JOEL HUBER and KLAUS ZWERINA*

Choice designs traditionally have been built under the assumption that all coefficients are zero. The authors show that if there are reasonable nonzero priors for expected coefficients, then these can be used to generate more statistically efficient choice designs, because the alternatives in their choice sets are balanced in utility—they have more similar choice probabilities. The authors demonstrate that the appropriate measure of choice design efficiency requires probability centering and weighting of the rows of the design matrix, and they illustrate how this criterion enables the analyst to appropriately trade off utility balance against three other principles: orthogonality, level balance, and minimal overlap. Two methods, swapping and relabeling attribute levels, provide complementary ways to increase the utility balance of choice designs. The authors apply a process for generating utility-balanced designs to five different choice designs and show that it reduces by 10–50% the number of respondents needed to achieve a specific error level around the parameters. A sensitivity analysis reveals that these gains are diminished, but still substantial, despite strong misspecifications of prior parameter estimates.

The Importance of Utility Balance in Efficient Choice Designs

Because of their ability to mimic realistic market decisions, choice experiments are increasingly used to model market demand (Carson et al. 1994). Choice experiments allow an estimate of the impact of product features on choice in a competitive context (Louviere and Woodworth 1983) and an assessment of the differential impact of price for brands in the presence of different competitors (Anderson and Wiley 1992; Batsell and Louviere 1991). What choice experiments gain in realism, compared to ratings-based conjoint, they lose in reliability. That is, reliable parameter estimates require pooling the choices from different respondents. A typical choice experiment might require pooling the choices of 100 respondents each of whom makes eight choices.

There are various ways to decrease the number of respondents necessary for reliable within-segment analysis. One way is to "explode" rank order data into a greater number of inferred choices (Chapman and Staelin 1982). Reliability also can be increased by adjusting ratings-based conjoint to correspond to choices (Huber et al. 1993). However, both of these methods depend on the correspondence between choices and either ranking or rating, which may not be justified (Ben-Akiva, Morikawa, and Shiroshi 1992).

Within the context of choice designs, several authors have proposed designs that decrease the number of respondents or choices per respondent needed to achieve an expected level of accuracy. For example, Bunch, Louviere, and Anderson (1994) provide statistically efficient main-effect designs, and Anderson and Wiley (1992) and Lazari and Anderson (1994) provide statistically efficient cross-effect designs. Kuhfeld, Tobias, and Garratt (1994) demonstrate the use of a computerized search to find nearly optimal designs for large, complex, and asymmetric problems. Choice-Based Conjoint Analysis (CBC) (Sawtooth Software 1993) provides computer-implemented randomized tasks that are reasonably efficient for main effects and allow for the estimation of all interactions.

All of these techniques generate reasonable designs if a person does not use prior information about the expected coefficients—that is, if the analyst expects all of the coefficients to be zero. However, with the level of knowledge typically available in most marketing studies, this strategy is certainly too conservative and may be inappropriate. We often know, for example, that low prices are generally preferred to high ones, that some brands are consistently desired over others, and that rankings of some product features are shared by different consumers. Furthermore, in practical marketing research, most surveys are tested with a small-sample pretest, the outcome of which provides rea-

*Joel Huber is a professor, Fuqua School of Business, Duke University. Klaus Zwerina is visiting scholar at Fuqua School of Business, Duke University and a doctoral student under Professor Friedhelm Bliemel, Universität Kaiserslautern, Germany. This paper emerged from Klaus Zwerina's dissertation research while at Duke. The authors thank Richard Staelin, Warren Kuhfeld, Richard Johnson, Timothy Renken, and Marcus Christen for their help.

Journal of Marketing Research
Vol. XXXIII (August 1996), 307–317

sonable priors for the choice model. We show how this prior information can be used to generate choice designs that are substantially more efficient than those heretofore available. The primary mechanism by which this improvement is achieved is by balancing the utilities of the alternatives offered in the choice sets. That is, we define alternatives within each choice set to have more equal choice probabilities.

Although it has not been reflected in published choice designs, the idea of utility balance is not new. The pair comparisons in the popular conjoint package, Adaptive Conjoint Analysis (ACA) (Johnson 1987), balance profiles so that respondents are as nearly indifferent to a pair as possible. In the pure choice realm, Thurstone's (1927) law of comparative judgments develops a discriminal process for choices among pairs of objects. The process, which is the intellectual precursor of contemporary choice models, seeks to place objects such that the probability of selecting one over another corresponds to their Euclidean distance, thus ensuring that objects with large preference differences will be far apart on the scale. To efficiently generate such a scale, Torgerson (1958, p. 138) provides strategies that focus on similar pairs while also undersampling dissimilar ones. The justification for undersampling dissimilar pairs derives from the nonlinearity of the cumulative normal distribution in the Thurstone model, and in general from any S-shaped cumulative probability function such as the logistic. In particular, the choice among a pair in which one alternative is chosen almost all of the time anchors the extreme values of the S-shaped cumulative probability function but provides little information on the "slope" of that function. Generally, choice sets that generate extreme probabilities are less effective at constraining the parameters of the choice model than are moderate ones.

Much in the spirit of Torgerson's (1958) sampling strategies, Krieger and Green (1991) recommend permuting attribute levels in choice sets to minimize instances of dominated alternatives. Dominated alternatives within a choice set are an extreme case in which the choice provides no new information, but the problem also occurs with nondominating choice sets in which one alternative is an easy choice. The utility-balanced choice designs that we discuss subsequently similarly limit the number of dominated alternatives, but they do so in a general context of improving the efficiency of the entire choice design. We provide two utility-balance strategies and demonstrate how increasing utility balance with these operations can produce designs with smaller expected error around the estimated parameters. At some point, however, balancing the utilities of the alternatives in each choice set can increase the colinearity in the design and thereby deteriorate its efficiency. To resolve this conflict between greater utility balance and nonorthogonality, we provide criteria for identifying the level of utility balance that produces more efficient choice designs.

Specifically, we seek to answer three questions:

1. What is the theoretically appropriate efficiency measure for a choice design?
2. What processes can be used to generate more efficient choice designs given prior knowledge of the parameter estimates?
3. How large is the efficiency gain from these modifications in utility balance, and how robust are these gains to misspecified priors?

To answer the first question, we show that the efficiency of a choice design depends on the design matrix in which alternatives are "centered" within choice sets. To answer the second question, we provide two mechanisms to generate more efficient designs through utility balance: attribute relabeling and swapping. Finally, to answer the third question, we show that efficiency gains of 10–50% are achieved for typical choice designs and that these gains are not sensitive to monotone misspecifications of the prior parameter estimates.

THEORETICAL DEVELOPMENT

We derive the measure of efficiency for a choice design that has a set of priors on the coefficients. The calculation of this measure is illustrated in the Appendix in conjunction with the first example design. The derivation basically follows the seminal work of McFadden (1974) and is summarized in a variety of sources (Ben-Akiva and Lerman 1985; Bunch, Louviere, and Anderson 1994; Madalla 1983). The standard random utility framework posits that choice probabilities are derived from a fixed and variable component that is associated with the evaluation of each alternative, namely,

$$(1) \qquad u_i = x_i\beta + e_i.$$

Here, u_i is the momentary utility of alternative i, x_i is a row vector of K characteristics describing i, β is a column vector of weights associated with each of those K characteristics, and e_i is a random error term with expectation zero and variance σ_e^2.

Next, consider a choice experiment with N choice sets, C_n, indexed by n = 1,...,N, where each set is characterized by a set of alternatives $C_n = \{x_{1n},...,x_{J_n n}\}$. Under the logit assumption that the e_i's are independently, identically Gumbel-distributed, the probability that a consumer will choose alternative i from the choice set C_n is

$$(2) \qquad P_{in}(X_n, \beta) = \frac{e^{x_{in}\beta}}{\sum_{j=1}^{J_n} e^{x_{jn}\beta}}.$$

Within choice set C_n, let x_{jn} denote the j^{th} row of a matrix X_n, where x_{jn} is the K-dimensional row vector associated with an alternative, $j \in C_n$. If $M = \Sigma_{n=1}^{N} J_n$ is the total number of alternatives in the choice experiment, then the choice design matrix, X, reflecting the concatenation of all submatrices, X_n, is of size $M \times K$. Furthermore, if Y is a matrix of choices with elements y_{in}, each of which equals one if alternative i is chosen in C_n and equals zero otherwise, and if we assume that each choice reflects an independent draw from a multinomial distribution, then the log-likelihood of a given sample, Y, is

$$(3) \quad L(Y \mid X, \beta) = \sum_{n=1}^{N} \sum_{j=1}^{J_n} y_{jn} \ln(P_{jn}(X_n, \beta)) + \text{constant}.$$

Here, $P_{jn}(X_n,\beta)$, generated from Equation 2, is the choice probability of alternative j in set C_n, which depends on the characteristics of the alternatives, X_n, and the true parameter vector β.

Maximizing Equation 3 yields the maximum likelihood estimator, $\hat{\beta}$, of a choice model, with a particular choice design. McFadden (1974) shows that the distribution of $\hat{\beta}$ is asymptotically normal with mean β and covariance matrix

$$\text{(4a)}\qquad \Omega_P^{-1} = (Z'\,PZ)^{-1} = [\sum_{n=1}^{N}\sum_{j=1}^{J_n} z'_{jn}P_{jn}z_{jn}]^{-1},$$

where P is an M × M diagonal matrix with elements P_{jn}, and Z is an M × K matrix with rows

$$\text{(4b)}\qquad z_{jn} = x_{jn} - \sum_{i=1}^{J_n} x_{in}P_{in}.$$

Equation 4 takes a particularly simple form when the best guess is that the β's are equal to zero. In this case, the choice probabilities of each alternative in choice set C_n are equal to $1/J_n$, and the covariance matrix simplifies to

$$\text{(5a)}\qquad \Omega_0^{-1} = (Z'\,P_0Z)^{-1} = [\sum_{n=1}^{N}\frac{1}{J_n}\sum_{j=1}^{J_n} z'_{jn}z_{jn}]^{-1},$$

where

$$\text{(5b)}\qquad z_{jn} = x_{jn} - \bar{x}_n \text{ with } \bar{x}_n = \frac{1}{J_n}\sum_{j=1}^{J_n} x_{jn}.$$

Thus, the appropriate design matrix with $\beta = 0$ centers each attribute within each choice set. To emphasize the two different ways in which the choice probabilities affect the covariance matrix of $\hat{\beta}$, we denote Equation 4 as a *probability-centered* estimate of error, and Equation 5 as a *zero-centered* or *utility-neutral* estimate of error.

Equations 4 and 5 are valuable in two ways. First, though the zero-centered estimate of error derives directly from the original McFadden (1974) article, to our knowledge it has not been put in this form. Indeed, zero centering the design matrix provides a general way to estimate efficiencies of choice designs under the null hypothesis that $\beta = 0$, and thus it is an appropriate first step in evaluating any choice design. Second, Equation 4 is routinely used in an ex post sense to estimate the accuracy of a choice model once a person has a set of coefficients, $\hat{\beta}$, and uses these to estimate the P_{in}'s. However, we demonstrate its usefulness ex ante in generating efficient choice designs with prior estimates of β.

For a Fisher information matrix, Ω, several established summary measures of error are useful in comparing designs (Bunch, Louviere, and Anderson 1994; Kuhfeld, Tobias, and Garratt 1994). We focus on minimizing error around the estimated parameters. Two measures, inversely related to A- and D-efficiency are

$$\text{(6)}\qquad \text{A-error} = (\text{trace } \Omega^{-1})/K,$$

and

$$\text{(7)}\qquad \text{D-error} = (\det\ \Omega^{-1})^{1/K}.$$

Our results are substantially unchanged with either measure; so, in keeping with previous practice (Bunch, Louviere, and Anderson 1994; Kuhfeld, Tobias, and Garratt 1994), we use D-error in the summaries of our results. We denote D_P-error as the probability-centered estimate of error, and D_0-error as the utility-neutral estimate of error that is appropriate when all β's are zero.

We next discuss general properties of efficient choice designs and show how prior information about the expected utilities can generate choice designs that are 10–50% more efficient.

WHAT MAKES CHOICE DESIGNS EFFICIENT?

Four properties characterize efficient choice designs. Two of these, *level balance* and *orthogonality,* also characterize linear designs. The third, *minimal overlap,* becomes relevant for choice designs, because each attribute level is only meaningful in comparison to others within a choice set. The fourth property is *utility balance.* We discuss the first three criteria and then show how our new criterion can improve designs still further.

Level balance is often just termed *balance* (Kuhfeld, Tobias, and Garratt 1994), but we retain the prefix to distinguish it from utility balance. Level balance is the requirement that the levels of an attribute occur with equal frequency. For example, each level of a three-level attribute should occur in precisely one-third of the cases.

Orthogonality, the second criterion, is satisfied when the joint occurrence of any two levels of different attributes appear in profiles with frequencies equal to the product of their marginal frequencies (Addelman 1962). Thus, if there is level balance, the joint occurrence of any combination of a three- and a four-level attribute must occur in exactly one-twelfth of the cases. However, for many design specifications, level balance and orthogonality conflict so that one cannot be satisfied without degrading the other. Consider, for example, a design with one three-level and one four-level attribute in which the number of alternatives is not a multiple of twelve. To create such a design with, say, 16 profiles, Addelman (1962) recommends taking a 4^{4-2} design and creating a three-level attribute by collapsing two of its four levels into one; for example, by making all level 4s be level 3. Collapsing levels preserves orthogonality but violates level balance, because the collapsed level occurs eight times, whereas the others occur only four times. Kuhfeld, Tobias, and Garratt (1994) show how the OPTEX procedure (SAS Institute 1995) can produce more efficient designs while achieving neither perfect level balance nor orthogonality. Note that, except for special cases, optimality for linear designs involves trading off incompatible criteria of level balance and orthogonality.

Minimal level overlap becomes important for choice designs, because the contrasts between attribute levels are only meaningful as differences *within* a choice set. The impact of this requirement is expressed in the centering of attributes within each choice set found in equations 4 and 5. Minimal overlap means that the probability that an attribute level repeats itself in each choice set should be as small as possible. The cost of violating this criterion can be seen most clearly when the levels of one attribute are the same across all alternatives within a choice set. Here, the choice set provides *no* information on the attribute's value. We use the criteria of level balance, orthogonality, and minimal overlap to generate optimal utility-neutral choice designs—designs that minimize D_0-error. We then show, based on D_P-error, that we can improve the efficiencies of these designs by balancing the utilities of the alternatives in each choice set.

Table 1
COMPARISION OF ORIGINAL AND SWAPPED $3^3/3/9$ DESIGN

		Original $3^3/3/9$ Design					*Swapped $3^3/3/9$ Design*				
Set	*Alternative*	*Attribute A*	*Attribute B*	*Attribute C*	*Sum*	*Probability*	*Attribute A*	*Attribute B*	*Attribute C*	*Sum*	*Probability*
1	I	1	1	1	3	.002	3	1	3	7	.665
	II	2	2	2	6	.047	2	2	2	6	.245
	III	3	3	3	9	.951	1	3	1	5	.090
2	I	1	2	2	5	.045	3	1	2	6	.333
	II	2	3	3	8	.910	2	3	1	6	.333
	III	3	1	1	5	.045	1	2	3	6	.333
3	I	1	3	3	7	.488	3	2	1	6	.333
	II	2	1	1	4	.024	2	1	3	6	.333
	III	3	2	2	7	.488	1	3	2	6	.333
4	I	2	1	3	6	.333	3	1	1	5	.090
	II	3	2	1	6	.333	1	3	3	7	.665
	III	1	3	2	6	.333	2	2	2	6	.245
5	I	2	2	1	5	.045	2	1	3	6	.245
	II	3	3	2	8	.910	3	3	1	7	.665
	III	1	1	3	5	.045	1	2	2	5	.090
6	I	2	3	2	7	.488	2	3	1	6	.245
	II	3	1	3	7	.488	3	2	2	7	.665
	III	1	2	1	4	.024	1	1	3	5	.090
7	I	3	1	2	6	.333	1	3	2	6	.245
	II	1	2	3	6	.333	3	1	1	5	.090
	III	2	3	1	6	.333	2	2	3	7	.665
8	I	3	2	3	8	.910	2	3	2	7	.665
	II	1	3	1	5	.045	3	2	1	6	.245
	III	2	1	2	5	.045	1	1	3	5	.090
9	I	3	3	1	7	.488	1	2	3	6	.333
	II	1	1	2	4	.024	3	1	2	6	.333
	III	2	2	3	7	.488	2	3	1	6	.333
D_p-error						.381					.280

GENERATION OF EFFICIENT CHOICE DESIGNS

We denote a choice design that is characterized by, say, three attributes each at two levels, and four alternatives in each of nine choice sets, by $2^3/4/9$. More complex design families can also be described: $3^2 \times 4^3/2/16 + 2^2 \times 4^3/3/15$ is a mixed design with 16 pairs and 15 triples in which attributes have different numbers of levels in the different subdesigns.

Our process of generating utility-balanced choice designs starts with an efficient utility-neutral design. Although there are many ways to generate D_0-efficient designs, we focus on shifted or cyclic designs, first developed for choice experiments by Bunch, Louviere, and Anderson (1994). Cyclic designs are easily generated and have perfect level balance, orthogonality, and minimal overlap. We consider five families of designs, beginning with simple designs to illustrate the principles, and then moving to more complex, but more managerially useful, designs.

The first design we consider is the $3^3/3/9$ design shown in the left-hand panel of Table 1. It has three attributes, each defined with three levels, and nine choice sets comprised of three alternatives. The first alternatives within each choice set represent a classic nine-profile, 3^{3-1} orthogonal array (Addelman 1962). Subsequent alternatives are constructed by adding cyclically generated alternatives to each set. The attribute levels of these new alternatives add one to the level of the previous alternative until it is at the highest level, at which point the assignment re-cycles to the lowest level. In this way, designs with the same number of alternatives as (or fewer alternatives than) the maximum number of levels can be generated easily.

These cyclically generated alternatives are beneficial because they mirror the perfect level balance and orthogonality of the seed array. That is, the frequency of occurrence of each pair of attribute levels is equivalent to the product of their marginal frequencies. In the design shown, each level occurs in one-third of the profiles, and any pair of levels occurs in precisely one-ninth of the profiles. Finally, because of the symmetry of the design, there is minimal overlap—each level occurs only once for each attribute within a choice set.

The three properties—level balance, orthogonality, and minimal overlap—mean that the original $3^3/3/9$ design is optimal with respect to the criterion of D_0-error. The application of equations 5 and 7 result in a D_0-error of .19, which assumes $\beta = 0$. To illustrate the impact of nonzero β's, we initially space the partworths for each of the attributes evenly between –1 and 1. This equal spacing is convenient, because the partworths become a translation of the level labels. Thus, a partworth of –1 corresponds to level 1, 0 to level 2, and 1 to level 3. Given this scaling, a measure for the utilities of each alternative is the sum measure shown in Table 1. We evaluate the logit probabilities that follow, using Equation 2 in which the β's are the original partworth levels. In the Appendix, we show some intermediate results in calculating D_P-error.[1] Note that the D_P-error of .38 is not directly comparable with the D_0-error of .19, because the former reflects errors around the partworths –1,0,1, whereas the latter reflects errors around the partworths 0,0,0.

In the second panel of Table 1, we show that the D_P-error can be improved by what we call a *swap*. A swap involves switching one level of an attribute within a choice set. Note that in the original design the first choice set is out of balance; alternative III dominates the first two, garnering 95% of the expected choices. The swaps shown switch the first and third levels of attributes A and C, thereby resulting in more equal choice probabilities. A computer program generated the design shown in Table 1 by evaluating all possible swaps for a given choice set and then executing that swap if it could improve D_P-error. The program then examined subsequent choice sets in the same way and continued to make swaps until no more improvement was possible. Although this sequential method did not guarantee the optimal design, one iteration consistently provided designs with efficiencies within 98% of the optimal swapped design.

In the first column of Table 2, we show the net gain from these swaps. Given the prior coefficients, D_P-error decreases from .38 to .28. Thus, the swapped design can use 27% fewer respondents and still have the same expected error around the parameters as does the original design. Note, however, that though D_P-error improves with swaps, D_0-error becomes 18% worse, which implies that if β is zero, swapping could reduce efficiency. Thus, the analyst must decide whether to set β to some nonzero prior vector. Any value, including zero, entails costs if it is wrong. We subsequently show that an analyst can have broad errors in nonzero priors and still gain from their use, thereby implying that there is generally a net benefit from using priors.

The analysis of the $3^3/3/9$ family is predicated on a particular set of coefficients with partworths arbitrarily ranging from –1 to 1. How big are the efficiency gains from utility balance if this range is increased or decreased? In general, the logit coefficients increase if respondents are more internally consistent and more homogeneous or if the model better specifies their choices. In any of these cases the scale of the solution increases (cf. Ben-Akiva and Lerman 1985; Swait and Louviere 1993). Looking horizontally across Table 2 shows what happens to the gain from utility balance if the coefficients are made larger or smaller by 25%.

[1]The SAS code to calculate D_P-error is available from the second author.

Table 2
GAINS IN EXPECTED ERRORS DUE TO UTILITY BALANCE FROM RELABELING AND SWAPPING

Design Family	*Original Values* ($\beta = \beta_0$)		*Less Noise* ($\beta = \beta_0 \times 1.25$)		*Greater Noise* ($\beta = \beta_0 \times .75$)	
	D_P-error [SD]	*D_o-error*	*D_P-error [SD]*	*D_o-error*	*D_P-error [SD]*	*D_o-error*
$3^3/3/9$						
Original	.381 [.002]	.192	.475 [.012]	.192	.305 [.000]	.192
Swapped	.280	.227	.311	.223	.259	.223
Total % Gains	27%	–18%	34%	–16%	15%	–16%
$3^4/2/15$						
Average relabeled[a]	.447 [.072]	.163	.611 [.140]	.163	.325 [.033]	.163
Best relabeled[b]	.297	—[d]	.335	—	.256	—
Swapped[c]	.253	.215	.265	.224	.231	.185
Total % Gains	43%	–32%	57%	–37%	29%	–14%
$4^4/4/16$						
Average relabeled	.307 [.024]	.157	.384 [.044]	.157	.244 [.012]	.157
Best relabeled	.263	—	.301	—	.222	—
Swapped	.198	.178	.208	.180	.188	.178
Total % Gains	36%	–13%	46%	–14%	23%	–13%
$4 \times 3^3/3/48$						
Average relabele[d]	.231 [.013]	.102	.295 [.024]	.102	.178 [.006]	.102
Best relabeled	.199	—	.238	—	.163	—
Swapped	.142	.112	.146	.116	.133	.113
Total % Gains	39%	–9%	50%	–13%	25%	–11%
$9 \times 8 \times 4 \times 3^4 \times 2^3/3/63$						
Average relabeled	.227 [.025]	.068	.302 [.043]	.068	.165 [.013]	.068
Best relabeled	.161	—	.196	—	.130	—
Swapped	.084	.071	.089	.071	.075	.075
Total % Gains	63%	–5%	71%	–5%	55%	–11%

[a]Average D_P-error of 1000 randomly relabeled designs. Standard deviations in brackets.
[b]Best design from relabeling.
[c]Best design from relabeling and subsequent swapping.
[d]Relabeling has no affect on expected errors when β = 0.

Enlarging the coefficients increases the efficiency gains from utility balance, whereas making them smaller has the opposite effect. This result makes intuitive sense. The large coefficients generate extreme choice probabilities, thus exacerbating any initial imbalance within choice sets. In contrast, when coefficients are close to zero, the choice probabilities are nearly equivalent, and the choice design has less need for utility balancing.

To summarize, we have shown that swapping to balance utilities produces an efficiency gain even for this small design of three attributes and nine choice sets and that this gain increases with the scale of the coefficients. However, the benefit from swaps is not without risk. If the coefficients are zero, then the lack of orthogonality engendered by the swapping produces a less efficient design. The next method we consider to increase utility balance is relabeling. Relabeling is less risky because it alters utility balance without affecting orthogonality.

Relabeling involves reassigning labels to the levels of a design, for example, replacing attribute levels 1, 2, and 3 with 3, 1, and 2 (Krieger and Green 1991; Kuhfeld, Tobias, and Garratt 1994). Relabeling can improve utility balance without degrading D_0-error. Unfortunately, relabeling does not work for some small designs, such as the $3^3/3/9$ family, in which it results in virtually the same choice sets in a different order. However, it is effective for larger designs, as is illustrated on a design of 15 pairs with four attributes each at three levels ($3^4/2/15$) in Table 3. We generated the 15 core stimuli for this design with a 3^4 main effect design in 15 profiles by using the OPTEX procedure (SAS Institute 1995), and we produced the second alternatives within choice sets using cycles, modulo 3. The three levels of each of the four attributes then can be relabeled in $3! = 6$ different ways, resulting in $6^4 = 1296$ possible designs. The design shown in Table 3, part A, is one of these designs, with a D_P-error of .45, which is the same error as the expected error of a design chosen at random from the set of relabeled designs. The best possible relabeled design has a D_P-error of .30—a 33% improvement of efficiency over the average design.

The probabilities in Table 3 demonstrate the way utility balance improves the efficiency of choice designs. Although both designs have some unbalanced choice sets, the original design has far more. The original design has 11 sets with maximum probabilities exceeding .85, whereas the relabeled design has only 4 sets.

Relabeling provides the best utility balance design while also preserving orthogonality, level balance, and minimal overlap. Swapping can result in an even better design by sacrificing some orthogonality to achieve more utility balance. In Table 2, we show for the $3^4/2/15$ family that the total efficiency gain grows from 33% to 43% with swapping, though D_0-error degrades by 32%, which suggests that if an analyst were not confident that $\beta \neq 0$, he or she might choose the best relabeled design over the one that includes both relabeling and swapping.

APPLICATION OF SWAPPING AND RELABELING TO MORE COMPLEX CHOICE DESIGNS

The purpose of the previous discussion was to introduce, with simple designs and detailed examples, the concepts of swapping and relabeling to achieve utility balance. We here apply the same principles to larger and more complex designs in which it is shown that the efficiency gains due to utility balance are even greater.[2]

Consider a choice design that might be used for experimental work: four attributes, each at four levels, represented in 16 choice sets with four alternatives ($4^4/4/16$). The design was developed as before with a core design of 16 profiles defined from a 4^{4-2} orthogonal array and supplemented by competitive alternatives generated with modulo 4 cycles. We then assumed a β vector with partworth values for each attribute of –1, –1/3, +1/3, and 1 and examined the distribution of D_P-error across 1000 randomly relabeled designs. As is shown in Table 2, the average error is .31, with a standard deviation of .024. The best relabeled design has a D_P-error of .26 resulting in an efficiency improvement of 16%. Swapping increases this improvement to a total of 36%.

Next, consider a brand-specific choice design. These are designs with estimable brand-by-attribute interactions. They are important when an analyst is concerned with modeling choices among brands that themselves have different attributes. As an example, we might have four soft drink brands—Coke, Pepsi, RC, and President's Choice—each at three different prices and each possibly having different flavors and containers. Accordingly, we need a choice design in which brand, brand × price, brand × container, and brand × flavor are all estimable. This leads to a $4 \times 3^3/3/48$ (interaction) design with 27 parameters (9 parameters for main effects and 18 for interactions).

To generate such choice designs with estimable interactions requires a modification of cycling used to build alternatives from a core choice set. In particular, we must assure that, for example, Coke at its lowest price is in a choice set against Pepsi at its lowest price, a combination that would not normally be built with cycling. To make the interactions estimable, we modified the choice sets to include several such contrasts. This design was then further improved by swapping under the condition that β is zero. This process does not appear to generate the optimal (utility-neutral) design with interactions, but it is sufficient to test whether utility balance works in the case of a design with interactions.

In Table 2, we show that the average relabeled design has a D_P-error of .23, which drops to .20 (a 14% gain) if the best relabeled design is adopted. Subsequent swaps reduce the D_P-error to .14, resulting in a total efficiency gain of 39%. Thus, the application of swapping and relabeling to achieve utility balance also is effective for designs with interactions.

Finally, in Table 2, we show the efficiency gains from a large design. This complex design ($9 \times 8 \times 4 \times 3^4 \times 2^3/3/63$) involves 63 choice sets with ten factors at varying levels, requiring 29 main-effect parameters. The nine-level attribute could be used as a measured factor in its own right or used to define a blocking variable so that each respondent receives seven choices comprising one-ninth of the total design. To build this design, we used the OPTEX procedure to create a core design with 63 profiles and completed each choice set by applying cycles with modulos corresponding to the number of levels of each attribute. The 1000 randomly relabeled versions of this design yield an average D_P-error of .23. The best of these relabeled designs shows an

[2]In the interest of brevity, the particular designs will not be displayed. They are, however, available from the authors.

Table 3
COMPARISON OF ORIGINAL AND RELABELED $3^4/2/15$ CHOICE DESIGN[a]

		A. Original $3^4/2/15$ Design						B. Relabeled $3^4/2/15$ Design					
Set	*Alternative*	*Attribute A*	*Attribute B*	*Attribute C*	*Attribute D*	*Sum*	*Probability*	*Attribute A*	*Attribute B*	*Attribute C*	*Attribute D*	*Sum*	*Probability*
1	I	1	3	1	3	8	.731	3	2	3	2	10	.953
	II	2	1	3	1	7	.269	1	3	2	1	7	.047
2	I	1	2	3	1	7	.119	3	1	2	1	7	.500
	II	2	3	2	2	9	.881	1	2	1	3	7	.500
3	I	1	1	3	2	7	.119	3	3	2	3	11	.998
	II	2	2	2	3	9	.881	1	1	1	2	5	.002
4	I	1	3	2	1	7	.731	3	2	1	1	7	.047
	II	2	1	1	2	6	.269	1	3	3	3	10	.953
5	I	1	2	2	3	8	.731	3	1	1	2	7	.500
	II	2	3	1	1	7	.269	1	2	3	1	7	.500
6	I	2	2	3	3	10	.731	1	1	2	2	6	.500
	II	3	3	2	1	9	.269	2	2	1	1	6	.500
7	I	2	3	1	2	8	.119	1	2	3	3	9	.500
	II	3	1	3	3	10	.881	2	3	2	2	9	.500
8	I	2	2	2	1	7	.119	1	1	1	1	4	.002
	II	3	3	1	2	9	.881	2	2	3	3	10	.998
9	I	2	1	2	2	7	.119	1	3	1	3	8	.500
	II	3	2	1	3	9	.881	2	1	3	2	8	.500
10	I	2	1	1	1	5	.007	1	3	3	1	8	.500
	II	3	2	3	2	10	.993	2	1	2	3	8	.500
11	I	3	3	3	3	12	.999	2	2	2	2	8	.500
	II	1	1	2	1	5	.001	3	3	1	1	8	.500
12	I	3	1	1	1	6	.119	2	3	3	1	9	.500
	II	1	2	3	2	8	.881	3	1	2	3	9	.500
13	I	3	3	3	2	11	.982	2	2	2	3	9	.500
	II	1	1	2	3	7	.018	3	3	1	2	9	.500
14	I	3	1	2	3	9	.982	2	3	1	2	8	.500
	II	1	2	1	1	5	.018	3	1	3	1	8	.500
15	I	3	2	1	2	8	.119	2	1	3	3	9	.500
	II	1	3	3	3	10	.881	3	2	2	2	9	.500
D_P-error							.449						.297

[a]Reassignments: A → [1 – 3, 2 – 1, 3 – 2], B → [1 – 3, 2 – 1, 3 – 2], C → [1 – 3, 2 – 1, 3 – 2], and D → [1 – 1, 2 – 3, 3 – 2].

efficiency gain of 30% over the average, and subsequent swapping increases the gain to a total of 63%.

To summarize, we have examined the impact of utility balance on five different designs. It is important to stress that these results are not unique to the design families shown but are typical of choice designs generally. Relabeling and swapping enable an analyst to generate utility-balanced designs with 10–50% fewer respondents that are just as accurate as utility-neutral designs. An examination of Table 2 indicates that these gains appear to be greater for large designs—presumably because, compared to designs with fewer choice sets, designs with more choice sets can be utility-balanced without greatly distorting their orthogonality. This conjecture is supported by the observation that the smallest loss in D_0-error from swapping comes from the designs that have the most excess degrees of freedom.

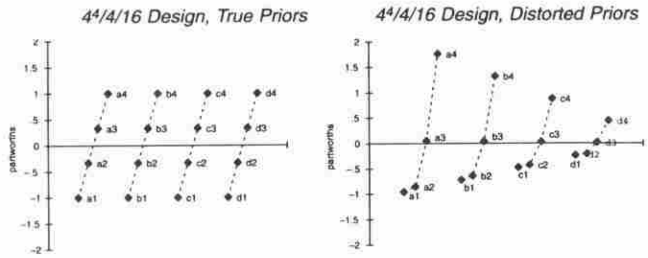
Notice that the efficiency gains depend significantly on the original utility-neutral design that is taken as the starting design. We took a design with average efficiency as the base—the expected efficiency with a random selection of one of the utility-neutral designs. Thus, not using priors subjects the analyst to both a lower expected efficiency and a substantial variation about that measure. Note further that these gains in efficiency occur despite minimal efforts on our part to find the most efficient design. We only sampled 1000 randomly relabeled designs, and swapping as operationalized is unlikely to find the absolute optimum. Indeed, we found that gains could have been even greater (though not more than 1–2%) if more thorough searches had been performed.

Scale of the coefficients is an important determinant of how much gain in efficiency is due to utility balance. In Table 2, we show that what was true for the $3^3/3/9$ family occurs generally. Multiplying all coefficients by 1.25 results in an increased efficiency gain of approximately 25%, whereas multiplying each by .75 reduces the efficiency gains by approximately 33%. In the extreme case, in which scale approaches zero, the utility-neutral design is best. Although scale has a strong impact on the extent of the efficiency gains possible, as will be made clear in the next section, being wrong about scale has relatively little impact on a person's ability to find an efficient design.

SENSITIVITY OF THESE RESULTS TO MISSPECIFICATION OF PRIOR BETAS

Because partworths can only be roughly estimated before the choice experiment, our goal here is to examine how the gains from utility balance change if the assumed priors are incorrect. We test the impact of three kinds of monotone misspecifications: skewness within attributes, relative weighting across attributes, and scale of the partworths. Misspecifications in skewness are generated by halving partworths less than zero and doubling those greater than zero (e.g., $-1, -.3, .3, 1 \Rightarrow -.5, -.15, .60, 2$); note that this changes the spacing, but not the ordering, of the levels. Weight misspecifications are represented by allowing the weights of the attributes to stretch uniformly between 1 and 4 (e.g., $1, 1, 1 \Rightarrow 1, 2.5, 4$). These distorted partworths are then normalized so that the sum of the squared β's is held constant, thus controlling for differences in the scale (Bunch, Louviere, and Anderson 1994). In Figure 1, we illustrate how the skewness and weight distortions greatly modify the partworths. Finally, the impact of misspecifications in scale are produced by multiplying the normalized partworths either by .5 or 1.5. To measure the impact of multiple misspecifications, we generated 12 designs for each design family that reflect all combinations of the three levels of scale and two levels each of attribute weighting and skewness.

Figure 1
COMPARISON BETWEEN TRUE PRIORS AND DISTORTED PRIORS BY WEIGHTS AND SKEWNESS



The loss in efficiency from these somewhat extreme misspecifications is gauged relative to the D_P-error of the best utility-balanced design with the true parameter vector β. We create the utility-balanced design, $\tilde{X}^*_\beta$, with respect to the misspecified β's, and evaluate this design under the true parameters. The ratio of the D_P-error for the best design, $D_P(\beta, X^*_\beta)$, divided by the D_P-error for the best design given we use distorted partworths, $D_P(\beta, \tilde{X}^*_\beta)$, is a measure of the relative efficiency of a design with misspecified prior coefficients.

We ran a descriptive ANOVA using this measure of relative efficiency as the dependent variable and using the five design families × three levels of scale × two levels of skewness and weighting and all two-way interactions as explanatory variables. The model accounted for variation extremely well; the resulting R^2 was 96%, and all main effects were highly significant ($p < .01$). Weight and skewness accounted for the largest variation, whereas design and scale were less important (percentage of explained variance = 41% and 21% versus 15% and 3%, respectively). The interactions were significant and accounted for 14% of variation. With the lack of importance of scale, a good way to display the results is to average over the three values of scale (.5, 1.0, 1.5). In the first column of Table 4, we indicate that misspecifying the scale produces designs approximately 95% as good as knowing the scale perfectly. Combining scale with skewness misspecifications results in design efficiencies of approximately 85%, whereas scale with weight misspecifications drops efficiencies to approximately 80%. Finally, combining all three types of misspecifications results in efficiencies of approximately 70% of the perfectly known priors condition. As bad as these losses may be, they are still substantially smaller than the efficiency losses from disregarding priors altogether, which is shown in the last column of Table 4. Thus, the analyst is better off being wrong about priors than disregarding them completely.

It is important to distinguish between the relatively minor impact of scale shown here and its major impact previously discussed and shown in Table 2. We showed that scale is important in determining the *potential gain in efficiency* from utility balance. However, in Table 4, we indicate that *identi-*

Table 4
RELATIVE EFFICIENCIES GIVEN MISSPECIFICATIONS OF PRIORS

	Level of Misspecification				
Design Family	*Scale Only*	*Scale & Skewness*	*Scale & Weight*	*Scale, Weight, & Skewness*	*Relative Efficiency of Utility-Neutral Designs ($\beta = 0$)*
$3^3/3/9$	95%[a]	89%	86%	83%	73%
$3^4/2/15$	97	86	82	70	57
$4^4/4/16$	97	90	85	81	64
$4 \times 3^3/3/48$ (interactions)	98	86	84	75	61
$9 \times 8 \times 4 \times 3^4 \times 2^3/3/63$	99	77	71	58	37

[a]Efficiencies are relative to utility-balanced designs using true priors. The numbers are averages across three scale levels.

fying a good design is not importantly dependent on scale. Put differently, making the wrong guess on scale means a person could be wrong about the extent of gain in efficiency, but this misspecification has relatively little impact on the relative efficiency of the optimal utility-balanced design.

DISCUSSION

We show that utility balance provides substantial improvement in the efficiency of choice designs. This efficiency gain arises because choices between similarly-valued alternatives provide better information about the coefficients. Thus, when priors are not zero, utility balance joins orthogonality, level balance, and minimal overlap as principles defining efficient choice designs. It is important to stress that though these four principles help us understand what makes a good choice design, in the region of the most efficient designs, they generally conflict with each other. Thus, it is necessary to use search routines coupled with design modification strategies to find designs that optimize D_P-error.

Prior estimates of the logistic coefficients are required to generate utility-balanced designs. There are several ways to generate a coherent and useful set of priors. Perhaps the best is through a small sample pretest that provides tentative logit coefficients. In this case, these coefficients can be entered directly into the estimate of D_P-error in Equation 4. Alternatively, the analyst could subjectively estimate probabilities for a provisional choice experiment and use the logit coefficients from that pseudoexperiment to provide priors. Finally, the analyst might allow experienced managers to guess the expected partworths for a study and use these directly as his or her estimate for β. The sensitivity analysis clearly shows that using even distorted estimates of β is better than assuming they are zero.

Once prior β's are assessed, it is straightforward, if tedious, to find an efficient choice design. The cyclic choice designs generated from orthogonal arrays provide a good starting point with minimal overlap, level balance, and orthogonality. From this starting design, zero-centered (D_0) and probability-centered (D_P) errors can be estimated. Relabeling attribute levels then allows for the exploration of designs that reduce D_P-error without altering D_0-error. Particularly for designs with large numbers of attributes or levels, relabeling alone can reduce the number of independent respondents needed to produce a specified error level by 10–40%.

Although swapping within choice sets normally increases overall efficiency, it needs to be more closely monitored. Swapping attribute levels can strongly reduce D_P-error, but it can increase colinearity and thereby worsen D_0-error. Thus, swapping is most appropriate when the analyst is confident in the prior *β's* used to estimate D_P-error. If they are inaccurate, then a less risky strategy might be to use relabeling only, which has no effect on D_0-error.

Our focus is to demonstrate the surprisingly strong impact of utility balance on the efficiency of choice designs. However, there are several questions that arise as a result of our explorations, which define an appropriate agenda for further research. Although we are confident that utility balance applies to models similar to those that were explored previously, more theoretical and empirical work is needed to generate designs that are appropriate for, say, a nested logit or a probit model with explicit heterogeneity terms. More work is also needed to apply utility balance to availability choice designs (Anderson and Wiley 1992; Kuhfeld, Tobias, and Garratt 1994; Lazari and Anderson 1994). These designs allow for the exploration of differential substitutability—the idea that the presence of one alternative in a choice set affects the utility of the others in the set. Modeling this common market behavior requires the use of universal logit (McFadden, Tye, and Train 1977). We do not consider such designs because they mainly require the use of a fixed alternative that is common to different choice sets. These fixed alternatives (e.g., "my current brand" or "I would keep shopping") facilitate the generation of estimable choice models but complicate the process of utility balance. There are interesting but as yet unanswered questions involved with a fixed alternative, particularly dealing with its inherent value: How popular should the fixed alternative be? How should designs be balanced if there are fixed alternatives? How much impact does the attractiveness of the fixed alternative have on the efficiency of the resultant design?

We explore ways to use prior information to generate more efficient designs. There are other, complementary ways that use prior knowledge to provide more accurate estimates. One approach is to use knowledge of functional forms to substitute linear or quadratic functions for partworth functions (Srinivasan, Jain, and Malholtra 1983): for example, replacing partworth price levels with one linear price term. Another approach is to use the priors directly in the estimation of the parameters within a Bayesian framework (Allenby, Arora, and Ginter 1995). Both of these strategies provide significant gains in design efficiency when applied to ratings-based conjoint analysis and are likely to increase the efficiency of choice experiments also, though that conjecture must be tested.

A final research issue concerns validating utility balance and other efficiency measures on consumers. The results presented here are relevant to ideal consumers, and it is important to assess the change in utility balance, for example, on the underlying preferences and error level of actual consumers. Although further research is always needed on the impact of human factors on choice designs, the idea that the task should mimic marketplace choices as closely as possible (Carson et al. 1994) makes utility-balance strategies appealing. To the extent that marketplace dynamics tend to force out competitors with low shares, market choice sets tend to be utility-balanced. Thus, utility balance may serve a dual role of increasing both the efficiency and the realism of a choice experiment.

Choice experiments have created excitement because of their ability to emulate consumer behavior. We show that balancing utilities within choice sets can increase the efficiency of choice designs; we provide a framework for generating such utility-balanced designs and a way to estimate their expected gains in efficiency. The utility balance developed here is not a tool in itself but should be seen as another way in which analysts can more intelligently explore consumer choices.

Appendix

DERIVATION OF PROBABILITY-CENTERED ERROR FOR A $3^3/3/9$ CHOICE DESIGN

Effects-Coded Design Matrix of Original $3^3/3/9$ Choice Design: X						Probabilities	*Probability-Centered Design Matrix: Z (from Equation 4b)*					
X1.1	X1.2	X2.1	X2.2	X3.1	X3.2	(from Equation 2)	Z1.1	Z1.2	Z2.1	Z2.2	Z3.1	Z3.2
1	0	1	0	1	0	.002	1.95	.90	1.95	.90	1.95	.90
0	1	0	1	0	1	.047	.95	1.90	.95	1.90	.95	1.90
–1	–1	–1	–1	–1	–1	.951	–.05	–.10	–.05	–.10	–.05	–.10
1	0	0	1	0	1	.045	1.00	–.86	.86	1.86	.86	1.86
0	1	–1	–1	–1	–1	.910	0.0	.14	–.14	–.14	–.14	–.14
–1	–1	1	0	1	0	.045	–1.00	–1.86	1.86	.86	1.86	.86
1	0	–1	–1	–1	–1	.488	1.00	.46	–.54	–1.00	–.54	–1.00
0	1	1	0	1	0	.024	0.0	1.46	1.46	0.0	1.46	0.0
–1	–1	0	1	0	1	.488	–1.00	–.54	.46	1.00	.46	1.00
0	1	1	0	–1	–1	.333	0.0	1.00	1.00	0.0	–1.00	–1.00
–1	–1	0	1	1	0	.333	–1.00	–1.00	0.0	1.00	1.00	0.0
1	0	–1	–1	0	1	.333	1.00	0.0	–1.00	–1.00	0.0	1.00
0	1	0	1	1	0	.045	.86	1.86	.86	1.86	1.00	–.86
–1	–1	–1	–1	0	1	.910	–.14	–.14	–.14	–.14	0.0	.14
1	0	1	0	–1	–1	.045	1.86	.86	1.86	.86	–1.00	–1.86
0	1	–1	–1	0	1	.488	.46	1.00	–1.00	–.54	.46	1.00
–1	–1	1	0	–1	–1	.488	–.54	–1.00	1.00	.46	–.54	–1.00
1	0	0	1	1	0	.024	1.46	0.0	0.0	1.46	1.46	0.0
–1	–1	1	0	0	1	.333	–1.00	–1.00	1.00	0.0	0.0	1.00
1	0	0	1	–1	–1	.333	1.00	0.0	0.0	1.00	–1.00	–1.00
0	1	–1	–1	1	0	.333	0.0	1.00	–1.00	–1.00	1.00	0.0
–1	–1	0	1	–1	–1	.910	–.14	–.14	0.0	.14	–.14	–.14
1	0	–1	–1	1	0	.045	1.86	.86	–1.00	–1.86	1.86	.86
0	1	1	0	0	1	.045	.86	1.86	1.00	–.86	.86	1.86
–1	–1	–1	–1	1	0	.488	–.54	–1.00	–.54	–1.00	1.00	.46
1	0	1	0	0	1	.024	1.46	0.0	1.46	0.0	0.0	1.46
0	1	0	1	–1	–1	.488	.46	1.00	.46	1.00	–1.00	–.54

$$\beta' = \begin{bmatrix} -1 & 0 & -1 & 0 & -1 & 0 \end{bmatrix}$$

Probability-Centered Covariance Matrix

$$\Omega_p^{-1} = (Z'PZ)^{-1}$$

(from Equation 4a)

	Z1.1	Z1.2	Z2.1	Z2.2	Z3.1	Z3.2
Z1.1	.67	–.30	.20	.02	.20	.02
Z1.2	–.30	.41	.02	–.02	.02	–.02
Z2.1	.20	.02	.67	–.30	.20	.02
Z2.2	.02	–.02	–.30	.41	.02	–.02
Z3.1	.20	.02	.20	.02	.67	–.30
Z3.2	.02	–.02	.02	–.02	–.30	.41

$D_P\text{-error} = (\det\Omega_p^{-1})^{1/6} = .38$

REFERENCES

Addelman, Sidney (1962), "Symmetrical and Asymmetrical Fractional Factorial Plans," *Technometrics*, 4 (February) 47–58.

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1995), "Incorporating Prior Knowledge Into the Design and Analysis of Conjoint Studies," *Journal of Marketing Research*, 17 (May), 152–62.

Anderson, Don A. and James B. Wiley (1992), "Efficient Choice Set Designs for Estimating Cross-Effects Models," *Marketing Letters*, 3 (October), 357–70.

Batsell, Randy R. and Jordan J. Louviere (1991), "Experimental Choice Analysis," *Marketing Letters*, 2, 199–214.

Ben-Akiva, Moshe and Steven Lerman (1985), *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

———, Takayuki Morikawa, and Fumiaki Shiroshi (1992), "Analysis of the Reliability of Preference Ranking Data," *Journal of Business Research*, 24 (March), 149–64.

Bunch, David S., Jordan J. Louviere, and Don Anderson (1994), "A Comparison of Experimental Design Strategies for Multinomial Logit Models: The Case of Generic Attributes" working paper, Graduate School of Management, University of California at Davis.

Carson, Richard T. et al. (1994), "Experimental Analysis of Choice," *Marketing Letters*, 5 (October), 351–67.

Chapman, Randall G. and Richard Staelin (1982), "Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model," *Journal of Marketing Research*, 19 (August), 288–301.

Huber, Joel, Dick R. Wittink, John A Fiedler, and Richard Miller (1993), "The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice," *Journal of Marketing Research*, 30 (February), 105–14.

Johnson, Richard M. (1987), "Adaptive Conjoint Analysis," in *Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis and Computer Interviewing*. Ketchum, ID: Sawtooth Software, 253–65.

Krieger, Abba B. and Paul E. Green (1991), "Designing Pareto Optimal Stimuli for Multiattribute Choice Experiments," *Marketing Letters*, 2 (November), 337–48.

Kuhfeld, Warren F., Randall D. Tobias, and Mark Garratt (1994), "Efficient Experimental Design With Marketing Research Applications," *Journal of Marketing Research*, 21 (November), 545–57.

Lazari, Andreas G. and Don A. Anderson (1994), "Designs of Discrete Choice Experiments for Estimating Both Attribute and Availability Cross Effects," *Journal of Marketing Research*, 31 (August), 375–83.

Louviere, Jordan J. and George Woodworth (1983), "Design and Analysis of Simulated Consumer Choice of Allocation Experiments: A Method Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350–67.

Madalla, G. S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zarembka, ed. New York: Academic Press, 105–42.

———, W. Tye, and K. Train (1977), "An Application of Diagnostic Tests for the Irrelevant Alternatives Property of the Multinomial Logit Model," *Transportation Research Record*, 637, 39–46.

SAS Institute (1995), *SAS/QC Software: Usage and Reference, Version 6*, Vol. 1, 1st ed. Cary, NC: SAS Institute.

Sawtooth Software (1993), *The CBC System for Choice-Based Conjoint Analysis*. Ketchum, ID: Sawtooth Software.

Srinivasan, V., Arun K. Jain, and Naresh K. Malholtra (1983), "Improving the Predictive Power of Conjoint Analysis by Constrained Parameter Estimation," *Journal of Marketing Research*, 20 (November) 433–38.

Swait, Joffre and Jordan J. Louviere (1993), "The Role of the Scale Parameter in the Estimation and Use of Generalized Extreme Utility Models," *Journal of Marketing Research*, 30 (August), 305–14.

Thurstone, L. L. (1927), "A Law of Comparative Judgment" *Psychological Review*, 34, 273–86.

Torgerson, Warren S. (1958), *Theory and Methods of Scaling*. New York: John Wiley & Sons.