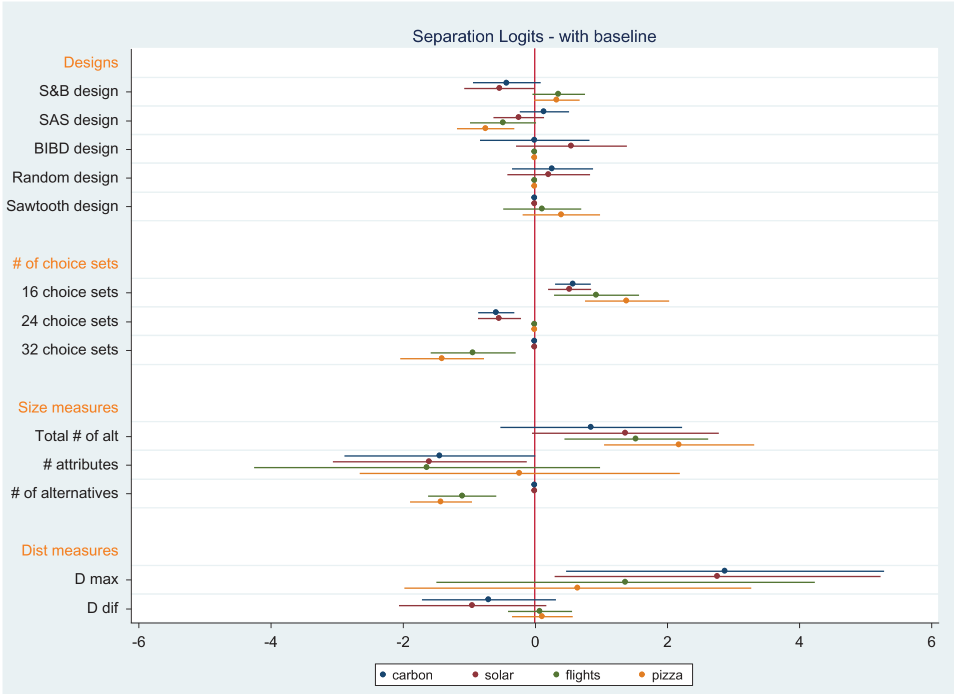

Figure 4. Logit results with baseline.

Table 5. Tests of equality between data sets.

Stacked data sets	Common intercept LR	DF	p Value	Different intercept LR	DF	p Value
Carbon & Solar	45.0	9	0.0000	7.9	8	**0.4419**
Carbon & Flights	67.3	6	0.0000	52.1	5	0.0000
Solar & Pizza	51.1	6	0.0000	49.5	5	0.0000
Solar & Flights	39.0	6	0.0000	33.7	5	0.0000
Carbon & Pizza	31.5	6	0.0000	31.4	5	0.0000
Flights & Pizza	8.5	9	**0.4886**	5.5	8	**0.7071**
Carbon & Solar & Flights	108.4	15	0.0000	57.8	13	0.0000
Carbon & Solar & Pizza	93.5	15	0.0000	55.8	13	0.0000
Carbon & Flights & Pizza	93.5	15	0.0000	82.5	13	0.0000
Solar & Flights & Pizza	56.0	15	0.0000	52.4	13	0.0000
Carbon & Solar & Flights & Pizza	139.8	24	0.0000	93.5	21	0.0000

Note. LR—Likelihood ratio; DF—degrees of freedom.

regarded as aberrant, finding more of such behavior is suggestive of a design-induced demand artifact sensitive to encouraging such behavior that is not confounded with the impacts of the design on the error variance.

We found that different designs induce rates of perfect separation at an individual level that exceed what was expected given their different statistical properties. This induced behavior may at best only impact the precision with which preference parameters are estimated. Nonetheless, this effect should be taken into account when choosing a design by looking at its statistical efficiency. In the worst case the induced behavior may also produce biases in estimates of preference parameters.

These results also have relevance for estimating the distribution of preferences. That is, designs that generate higher separation require more reliance on restrictions across individuals in the form of higher degrees of smoothing to estimate distributions of preference parameters compared with designs that can produce more identification of preferences at the individual level.

We believe that our test for a separating hyperplane can potentially be adapted to look at which generalizations of the conditional multinomial logit model are most consistent with the underlying individual data. For instance, the popular latent class specification effectively assumes there are a relatively small number of "preference" types and that all individuals can be completely characterized by one of them. This suggests any individual's choices could be characterized being perfectly explained by a one of a small number of distinct separating hyperplanes. Likewise, the popular random parameter (mixed) logit model with the usual continuous normally distributed parameter assumption has implications for the distribution of separating hyperplanes across individuals. The challenge in using the separating hyperplane test in this way is the development of experimental designs that would clearly distinguish between competing hypotheses about the distribution of preference parameters.

References

Adamowicz, V., D. Dupont, and A. Krupnick. 2006. Willingness to pay to reduce community health risks from municipal drinking water, a stated preference study. Paper presented at the Third World Congress of Environmental and Resource Economics, Kyoto, Japan, August.

Anderson, D. A., and J. B. Wiley. 1992. Efficient choice set designs for estimating availability cross-effects models. *Marketing Letters* 3 (4):357–70.

Arentze, T., A. Borgers, H. Timmermans, and R. Del Mistro. 2003. Transport stated choice responses: Effects of task complexity, presentation format and literacy. *Transportation Research Part E* 39 (3):229–44.

Atkinson, A. C., and L. M. Haines. 1996. Designs for nonlinear and generalized linear models. In *Handbook of statistics*, ed. S. Ghosh and C. R. Rao, Vol. 13, 437–75. Amsterdam, The Netherlands: Elsevier.

Bateman, I., R. T. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, M. Jones-Lee, G. Loomes, S. Mourato, E. Ozdemiroglu, D. W. Pearce, R. Sugden, and J. Swanson. 2003. *Economic valuation with stated preference techniques: A manual*. Northampton, MA: Edward Elgar.

Beck, M., T. Kjaer, and J. Lauridsen. 2011. Does the number of choice sets matter?: Results from a Web survey applying a discrete choice experiment. *Health Economics* 20 (3):273–86.

Bliemer, M. C. J., and J. M. Rose. 2010. Serial choice conjoint analysis for estimating discrete choice models. In *Choice modelling: The state of the art and the state of practice*, eds. S. Hess and A. Daly, 139–61. Northampton, MA: Edward Elgar.

Bliemer, M. C. J., and J. M. Rose. 2011. Experimental design influences on stated choice outputs: An empirical study in air travel choice. *Transportation Research Part A* 45 (1):63–79.

Bliemer, M. C. J., J. M. Rose, and S. Hess. 2008. Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling* 1 (1):98–127.

Box, G. E., J. S. Hunter, and W. G. Hunter. 2005. *Statistics for experimenters: Design, innovation, and discovery*, 2nd ed. New York, NY: Wiley-Interscience.

Boxall, P., W. L. Adamowicz, and A. Moon. 2009. Complexity in choice experiments: Choice of the status quo alternative and implications for welfare measurement. *Australian Journal of Agricultural and Resource Economics* 53 (4):503–19.

Boyle, K. J., and S. Özdemir. 2009. Convergent validity of attribute-based, choice questions in stated-preference studies. *Environmental Resource Economics* 42 (2):247–64.

Brazell, J. D., C. G. Diener, E. Karniouchina, W. L. Moore, V. Severin, and P. F. Uldry. 2006. The no-choice option and dual response choice designs. *Marketing Letters* 17 (4):255–68.

Brazell, J. D., and J. J. Louviere, 1998. Length effects in conjoint choice experiments and surveys: An explanation based on cumulative cognitive burden. Working paper, Department of Marketing, University of Sydney, Sydney, Australia, July.

Carlsson, F., and P. Martinsson. 2003. Design techniques for stated preference methods in health economics. *Health Economics* 12 (4):281–94.

Carson, R. T., and T. Groves. 2007. Incentive and informational properties of preference questions. *Environmental and Resource Economics* 37 (1):181–210.

Carson, R. T., and T. Groves. 2011. Incentive and information properties of preference questions: Commentary and extensions. In *The international handbook on non-market environmental valuation*, ed. J. Bennett, 300–21. Northampton, MA: Edward Elgar.

Caussade, S., J. De D'Ortúzar, L. Rizzi, and D. A. Hensher. 2005. Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B* 39 (7):621–40.

Chaloner, K., and I. Verdinelli. 1995. Bayesian experimental design: A review. *Statistical Science* 10 (3):273–304.

Day, B., I. J. Bateman, R. T. Carson, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa, and P. Wang. 2012. Ordering effects and choice set awareness in repeat-response stated preference studies. *Journal of Environmental Economics and Management* 63 (1):73–91.

Dellaert, B. G. C., J. D. Brazell, and J. J. Louviere. 1999. The effect of attribute variation on consumer choice consistency. *Marketing Letters* 10 (2):139–47.

DeShazo, J. R., and G. Fermo. 2002. Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *Journal of Environmental Economics and Management* 44 (1):123–43.

Dhar, R. 1997. Consumer preference for a no-choice option. *Journal of Consumer Research* 24 (2):215–31.

Dhar, R., and I. Simonson. 2003. The effect of forced choice on choice. *Journal of Marketing Research* 40 (2):146–60.

Farrar, S., and M. Ryan. 1999. Response-ordering effects: A methodological issue in conjoint analysis. *Health Economic Letters* 8 (1):75–79.

Finney, D. J. 1978. *Statistical method in biological assay*. London, UK: Charles Griffin.

Goegebeur, Y., P. Goos, and M. L. Vandebroek, 2007. A hierarchical Bayesian approach to robust parameter design. Working paper, Department of Decision Sciences and Information Management, Katholieke Universiteit Leuven, Leuven, Belgium.

Green, P. E., and V. Srinivasan. 1990. Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing* 54 (1):3–19.

Hensher, D. A. 2004. Accounting for stated choice design dimensionality in willingness to pay for travel time savings. *Journal of Transport Economics and Policy* 38 (3):425–46.

Hensher, D. A. 2006. How do respondents process stated choice experiments? Attribute consideration under varying information load. *Journal of Applied Econometrics* 21 (6):861–78.

Hensher, D. A., J. M. Rose, and W. H. Greene. 2005. *Applied choice analysis: A primer*. New York, NY: Cambridge University Press.

Hensher, D. A., P. R. Stopher, and J. J. Louviere. 2001. An exploratory analysis of the effects of numbers of choice sets in designed choice experiments: An airline choice application. *Journal of Air Transport Management* 7 (6):373–79.

Huber, J., and K. Zwerina. 1996. The importance of utility balance in efficient choice designs. *Journal of Marketing Research* 33 (3):307–17.

Kanninen, B. J. 2002. Optimal design for multinomial choice experiments. *Journal of Marketing Research* 39 (2):214–17.

Kessels, R., B. Jones, P. Goos, and M. Vandebroek. 2009. An efficient algorithm for constructing Bayesian optimal choice designs. *Journal of Business and Economic Statistics* 27 (2):279–91.

Kjaer, T., M. Bech, D. Gyrd-Hansen, and K. Hart-Hansen. 2006. Ordering effect and price sensitivity in discrete choice experiments: Need we worry? *Health Economics* 15 (11):1217–28.

Krosnick, J. A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5 (3):213–36.

Kuhfeld, W. F., R. D. Tobias, and M. Garratt. 1994. Efficient experimental design with marketing research applications. *Journal of Marketing Research* 31 (4):545–57.

Lazari, A. G., and D. A. Anderson. 1994. Designs of discrete choice set experiments for estimating both attribute and availability cross effects. *Journal of Marketing Research* 31 (3):375–83.

Louviere, J. J., 2011. Modelling single individuals: The journey from psych lab to app store. Keynote Address, International Choice Modelling Conference, Leeds, UK.

Louviere, J. J. 2013. Modelling single individuals: The journey from psych lab to app store. In *Choice modelling: The state of the art and the state of practice*, ed. S. Hess and A. Daly, 1–47. Northhampton, MA: Edward Elgar.

Louviere, J. J., D. A. Hensher, and J. D. Swait. 2000. *Stated choice methods: Analysis and applications*. New York, NY: Cambridge University Press.

Louviere, J. J., T. Islam, N. Wasi, D. Street, and L. Burgess. 2008. Designing discrete choice experiments: Do optimal designs come at a price? *Journal of Consumer Research* 35 (2):360–75.

Louviere, J. J., D. Pihlens, and R. T. Carson. 2011. Design of discrete choice experiments: A discussion of issues that matter in future applied research. *Journal of Choice Modelling* 4 (1):1–8.

Mangasarian, O. L. 1965. Linear and nonlinear separation of patterns by linear programming. *Operations Research* 13 (3):444–52.

Manski, C. F. 1977. The structure of random utility models. *Theory and Decision* 8 (3):229–54.

Mazzotta, M., and J. Opaluch. 1995. Decision making when choices are complex: A test of Heiner's hypothesis. *Land Economics* 71 (4):500–15.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in economics*, ed. P. Zarembka, 105–42. New York, NY: Academic Press.

Mellers, B. A., and K. Biagini. 1994. Similarity and choice. *Psychological Review* 101 (3):505–18.

Meyer, R. J., and T. C. Eagle. 1982. Context-induced parameter instability in a disaggregate-stochastic model of store choice. *Journal of Marketing Research* 19 (1):62–71.

Ohler, T., A. Li, J. J. Louviere, and J. Swait. 2000. Attribute range effects in binary response tasks. *Marketing Letters* 11 (3):249–60.

Orne, M. T. 1962. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist* 17 (11):776–83.

Orne, M. T. 1969. Demand characteristics and the concept of quasi-controls. In *Artifact in behavioral research*, ed. R. Rosenthal and R. Rosnow, 143–79. New York, NY: Academic Press.

Rolfe, J., and J. Bennett. 2009. The impact of offering two versus three alternatives in choice modelling experiments. *Ecological Economics* 68 (4):1140–48.

Rose, J. M., and M. C. Bliemer. 2009. Constructing efficient stated choice experimental designs. *Transport Reviews* 29 (5):587–617.

Rose, J. M., D. A. Hensher, S. Caussade, J. De D'Ortuzar, and J. Rong-Chang. 2009. Identifying differences in preferences due to dimensionality in stated choice experiments: A cross cultural analysis. *Journal of Transport Geography* 17 (1):21–29.

Rose, J. M., and S. Hess. 2009. Dual response choices in reference alternative related stated choice experiments. *Transportation Research Records* 2135 (1):25–33.

Sandor, Z., and M. Wedel. 2001. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research* 38 (4):430–44.

Street, D. J., D. S. Bunch, and B. J. Moore. 2001. Optimal designs for 2 k paired comparison experiments. *Communications in Statistics-Theory and Methods* 30 (10):2149–71.

Street, D. J., and L. Burgess. 2007. *The construction of optimal stated choice experiments: Theory and methods*. New York, NY: John Wiley.

Swait, J., and W. Adamowicz. 2001. Choice environment, market complexity, and consumer behavior: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organizational Behavior and Human Decision Processes* 86 (2):141–67.

Train, K. E. 2009. *Discrete choice methods with simulation*, 2nd ed. New York, NY: Cambridge University Press.

Tudela, A., and G. Rebolledo. 2006. Optimal design of stated preference experiments when using mixed logit models. Proceedings of the European Transport Conference, Leiden, The Netherlands.

Vincent, J. R., R. T. Carson, J. R. DeShazo, K. A. Schwabe, I. Ahmad, S. K. Chong, Y. T. Chang, and M. D. Potts. 2014. Tropical countries may be willing to pay more to protect their forests. *Proceedings of the National Academy of Sciences* 111 (28):10113–18.

Viney, R., E. Savage, and J. J. Louviere. 2005. Empirical investigation of experimental design properties of discrete choice experiments in health care. *Health Economics* 14 (4):349–62.

Wittink, D. R., J. Huber, P. Zandan, and R. M. Johnson. 1992. The number of levels effect in conjoint: Where does it come from and can it be eliminated? *Sawtooth Software Conference Proceedings*, 355–64. Ketchum, ID: Sawtooth Software.

Wittink, D. R., L. Krishnamurthi, and D. Reibstein. 1989. The effect of differences in the number of attribute levels in conjoint results. *Marketing Letters* 1(2):113–23.

Yu, J., P. Goos, and M. Vandebroek. 2008. Model-robust design of conjoint choice experiments. *Communications in Statistics—Simulation and Computation* 37 (8):1603–21.

Yu, J., P. Goos, and M. Vandebroek. 2010. Comparing different sampling schemes for approximating the integrals involved in the efficient design of stated choice experiments. *Transportation Research Part B: Methodological* 44 (10):1268–89.

Yu, J., P. Goos, and M. Vandebroek. 2011. Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity. *International Journal of Research in Marketing* 28 (4):378–88.

Appendix

Table A1. Attributes of carbon data set.

Attribute number	Attribute description	Attribute levels 1	2	3	4
1	Start year	2011	2013		
2	Revenue allocation	Lower GST	Give to low-income households/seniors	Reduce business taxes	Improve government services and reducing deficit
3	Invest 20% of revenue in research and development (R&D)	No	Yes		
4	Transportation exempted for first 3 years	No	Yes		
5	Special treatment for energy sectors for first 3 years	No	Yes		
6	Method of implementing carbon reductions	Carbon trading scheme	Carbon tax	Technology standard	Hybrid scheme
7	International role for Australia	Begin large carbon reductions now	Wait until China and United States commit		
8	2020 Emission reduction target*	3%/5%	10%	20%	25%/30%

GST, goods and services tax.
*Designs with wider range use extreme values (3%/30%).

Table A2. Attributes of solar data set.

Attribute number	Attribute description	Attribute levels 1	2	3	4
1	Country of origin	Australia	China	USA	Germany
2	Capacity	1.0 kW	1.5 kW	2.0 kW	2.5 kW
3	Government rebate	\$400 × (capacity − 1) + 2000	\$2000 × (capacity − 1) + 2500	\$2500 × (capacity − 1) + 4000	\$3000 × (capacity − 1) + 5500
4	Production output warranty	15 years	30 years		
5	Size	(capacity/0.140) per kW/m^2	(capacity/0.050) per kW/m^2		
6	Product warranty	5 Year warranty—\$0	10 Year warranty—\$1000		
7	Purchase price	\$(5500 × capacity) per kWh	\$(6500 × capacity) per kWh	\$(7500 × capacity) per kWh	\$(8500 × capacity) per kWh
8	Payback time	Formula 1	Formula 2		

Table A3. Attributes of flights data set.

Attribute number	Attribute description	Attribute levels 1	2	3	4
1	Total travel time	4	5	6	7
2	Round-trip airfare (excluding tax)	\$350	\$450	\$550	\$650
3	Number of stops	None	One		
4	Juice/water/soft drinks	Not available	All free		
5	Airline	Qantas	Virgin Blue	JetStar	Australian Airlines
6	Frequent flyer club	No	Yes		
7	Food	None	Free hot meal	Free snack	Food can be purchased
8	Audio/video entertainment	Not available	Free	\$3	\$6
9	Wait in baggage claim for bags	10 minutes	20 minutes	30 minutes	40 minutes
10	Percent of time flight departs on time	100%	80%		
11	Typical wait to check in	5 minutes	10 minutes	20 minutes	40 minutes
12	Wine/beer	\$6 each	Both free		

Table A4. Attributes for pizza data set.

Attribute number	Attribute description	Attribute levels 1	2	3	4
1	Brand	Pizza Hut	Dominos	Eagle Boys	Pizza Haven
2	Price	\$12.00	\$14.00	\$16.00	\$18.00
3	Number of toppings	1	2	3	4
4	Average delivery time	10 minutes	20 minutes	30 minutes	40 minutes
5	Likely range in delivery time	Plus or minus 10%	Plus or minus 20%	Plus or minus 30%	Plus or minus 40%
6	How often pizza arrives hot	10/10	8/10	6/10	4/10
7	Type of crust	Regular	Thick	Cheese stuffed	Thin
8	Free Coke or Pepsi	No	Yes		
9	Free dessert	No	Yes		
10	Free garlic bread/bread sticks	No	Yes		
11	Free side salad	No	Yes		
12	Free hot chicken wings	No	Yes		

Table A5. Designs for carbon data set.

Version	Design	Choice tasks per respondent	Levels per attribute: Attribute number 1	2	3	4	5	6	7	8	Attributes with 2 levels	Attributes with 4 levels	Attributes total	Number of responses	Number of distinct alternatives
1	SB	16	2	4	0	0	0	4	2	2	3	2	5	55	56
2	SAS	16	2	4	0	0	0	4	2	2	3	2	5	55	60
3	RD	16	2	4	0	0	0	4	2	2	3	2	5	55	44
4	SB	24	2	4	0	0	0	4	2	2	3	2	5	55	74
5	SAS	24	2	4	0	0	0	4	2	2	3	2	5	55	70
6	RD	24	2	4	0	0	0	4	2	2	3	2	5	55	73
7	BIBD	20	2	4	0	0	0	4	2	2	3	2	5	55	16
8	BIBD	20	2	4	0	0	0	4	2	2	3	2	5	55	16
1	SB	16	2	4	0	0	0	4	2	4	2	3	5	55	64
2	SAS	16	2	4	0	0	0	4	2	4	2	3	5	55	59
3	RD	16	2	4	0	0	0	4	2	4	2	3	5	55	62
4	SB	24	2	4	0	0	0	4	2	4	2	3	5	55	83
5	SAS	24	2	4	0	0	0	4	2	4	2	3	5	55	75
6	RD	24	2	4	0	0	0	4	2	4	2	3	5	55	90
7	BIBD	20	2	4	0	0	0	4	2	4	2	3	5	55	16
8	BIBD rg	20	2	4	0	0	0	4	2	4	2	3	5	55	16
1	SB	16	2	4	2	2	2	4	2	2	6	2	8	55	64
2	SAS	16	2	4	2	2	2	4	2	2	6	2	8	55	61
3	RD	16	2	4	2	2	2	4	2	2	6	2	8	55	64
4	SB	24	2	4	2	2	2	4	2	2	6	2	8	55	93
5	SAS	24	2	4	2	2	2	4	2	2	6	2	8	55	82
6	RD	24	2	4	2	2	2	4	2	2	6	2	8	55	96
7	BIBD	20	2	4	2	2	2	4	2	2	6	2	8	55	16
8	BIBD rg	20	2	4	2	2	2	4	2	2	6	2	8	55	16
1	SB	16	2	4	2	2	2	4	2	4	5	3	8	55	64
2	SAS	16	2	4	2	2	2	4	2	4	5	3	8	55	61
3	RD	16	2	4	2	2	2	4	2	4	5	3	8	55	64
4	SB	24	2	4	2	2	2	4	2	4	5	3	8	55	96
5	SAS	24	2	4	2	2	2	4	2	4	5	3	8	55	81
6	RD	24	2	4	2	2	2	4	2	4	5	3	8	55	96
7	BIBD	20	2	4	2	2	2	4	2	4	5	3	8	55	16
8	BIBD rg	20	2	4	2	2	2	4	2	4	5	3	8	55	16

Note. SB—Street & Burgess design; SAS—SAS software design; RD—random design; BIBD—balanced incomplete block design; BIBD rg—BIBD design with a wide range for attribute.

Table A6. Designs for solar data set.

Version	Design	Choice tasks per respondent	Levels per attribute (Attribute number)								Attributes with 2 levels	Attributes with 4 levels	Attributes total	Number of responses	Number of distinct alternatives
			1	2	3	4	5	6	7	8					
1	SB	16	0	4	2	2	0	0	4	2	3	2	5	50	56
2	SAS	16	0	4	2	2	0	0	4	2	3	2	5	50	60
3	RD	16	0	4	2	2	0	0	4	2	3	2	5	50	44
4	SB	24	0	4	2	2	0	0	4	2	3	2	5	50	74
5	SAS	24	0	4	2	2	0	0	4	2	3	2	5	50	70
6	RD	24	0	4	2	2	0	0	4	2	3	2	5	50	73
7	BIBD	20	0	4	2	2	0	0	4	2	3	2	5	50	16
8	BIBD rg	20	0	4	2	2	0	0	4	2	3	2	5	50	16
9	SB	16	0	4	4	2	0	0	4	2	2	3	5	50	64
10	SAS	16	0	4	4	2	0	0	4	2	2	3	5	50	59
11	RD	16	0	4	4	2	0	0	4	2	2	3	5	50	62
12	SB	24	0	4	4	2	0	0	4	2	2	3	5	50	83
13	SAS	24	0	4	4	2	0	0	4	2	2	3	5	50	75
14	RD	24	0	4	4	2	0	0	4	2	2	3	5	50	90
15	BIBD	20	0	4	4	2	0	0	4	2	2	3	5	50	16
16	BIBD rg	20	0	4	4	2	0	0	4	2	2	3	5	50	16
17	SB	16	2	4	2	2	2	2	4	2	6	2	8	50	64
18	SAS	16	2	4	2	2	2	2	4	2	6	2	8	50	61
19	RD	16	2	4	2	2	2	2	4	2	6	2	8	50	64
20	SB	24	2	4	2	2	2	2	4	2	6	2	8	50	93
21	SAS	24	2	4	2	2	2	2	4	2	6	2	8	50	82
22	RD	24	2	4	2	2	2	2	4	2	6	2	8	50	96
23	BIBD	20	2	4	2	2	2	2	4	2	6	2	8	50	16
24	BIBD rg	20	2	4	2	2	2	2	4	2	6	2	8	50	16
25	SB	16	4	4	2	2	2	2	4	2	5	3	8	50	64
26	SAS	16	4	4	2	2	2	2	4	2	5	3	8	50	61
27	RD	16	4	4	2	2	2	2	4	2	5	3	8	50	64
28	SB	24	4	4	2	2	2	2	4	2	5	3	8	50	96
29	SAS	24	4	4	2	2	2	2	4	2	5	3	8	50	81
30	RD	24	4	4	2	2	2	2	4	2	5	3	8	50	96
31	BIBD	20	4	4	2	2	2	2	4	2	5	3	8	50	16
32	BIBD rg	20	4	4	2	2	2	2	4	2	5	3	8	50	16

Note. SB—Street & Burgess design; SAS—SAS software design; RD—Random design; BIBD—Balanced incomplete block design; BIBD rg—BIBD design with a wide range for attribute.

Table A7. Designs for flights data set.

Version	Block	Design	Choice tasks per respondent	Alternatives per choice set	Levels per attribute: Attribute number 1	2	3	4	5	6	7	8	9	10	11	12	Attributes with 2 levels	Attributes with 4 levels	Attributes total	Num. of complete responses	Number of incomplete responses	Num. of responses total	Percentage complete	Number of distinct alternatives
1	1	S&B	16	3	4	4	2	2	4	2	0	0	0	0	0	0	3	3	6	21	1	22	95.5%	48
2	1	S&B	16	4	4	4	2	2	4	2	0	0	0	0	0	0	3	3	6	21	3	24	87.5%	64
3	1	S&B	16	5	4	4	2	2	4	2	0	0	0	0	0	0	3	3	6	20	2	22	90.9%	46
4	2	SAS	16	3	4	4	2	2	4	0	4	0	0	0	0	0	2	4	6	22	4	26	84.6%	44
5	2	SAS	16	4	4	4	2	2	4	0	4	0	0	0	0	0	2	4	6	21	3	24	87.5%	64
6	2	SAS	16	5	4	4	2	2	4	0	4	0	0	0	0	0	2	4	6	20	6	26	76.9%	80
7	3	S&B	32	3	4	4	2	2	4	0	4	4	4	0	0	0	2	6	8	21	3	24	87.5%	96
8	3	S&B	32	4	4	4	2	2	4	0	4	4	4	0	0	0	2	6	8	21	4	25	84.0%	128
9	3	S&B	32	5	4	4	2	2	4	0	4	4	4	0	0	0	2	6	8	22	4	26	84.6%	160
10	4	S&B	16	3	4	4	2	2	2	2	2	0	0	2	0	0	6	2	8	21	5	26	80.8%	48
11	4	S&B	16	4	4	4	2	2	2	2	2	0	0	2	0	0	6	2	8	21	15	36	58.3%	64
12	4	S&B	16	5	4	4	2	2	2	2	2	0	0	2	0	0	6	2	8	21	6	27	77.8%	80
13	5	S&B	32	3	4	4	2	2	4	2	4	4	4	0	4	0	3	7	10	21	2	23	91.3%	96
14	5	S&B	32	4	4	4	2	2	4	2	4	4	4	0	4	0	3	7	10	21	5	26	80.8%	128
15	5	S&B	32	5	4	4	2	2	4	2	4	4	4	0	4	0	3	7	10	21	10	31	67.7%	160
16	6	SAS	32	3	4	4	2	2	4	2	2	0	2	2	2	0	7	3	10	22	7	29	75.9%	92
17	6	SAS	32	4	4	4	2	2	4	2	2	0	2	2	2	0	7	3	10	20	5	25	80.0%	128
18	6	SAS	32	5	4	4	2	2	4	2	2	0	2	2	2	0	7	3	10	20	15	35	57.1%	160
19	7	S&B	32	3	4	4	2	2	4	2	4	4	2	2	4	2	6	6	12	20	5	25	80.0%	96
20	7	S&B	32	4	4	4	2	2	4	2	4	4	2	2	4	2	6	6	12	20	8	28	71.4%	128
21	7	S&B	32	5	4	4	2	2	4	2	4	4	2	2	4	2	6	6	12	20	8	28	71.4%	160
22	8	SAS	32	3	4	4	2	2	4	2	2	2	2	2	2	2	9	3	12	20	8	28	71.4%	92
23	8	SAS	32	4	4	4	2	2	4	2	2	2	2	2	2	2	9	3	12	20	9	29	69.0%	128
24	8	SAS	32	5	4	4	2	2	4	2	2	2	2	2	2	2	9	3	12	21	8	29	72.4%	160
25	9	OTH	16	3	4	4	2	2	4	0	4	0	0	0	0	0	2	4	6	20	3	23	87.0%	32
26	9	OTH	16	4	4	4	2	2	4	0	4	0	0	0	0	0	2	4	6	20	10	30	66.7%	32
27	9	OTH	16	5	4	4	2	2	4	0	4	0	0	0	0	0	2	4	6	20	8	28	71.4%	32
28	10	OTH	32	3	4	4	2	2	4	2	4	4	4	0	4	0	3	7	10	20	3	23	87.0%	48
29	10	OTH	32	4	4	4	2	2	4	2	4	4	4	0	4	0	3	7	10	22	9	31	71.0%	48
30	10	OTH	32	5	4	4	2	2	4	2	4	4	4	0	4	0	3	7	10	20	15	35	57.1%	32
31	11	OTH	32	3	4	4	2	2	4	2	4	4	2	2	4	2	6	6	12	22	4	26	84.6%	32
32	11	OTH	32	4	4	4	2	2	4	2	4	4	2	2	4	2	6	6	12	20	10	30	66.7%	32
33	11	OTH	32	5	4	4	2	2	4	2	4	4	2	2	4	2	6	6	12	21	13	34	61.8%	48

Table A8. Designs for pizza data set.

					Levels per attribute																			
					Attribute number																			
Version	Block	Design	Choice tasks per respondent	Alternatives per choice set	1	2	3	4	5	6	7	8	9	10	11	12	Attributes with 2 levels	Attributes with 4 levels	Attributes total	Num. of complete responses	Number of incomplete responses	Num. of responses total	Percentage complete	Number of distinct alternatives
1	1	S&B	16	3	4	4	4	0	0	0	0	2	2	2	0	0	3	3	6	21	4	25	84.0%	48
2	1	S&B	16	4	4	4	4	0	0	0	0	2	2	2	0	0	3	3	6	20	5	25	80.0%	64
3	1	S&B	16	5	4	4	4	0	0	0	0	2	2	2	0	0	3	3	6	20	2	22	90.9%	46
4	2	SAS	16	3	4	4	4	4	0	0	0	2	2	0	0	0	2	4	6	21	0	21	100.0%	44
5	2	SAS	16	4	4	4	4	4	0	0	0	2	2	0	0	0	2	4	6	20	9	29	69.0%	64
6	2	SAS	16	5	4	4	4	4	0	0	0	2	2	0	0	0	2	4	6	22	2	24	91.7%	80
7	3	S&B	32	3	4	4	4	4	4	4	0	2	2	0	0	0	2	6	8	21	1	22	95.5%	96
8	3	S&B	32	4	4	4	4	4	4	4	0	2	2	0	0	0	2	6	8	20	4	24	83.3%	128
9	3	S&B	32	5	4	4	4	4	4	4	0	2	2	0	0	0	2	6	8	21	6	27	77.8%	160
10	4	S&B	16	3	4	4	2	2	0	0	0	2	2	2	2	0	6	2	8	21	3	24	87.5%	48
11	4	S&B	16	4	4	4	2	2	0	0	0	2	2	2	2	0	6	2	8	21	5	26	80.8%	64
12	4	S&B	16	5	4	4	2	2	0	0	0	2	2	2	2	0	6	2	8	22	5	27	81.5%	80
13	5	S&B	32	3	4	4	4	4	4	4	4	2	2	2	0	0	3	7	10	21	4	25	84.0%	96
14	5	S&B	32	4	4	4	4	4	4	4	4	2	2	2	0	0	3	7	10	21	4	25	84.0%	128
15	5	S&B	32	5	4	4	4	4	4	4	4	2	2	2	0	0	3	7	10	20	8	28	71.4%	160
16	6	SAS	32	3	4	4	4	2	0	2	2	2	2	2	2	0	7	3	10	22	2	24	91.7%	92
17	6	SAS	32	4	4	4	4	2	0	2	2	2	2	2	2	0	7	3	10	20	6	26	76.9%	128
18	6	SAS	32	5	4	4	4	2	0	2	2	2	2	2	2	0	7	3	10	20	7	27	74.1%	160
19	7	S&B	32	3	4	4	4	4	4	2	4	2	2	2	2	2	6	6	12	22	3	25	88.0%	96
20	7	S&B	32	4	4	4	4	4	4	2	4	2	2	2	2	2	6	6	12	20	4	24	83.3%	128
21	7	S&B	32	5	4	4	4	4	4	2	4	2	2	2	2	2	6	6	12	20	10	30	66.7%	160
22	8	SAS	32	3	4	4	4	2	2	2	2	2	2	2	2	2	9	3	12	22	2	24	91.7%	92
23	8	SAS	32	4	4	4	4	2	2	2	2	2	2	2	2	2	9	3	12	20	7	27	74.1%	128
24	8	SAS	32	5	4	4	4	2	2	2	2	2	2	2	2	2	9	3	12	21	13	34	61.8%	160
25	9	OTH	16	3	4	4	4	4	0	0	0	2	2	0	0	0	2	4	6	21	3	24	87.5%	32
26	9	OTH	16	4	4	4	4	4	0	0	0	2	2	0	0	0	2	4	6	21	5	26	80.8%	32
27	9	OTH	16	5	4	4	4	4	0	0	0	2	2	0	0	0	2	4	6	21	9	30	70.0%	32
28	10	OTH	32	3	4	4	4	4	4	4	4	2	2	2	0	0	3	7	10	21	2	23	91.3%	48
29	10	OTH	32	4	4	4	4	4	4	4	4	2	2	2	0	0	3	7	10	21	5	26	80.8%	48
30	10	OTH	32	5	4	4	4	4	4	4	4	2	2	2	0	0	3	7	10	20	7	27	74.1%	32
31	11	OTH	32	3	4	4	4	4	4	2	4	2	2	2	2	2	6	6	12	21	3	24	87.5%	32
32	11	OTH	32	4	4	4	4	4	4	2	4	2	2	2	2	2	6	6	12	20	5	25	80.0%	32
33	11	OTH	32	5	4	4	4	4	4	2	4	2	2	2	2	2	6	6	12	20	15	35	57.1%	48