Compensatory Choice Models of Noncompensatory Processes: The Effect of Varying Context

ERIC J. JOHNSON
ROBERT J. MEYER*

The sensitivity of the parameters and fit of compensatory choice models to contextual variations in information processing strategies is examined. A set of predictions is derived concerning specification errors which may arise when a compensatory model misrepresents a "true," noncompensatory choice process. These predictions are then tested in an experimental analysis of apartment choice behavior. Logit analysis and protocol analysis are employed to assess how the parameters and fit of a compensatory model vary in light of changes in the underlying pattern of information processing across choice sets of differing sizes. Although attribute usage and parameter variation across set sizes conformed to theoretical expectations, a hypothesized decrease in predictive accuracy was not supported.

An important part of understanding consumer behavior is the construction of formal representations of choice processes. Such formalisms as decision nets, conjoint analysis, and discrete choice analysis all attempt to model the relationship between characteristics of a product and observed choices.

Most applied forecasting methods used in marketing represent choice processes in terms of some form of algebraic model. A common assumption is that the parameters of these models are independent of the particular set of alternatives under study. Process analyses of choice, however, suggest that this assumption may not always be valid: changes in the number of alternatives and their attributes have been known to affect reports of the way choices are made.

Although protocols suggest that decision strategies are likely to be sensitive to context, the extent to which this will be reflected in the parameters of algebraic models has not been systematically investigated. The purpose of this paper is to explore this issue. Specifically, we examine the relationship between the parameters and fit of an algebraic model and changes in decision strategies across contexts as revealed through concurrent protocols.

*Eric J. Johnson is an Assistant Professor at the Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA 15213. Robert J. Meyer is an Assistant Professor at the Graduate School of Management, University of California, Los Angeles, CA 90024. The order of authorship is alphabetical; both contributed equally to this research. The authors thank Lee Cooper, Jordan Louviere, Michael Menasco, three anonymous reviewers for their helpful comments, and Matthew Saltzman for his programming assistance.

Our decision is organized in four sections, that (1) introduce the problem in detail; (2) derive a set of theoretical predictions concerning how changes in processing heuristics may affect the parameters and fit of an algebraic model; (3) report an empirical test; and (4) discuss the implications of our findings.

BACKGROUND

Underlying most applied work in consumer choice analysis is a simple theory of decision making: consumers are hypothesized to approach choice situations with a predefined algebraic judgment policy or utility function. This function defines how the observed attributes of products will be integrated to form overall evaluations of desirability. After independently evaluating each alternative, consumers are hypothesized to choose the option with the highest overall utility or value (e.g., McFadden 1981).

The intuitive appeal of the paradigm rests in its implications for forecasting. If an analyst is able to specify an individual's multiattribute judgment policy, it should be possible to predict evaluations of new or untested product options (e.g., Green, Carroll, and Goldberg 1981).

In recent years, the prospect of being able to make such forecasts has inspired a rapid growth in the development of improved methods for inferring multiattribute judgment policies from judgment and choice data (e.g., Green 1974; Hensher and Johnson 1981; Keeney and Raiffa 1976; Louviere and Woodworth 1983). Although current methods often vary widely in philosophy and structure, most share two common assumptions:

1. Judgment policies can be represented by a compensatory model.
2. Judgment policies are not contingent upon the evaluated set of alternatives.

The first assumption may not be not a particularly limiting one. Rephrased, it states that analysts usually restrict their attention to algebraic functions which are linear or multilinear in form—functions which allow tradeoffs between attributes (Hensher and Johnson 1981; Keeney and Raiffa 1976; McFadden 1981). Although it is likely that consumers may often use noncompensatory strategies—that is, eliminate alternatives without examining all attributes—such functions can usually be approximated by an interactive form of linear model (e.g., Billings and Marcus 1983; Einhorn 1970; Keeney and Raiffa 1976; Louviere 1979). Although a discussion of the reasons for the robustness of linear models is beyond the scope of this paper, it is sufficient to note that interactive or log-linear approximations of noncompensatory processes will usually provide a good description of response data which are not error-free (Einhorn 1970).

The assumption that judgment policies are not contingent upon the characteristics of the evaluated alternatives is more problematic. Although a linear or multilinear function may provide a good description of judgment or choice behavior in one context, the same function may not describe behavior in another. Given that such models are increasingly popular as tools for forecasting, this point appears critical: if individuals' judgment policies are influenced by the characteristics of the choice problem, then derived models may have limited transferability across contexts.

Research which has examined the effect of context on choice suggests that these concerns may be serious ones. The strongest evidence comes from analyses of decision protocols—studies in which individuals are asked to "think aloud" while engaged in choices made in differing contexts (e.g., Billings and Marcus 1983; Einhorn, Kleinmuntz, and Kleinmuntz 1979; Payne 1982; Svenson 1979). The view of decision making that emerges from this work is quite different from that which underlies most algebraic models. Instead of a single, context-free judgment rule, individuals appear to possess an assortment of contingent judgmental heuristics. The heuristics employed in choice depend upon the external representation of the choice problem, and vary over time as the structure of the choice set changes (e.g., Bettman 1979; Lichtenstein and Slovic 1971; Payne 1976, 1982).

As an example, when faced with simple binary choices, individuals often form preferences using dimensional comparisons (Russo and Dosher 1983). As the number of alternatives increases, there are changes in processing tactics: individuals make greater use of elimination strategies (e.g., Bettman and Jacoby 1976; Olshavsky 1979; Payne 1976; Wright and Barbour 1977). Because these two sets of processes are best described by differing forms of compensatory models, a model derived by observing choices made from sets of one size may do quite poorly when applied to predicting choices from sets of another.

Thus we appear to be faced with a paradox. On one hand, protocol analyses suggest that choices are likely to be made by a wide variety of strategies which are contingent upon characteristics of the choice alternatives; on the other, our technology for consumer forecasting tends to employ models which assume a single (multiattribute) process defined independently of context. For the applied researcher, the problem is a simple one: given that compensatory models approximate the trace of a more complex, contingent choice process, will changes in the structure of this process affect the parameters and fit of a compensatory model?

Empirical evidence on the robustness compensatory models is incomplete. It is well known, for example, that additive models can correlate well with judgment or rating data which are generated by a nonadditive source (Birnbaum 1973; Dawes and Corrigan 1974; Olshavsky and Acito 1980). The conditions for this appear to be rather general:

1. The predictor variable must have a monotonic relationship with the criterion.
2. There must be error in both sets of variables (Dawes and Corrigan 1974).

If these conditions are satisfied across contexts, compensatory models should provide a reasonably robust description of judgmental responses.

Evidence for the robustness of compensatory models for choice data—the criterion of interest in most applied work—is not as extensive. Although there is some evidence concerning the predictive validity of compensatory models when parameters are misspecified (Curry, Louviere, and Augustine 1981; Green, DeSarbo, and Kedia 1980), little is known about the sensitivity of such parameters to changes in underlying processing strategies.

Here we examine this issue. Our approach is both theoretical and empirical. We first derive a set of theoretical predictions concerning the types of specification errors that should arise when misrepresenting a noncompensatory (contingent) evaluation process in terms of a compensatory choice model. We then test these predictions by noting how the parameters and fit of a compensatory model vary under a manipulation of choice context known to induce noncompensatory processing: choice set size.

THEORETICAL ANALYSIS

In this section we explore the problem of specification errors that may arise when a compensatory choice model misrepresents the "true" pattern of information processing. We consider a "worst case" scenario of model misspecification: an analyst uses a linear compensatory model to predict choices generated by a noncompensatory source.

We first define a sequential elimination policy which represents the underlying choice process. The representation is a general one that subsumes a number of noncompensatory choice rules, most notably the conjunctive (e.g., Einhorn 1970). We then derive a set of predictions concerning the types of specification errors that should arise when attempting to represent data generated by this process in terms of a linear compensatory choice model.

The Stochastic Elimination Model

Basic Process. We characterize the individual's choice process as consisting of k discrete elimination stages, $k = 1, \ldots, m$. At each stage the decision maker defines an elimination policy which reduces the set of alternatives to smaller number of candidates. This iterative process continues until one candidate remains (a choice is defined).

Our formal model makes three assumptions:

A1. Let V_{ik} be the value or utility associated with alternative i at stage k, and let X_i be a vector of uniformly scaled attribute values associated with i. For each alternative i at each stage, V_{ik} is represented by a value mapping $\beta_k' X_i$ such that:

$$V_{ik} = \beta_k' X_i + \epsilon_i \tag{1}$$

where ϵ_i is an independent disturbance term, and β_k' is a parameter vector with elements that reflect the relative importance of each attribute in the evaluation process at stage k. Note that if the weight vector β_k' contains only one nonzero element for each stage (i.e., only one attribute is considered), the present model represents a conjunctive choice process.

A2. Associated with each alternative at each stage in the process is an independent, nonnegative probability of being retained as a candidate for choice. The probability that an alternative i will be a member of the candidate set (S_k) at stage k ($P(i \in S_k)$) is given by:

$$P(i \in S_k) = P(V_{ik} > T_k + \epsilon_{Tk}) \tag{2}$$

$$= P(\beta' X_i - T_k > \epsilon_{Tk} - \epsilon_i) \tag{3}$$

where T_k is a subjective value threshold and ϵ_{Tk} an associated random disturbance. This assumption states that on each stage, each alternative is evaluated by comparing its value at that stage to a stochastic threshold, $T_k + \epsilon_{Tk}$. The alternative is retained if its value exceeds this threshold.

A3. The random variables $\epsilon_{Tk} - \epsilon_{ik}$ for each alternative have independent, identically distributed (i.i.d.) logistic densities. Hence:

$$P(i \in S_k) = \frac{1}{1 + e^{-(\beta_k' X_i - T_k)}} \tag{4}$$

The final assumption provides a closed form for the probability in Equation 3. The i.i.d. restriction reaffirms the assumption that each candidate set decision is made independently for each alternative. Equation 4 implies that the probability of membership in the candidate set increases either with increases in overall value ($\beta_k' X_i$) or with decreases in threshold utility (T_k).

Note that we make formal assumptions neither about how threshold values (T_k) will be determined by a decision maker nor about the number of stages (m) required to yield a unique choice. A normative strategy would be to choose that set of thresholds which maximizes the overall utility of the chosen alternative while minimizing a cost function of m (the length of time spent making the choice; e.g., Grether and Wilde 1984). It is empirically more reasonable simply to assume that threshold values (and hence m) will be determined through trial and error; if one criterion eliminates too few or too many alternatives, the decision maker will "go back" and attempt some other.

The Formal Choice Model. Under the assumption of the independent probabilities of candidate set membership (Assumption 2), the probability that an alternative will be considered as a candidate for choice after m stages of elimination $P(i|k = 1, \ldots, m)$ is given by the product:

$$P(i|k = 1, \ldots, m) = \prod_{k=1}^{m} \frac{1}{1 + e^{-(\beta_k' X_i - T_k)}} \tag{5}$$

$$= \frac{e^{\sum_{k=1}^{m} \beta_k' X_i - T_k}}{\prod_{k=1}^{m} (1 + e^{\beta_k' X_i - T_k})} \tag{6}$$

The probability that alternative i will be chosen from a set of $j = 1, \ldots, N$ competitors ($P(i|j = 1, \ldots, N)$) is the normalized intersection of these candidate probabilities for all alternatives.[1] Formally:

$$P(i|j = 1, \ldots, N) = \frac{\dfrac{e^{\sum_{k=1}^{m} \beta_k' X_i - T_k}}{\prod_{k=1}^{m} (1 + e^{\beta_k' X_i - T_k})}}{\sum_{j=1}^{N} \dfrac{e^{\sum_{k=1}^{m} \beta_k' X_i - T_k}}{\prod_{k=1}^{m} (1 + e^{\beta_k' X_i - T_k})}} \tag{7}$$

Specification Errors in a Compensatory Approximation

We assume that an analyst observes repeated choices made from a set of $j = 1, \ldots, N$ alternatives generated by the stochastic elimination process summarized in Equation 7. We examine the types of specification errors which arise when attempting to represent these data in terms of the multinominal logit model:

---

[1]By using Equation 2 we are implicitly assuming that a choice of some alternative will be made—that is, the probability of the candidate set being a null set is zero.

$$P(i|j = 1, \ldots, N) = \frac{e^{\beta_o' X_i + \epsilon_{oi}}}{\sum_{j=1}^{N} e^{\beta_o' X_j + \epsilon_{oj}}} \tag{8}$$

where β_o' is a parameter vector, x_i is a vector of characteristics attributes of i, and ϵ_{oi} is an asymptotically normal error that the analysis assumes is independently, identically distributed for all i.

The multinomial logit is a form of fully compensatory (nonelimination) choice model which is used in a wide range of applied work (e.g., Urban and Hauser 1980). Although we limit our discussion of specification errors to the logit, the results should extend to any form of choice models assuming noncontingent utility functions, such as the multinomial probit (e.g., Currim 1982).

Our analysis centers on two issues:

1. How does the attribute salience vector (β_o') in the logit approximation relate to the saliency vectors (β_k') defined for each stage in the elimination model?
2. What are the properties of the specification error, if any, in the stochastic error ϵ_{oi} in the logit?

We derive two relevant results by equating Equations 7 and 8:

$$\beta_o \cong \frac{1}{m} \sum_{k=1}^{m} \beta_k' \tag{9}$$

$$\epsilon_{oi} \cong \frac{e^{\sum_{k=1}^{m} \beta_k' X_i - T_k}}{\prod_{k=1}^{m} (1 + e^{\beta_k' X_i - T_k})} - e^{\sum_{k=1}^{m} \beta_k' X_i} + \epsilon_{oi}^* \tag{10}$$

where ϵ_{oi}^* is an independent stochastic disturbance.

Equation 9 states that when the multinomial logit is used to represent a sequential elimination process, the revealed salience of attributes in the logit (β_o) will be approximately equal to their mean salience across the stages of elimination (β_k). Equation 10 defines the characteristics of the error in this approximation: it will always be nonpositive, increasing with both the number of elimination stages (m) and the overall utility of an alternative ($\beta_o' X_i$).

The result that ϵ_{oi} increases with m (the number of stages) is an intuitive one: it simply says that the more pervasive staged eliminations are in choice, the less appropriate the logit becomes as a representation of the choice process. Less obvious is the fact that the specification error will always be nonpositive and increase with the value of $\beta_o' X_i$. Basically, these results stem from the fact that the elimination model is a *satisficing* rather than *optimizing* model of choice. The probability of an alternative being chosen is the probability that it is "acceptable" on all criteria—not the probability that it is the "best available" (the criterion modeled by the multinomial logit). Formally, the numerator of the elimination model (Equation 7) is a zero-to-one bounded probability of *candidate set membership.* By contrast, the numerator of the multinomial logit is a zero-to-infinity measure of absolute utility. The more positive this absolute utility, the greater the (negative) difference between the predictions of the elimination model and the logit.

Implications for Cross-Task Transferability

Our theoretical analysis can be used to deduce the general conditions under which a compensatory model should fail in cross-task prediction. Our approach begins with the assumption that the attribute salience vectors β_k in the elimination model can be viewed as a set of measures of the extent to which each attribute is considered at each stage in choice. Thus Expression 9 can be restated as implying that the more frequently an attribute is attended to during choice, the higher its salience should be as defined in a logit model. Hence, if changes in task induce changes in the relative frequency with which attributes are considered, there should be a corresponding change in the parameters of the best fitting compensatory model.

To illustrate, when given a choice between two alternatives described by three attributes, an individual might compare them once on each dimension, and then choose the one which is best on at least two of the three (Russo and Dosher 1983). If the choice set were to be expanded—say to eight alternatives—a different strategy would probably be employed: the consumer might first use one of the attributes to screen the choice set, reducing it to a manageable size, and then carefully examine the remaining candidates on all three attributes (e.g., Wright and Barbour 1977). Because the attribute which was used for screening would have been referred to more frequently in the second scenario then the first, we would predict that its revealed salience will be greater in the second.

EMPIRICAL ANALYSIS

In this section we empirically examine the consequences of representing contingent choice processes with compensatory models. We focus on two implications of our theoretical analysis:

1. The salience of attributes revealed in a compensatory model should reflect the frequency of their use in an elimination process.
2. The overall fit of a compensatory model should decrease as eliminations become more commonplace.

We tested these predictions by monitoring individuals' choices from sets of varying sizes. This manipulation was designed to induce changes in the way individuals processed information across a set of contexts. Individuals were hypothesized to use compensatory strategies when evaluating one or two alternatives, but to make more extensive use of sequential elimination strategies when faced with larger choice sets.

We summarize these objectives in four hypotheses:

H1: As set sizes increase, sequential elimination strategies in choice should increase.

H2: As set sizes increase, the pattern of attribute usage will change. Specifically, some "initial filtering" attributes will be used relatively more frequently and other "final comparison" attributes will be used relatively less frequently.

H3: As set sizes increase, the distribution of revealed attribute weights in a compensatory (logit) model will follow the pattern of attribute use. Specifically, more important "filtering" attributes will increase in revealed importance and less important "comparison" attributes will decrease in importance.

H4: As set sizes increase, the overall fit of the compensatory model should decrease.

Two methodologies were employed to test these hypotheses: protocol analysis and logit analysis. Protocol analysis was used to examine Hypotheses 1 and 2, the necessary conditions for a test of specification errors due to a mismatch of model to process. Logit analysis was then employed to test Hypotheses 3 and 4, the hypothesized effects of this mismatch on the parameters and fit of a compensatory (logit) model.

Experimental Design

The context of our experiment was students' selections of apartments. Subjects were presented with profiles of hypothetical apartments, each described in terms of four attributes: monthly rent (a dollar figure), walking time to campus (in minutes), general appearance of the unit and its associated neighborhood (a brief verbal description), and its maintenance quality (a brief verbal description). Subjects were asked to make choices and judgments in choice sets of either 1, 2, 4, or 8 alternatives. For single-alternative choice sets, subjects provided a preference rating on a 10-point scale of overall desirability. For larger sets, subjects chose their most preferred apartment.

For each of the four levels of set size, we constructed eight different choice sets using a two-stage procedure:

1. An unlabeled "target" apartment in each set was constructed by assigning "high" and "low" attribute levels according to a fractional experimental design.
2. Remaining "competing" apartments were randomly assigned attribute levels drawn from a uniform distribution of mid-range values.

To ensure that the target alternative would not be transparent in the experiment, "high" and "low" attribute levels were drawn randomly from bounded (uniform) distributions around each level. Hence, although in the aggregate the characteristics of a given target were constant throughout the experiment, its exact description varied slightly from choice set to choice set for each subject.[2]

Examples of the experimental stimuli for choice sets of size 1 and 4 form Figure A. A more complete description of the experimental instructions and stimuli is available from the authors upon request.

The targets represented cells in one-half fraction of a 2^4 factorial. In this design, four apartment attributes were varied at either a "high" or "low" level. The design permitted independent estimation of all four attribute main effects, as well as three two-way interactions: rent with distance, maintenance, and appearance.[3]

Competing alternatives were assigned attribute levels between the high and low levels to control for some of the features of choice sets that would otherwise be confounded with set size. In larger set sizes, there will be a higher likelihood that dominated alternatives will exist and will be excluded from the set before the subject makes serious evaluations. The distributional characteristics (mean and variance) of the competitive set may also vary.

By using mid-range levels, we controlled for both of these effects. Specifically:

1. The relative incidence of dominance of the target was constant in each set-size condition: the target was dominant in one choice set, was dominated in one, and was in the pareto-optimal or feasible set in the remainder.[4]
2. The mean and variance of attribute levels was constant across set-size conditions.

The analysis examined changes in choice behavior across set sizes. In the protocol analysis, we inferred choice processes using subjects' self-reports. In the logit analysis, choice processes were examined by estimating the statistical effects of "rent," "distance," "maintenance," and "appearance" on observed choice frequencies associated with the target between conditions.

Protocol Analysis

This analysis tested the two hypotheses that should exist in order to demonstrate bias due to model misspecification—i.e., Hypothesis 1, which states that eliminations will become more frequent as set size increases, and Hypothesis 2, which suggests that screening attributes will be used relatively more frequently as set size increases.

[2]A copy of the PASCAL source code used in constructing the experimental scenarios is available from the authors.

[3]In this design, all unestimated two-and-higher-level interactions were assumed to be negligible. Hence, while all main effects and estimated two-way interactions were mutually independent, they would potentially be confounded by significant higher-order effects (Hahn and Shapiro, 1966).

[4]The dominant and dominated targets were the "low rent, low distance, high maintenance, and high appearance" and the "high rent, high distance, low maintenance, and low appearance" cells of the experimental design, respectively.

FIGURE A

EXEMPLARY APARTMENT SCENARIOS FOR JUDGMENT (N = 1) AND CHOICE (N = 4) CONDITIONS[a]

Judgment condition

Consider the following apartment:

Apartment 1: Rent: 315 dollars
Distance: 20 minutes to campus
Maintenance quality & availability: MINIMAL; Maintenance person difficult to contact.
Appearance: Limited-sized apt. in older residential area, no a/c, some conveniences.

How would you rate this apartment in terms of the best and worst you can imagine?
worst 1 2 3 4 5 6 7 8 9 10 best

Choice condition

Consider the following 4 apartments:

Apartment 1: Rent: 282 dollars
Distance: 27 minutes to campus
Maintenance quality & availability: FAIR; Maintenance person can be contacted most regular business hours.
Appearance: Reasonable-sized, clean apt. in quiet neighborhood, equipped, ww/c.

Apartment 2: Rent: 240 dollars
Distance: 23 minutes to campus
Maintenance quality & availability: GOOD; Maintenance person usually available.
Appearance: Reasonable-sized, clean apt. in quiet neighborhood, equipped, ww/c.

Apartment 3: Rent: 225 dollars
Distance: 22 minutes to campus
Maintenance quality & availability: FAIR; Maintenance person can be contacted most regular business hours.
Appearance: Moderate-sized, tidy apt. in safe neighborhood, equipped, ww/c.

Apartment 4: Rent 326 dollars
Distance: 11 minutes to campus
Maintenance quality & availability: MINIMAL; Message must be left with answering service.
Appearance: Spacious, tidy apt. in good location, fully equipped, ww/c, a/c.

Which one would you choose? ______

[a]In the choice condition, the target is Apartment 4.

Subjects. Because protocol analysis is a highly informative but labor intensive procedure, we examined the behavior of a small number of subjects. Nine subjects participated in response to signs posted on campus. Subjects were paid $4.00 to participate in an hour-long experimental session, which were run individually over a one-week period.

Procedure. Subjects were presented with three replications of the design described above. They generated concurrent verbal protocols (''thought aloud'') during the first four choices or judgments in each set size. Protocols were collected in one of the three replications. Order was counterbalanced: three subjects generated verbal reports in each replication. This was preceded by a brief warm-up task familiarizing subjects with verbal reports.

Analysis. The protocols were transcribed and segmented into complete thoughts by a research assistant who was unaware of the experimental hypotheses (Newell and Simon 1972). Two raters, also unaware of the hypotheses, coded each statement into one of six categories:

- *Read:* Verbatim reading of information from the choice scenario.
- *Evaluation:* An evaluation of an attribute or alternative, without reference to a standard or to another attribute or alternative.
- *Comparison:* A judgment of the relative value of an attribute or alternative relative to another.
- *Eliminations:* Statements that an alternative will be eliminated from consideration.
- *Choice and rating statements:* Statements announcing either the overall choice of an alternative or the rating assigned to an alternative.
- *Strategy statements:* Statements describing aspects of the decision process unrelated to the current choice set, such as the importance of an attribute or a description of a method for making a choice.

The coding was done using a computer program that presented a coder with the statement in question and asked a series of questions which resulted in the assignment of the statement to one of these six categories, or classified the statement as uncodable. Only 0.7 percent of the statements were classified as uncodable by both judges. Although judges initially agreed upon 79 percent of the codings, a clarification of the coding scheme on several points increased independent agreement to 96.5 percent, and the remaining conflicts were resolved by discussion. All six categories showed inter-rater agreement exceeding 90 percent. The raters also coded, where possible, the identity of the alternatives and attributes mentioned in the statement. These codings also showed high inter-rater agreement (95.8 and 95.0 percent, respectively).

Analyses Related to Hypothesis 1: Changes in Processing Strategies. We first examined the relationship between choice set size and the use of elimination strategies. Three different analyses examined the impact of increased set size. All three analyses would suggest the use of noncompensatory rules (such as the conjunctive and elimination-by-aspects):

1. *An analysis of the decision processes described by the protocols:* As set size increased, we expected an increase in the proportion of statements which eliminated alternatives, an increase in comparisons, and a relative decrease in the number of evaluation statements.
2. *An analysis of information search as described in the protocol:* As set size increased, we expected an increase

in the proportion of statements examining the alternative eventually chosen.

3. *An analysis of the pattern of information search:* As set size increased, we expected more attribute-based search.

To explore the decision process portrayed in the protocols, we conducted an analysis of variance upon the proportion of statements assigned to each category. This analysis used the position of the verbal report (first, second, third replication) as a between-subjects factor, and choice set size (1, 2, 4, or 8) as a within-subjects factor. Table 1 presents the relevant results, the proportion of each statement assigned into each coding category tabulated by set size. The columns on the right present significance tests of the effect of choice set size in each analysis of variance. A multivariate analysis of variance confirmed the significance of the effect of set size for each statement type.[5]

Increases in set size had marked effects on the composition of the protocols. Specifically, as the number of alternatives increased, the protocols contained relatively more statements reading information, less evaluation of that information, and somewhat more eliminations. Not surprisingly, comparisons were common in choice ($N > 2$) but not in judgment ($N = 1$). The increased frequency of eliminations in larger set sizes was consistent with the hypothesis that increases in set size should be associated with increases in the use of noncompensatory strategies.

Additional evidence was provided by an analysis of the proportion of statements mentioning the chosen brand. Johnson and Russo (1984) have suggested that noncompensatory decision strategies will result in an unbalanced search of the brand × attribute matrix. Specifically, noncompensatory strategies eliminate inferior alternatives from consideration, resulting in increased search of the brand eventually chosen. If noncompensatory processes become more common with larger set sizes, then we expect search to be relatively more concentrated for larger set sizes.

We calculated an index of the concentration of search, $P = (C - (1/n))$, which represents the proportion of statements mentioning search of the chosen alternative (C) in comparison to that predicted by equal search of all alternatives ($1/n$). A positive value of this index would indicate that search of the chosen alternative exceeded that predicted by chance. The value of this index increased with set size: for 2 alternatives $P = 0.03$, while for 4 and 8 alternatives, the index was 0.05 and 0.13, respectively. An analysis of variance, similar to the one above, was conducted on the proportion of statements referring to the chosen alternative for set sizes of 2, 4, and 8. A linear post-hoc (Scheffe) contrast, which compared the value of P across conditions, was significant at $F\ (2, 133) = 28.9\ (P \leq 0.001)$. Thus, the increase in the concentration of search also supported the hypothesis that the use of noncompensatory strategies increases with set size.

TABLE 1

PROPORTION OF STATEMENTS OF EACH TYPE OF SET SIZE

Statement type	Set size 1	2	4	8	F(3, 121)	P
Read	.48	.56	.62	.63	4.16	.01
Evaluations	.35	.13	.10	.09	29.99	.001
Comparisons	.03	.21	.19	.19	13.36	.001
Eliminations	.00	.00	.02	.05	3.90	.02
Strategy	.01	.01	.01	.02	3.38	.03
Rating	.12	.08	.05	.04	28.12	.001

Finally, an examination of the transitions in the protocol was also consistent with the hypothesized shift to noncompensatory processes in larger set sizes. Using only "read" statements (the closest indicator of the brand or attribute information that was being processed by the subject in a given scenario), we calculated the indices of brand and attribute processing suggested by Bettman and Jacoby (1976):

$$\text{Normalized proportion transitions} = \frac{(\text{Number of same brand transitions})/M}{(\text{Total number of transitions})}$$

where M equals:

$$\frac{(\text{Number of total transitions} - \text{Number of attributes searched})}{(\text{Number of total transitions})}$$

or the maximum number of brand transitions possible.

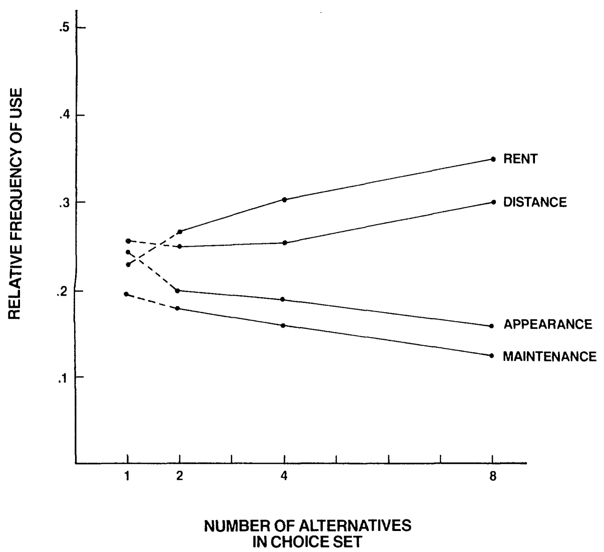
The equivalent measure was developed for attribute transitions. An analysis of variance on these dependent measures confirmed that attribute-based processes were more common in larger set sizes. The mean percentages of attribute transitions were 0.26, 0.36, and 0.40 for 2, 4, and 8 alternatives, respectively, while the proportions of brand transitions were 0.73, 0.61, and 0.58. Linear post-hoc contrasts confirmed the significance of these trends with $F\ (2, 116) = 7.74$ and 21.23, $p \leq 0.001$ for both. Although one compensatory rule—additive differences—could have produced a high proportion of attribute transitions, an examination of the protocols showed that use of this rule was rare. Moreover, additive differences would not produce the observed differences in the content of the protocols or an increased concentration of search as set size increases.

In sum, evidence from the content of the protocols, the proportion of search devoted to the chosen alternative, and the amount of search by brand and by attribute provided evidence consistent with the hypothesis that increases in set size were accompanied by increased use of noncompensatory processes.

[5]The MANOVA was conducted using five of the six categories, since the remaining category is a linear combination of the remaining five. The analysis was also conducted upon the proportions after an arcsin transformation, with no substantial change in the results.

FIGURE B

RELATIVE ATTRIBUTE USAGE AS A FUNCTION OF SET SIZE



Analyses Related to Hypothesis 2: Changes in Attribute Use. Our second major analysis focused on the hypothesis that the frequency with which major "screening" attributes are referred to during the choice process should increase with increases in set size. In Figure B, we plot the proportion of statements referring to each of the four attributes for each set size. As this figure demonstrates, the proportion of references to rent increased as the set size increased, while maintenance and appearance were mentioned proportionately less often. Distance also showed a proportional increase in frequency of mentions in the largest size choice set. The effect of set size was confirmed by an analysis of variance conducted on the individual statements. This analysis treated attributes and set size as within-subject factors. Overall, the interaction between set size and attribute use was significant (F (3, 2103) = 5.67 $p \leq 0.001$), as was the post-hoc linear contrast for all four categories ($p \leq 0.001$).

We might also point out that while set size induced changes in the relative frequency with which attributes were mentioned, it did not appear to induce changes in the order of referrals. For set sizes of two or more, the rank order of attribute usage remained constant despite changes in the raw relative frequencies (see Figure B).

Summary. The results of the protocol analyses suggested that the desired necessary conditions for testing Hypotheses 3 and 4 existed in the experiment. In particular, when confronted with larger set sizes, subjects tended to make more frequent use of noncompensatory strategies and to alter the relative frequency with which they referred to attributes. Hypotheses 3 and 4 are theoretical predictions concerning how these changes should impact upon the parameters and fit of a compensatory (logit) model.

Logit Analysis

The logit analysis examined Hypotheses 3 and 4—the effect of choice set size on parameter estimates of the compensatory model parameters, and overall fit. Although it would have been desirable to test these hypotheses at the individual level, the sparse nature of discrete choice data would not permit convergent parameter solutions of individual models.[6] Hence it was necessary to test the hypotheses at the group level using an expanded subject pool. The possible ambiguities induced by conducting our tests at the group level are discussed as we consider each hypothesis in turn.

Subjects. Ninety-one subjects participated in response to signs posted on campus, and were again paid $4.00 to participate in an hour-long experimental session. Subjects were run in groups that ranged in size from three to 20, over a one-week period.

Procedure. The experimental procedure basically mirrored that used in the individual-level analyses, with three modifications:

1. Subjects completed only two replications of the design.
2. The order of set-size presentation was randomized between subjects.
3. Verbal protocols were not elicited.

Analyses Related to Hypothesis 3: The Effect of Set Size on Revealed Measures of Attribute Importance. We wished to test the hypothesis that increases in the size of a choice set would cause the distribution of attribute salience parameters in a compensatory choice model to become increasingly skewed or polarized between "filtering" and "final comparison" attributes. An implicit assumption when testing this hypothesis at the group level is that the subject pool is relatively homogeneous in terms of their decision-making processes. While each subject might use staged processes for larger choice sets, if individuals use different attributes for "filtering" the set, upon aggregation we may observe no change in the mean salience of an attribute across levels of set size. Hence, while the parameter-change hypothesis may be true at the individual level, it may well disappear when studying the behavior of a highly heterogeneous aggregate. Violations of our assumption result in a conservative test.

With this caveat in mind, we examined changes in the statistical effect of attributes across set sizes. To ensure compatibility between desirability ratings collected for sizes

[6] A traditional limitation of discrete choice analyses is that model estimation invariably requires aggregation across a subject pool. Some researchers have attempted to avoid this problem by having subjects "allocate" several choices among alternatives (e.g., Batsell 1980), but there is empirical evidence that subjects' behavior in allocation tasks is not equivalent to that in analogous discrete choice tasks (e.g., Meyer and Eagle, 1982).

of one and discrete choices (collected for set sizes of two or more), ratings were converted to a binary scale using the mean rating response as a cut off; specifically:[7]

$$\text{Pr(Target)} = \begin{cases} 1 \text{ if Rating} \geq \text{overall mean rating} \\ 0 \text{ otherwise} \end{cases}$$

Hence, the dependent variable in all cases was a binary indicator of preference for the target alternative. Two analyses were performed on the data:

1. *A graphical analysis* of changes in the impact of the attributes (a measure of attribute importance) across the various set size conditions.
2. *A statistical analysis* of these changes using a discrete binary logit model of the individual choice data.

Graphical Analysis. To provide an initial, visual examination of the parameter-change hypothesis, choice proportions associated with the target (experimental) alternative were computed for each cell of the design (1982 discrete observations per cell). Because a logit model (Equation 8) was assumed to underlie these data, proportions were subjected to a logit transformation of the form ln(Pr(Target/ 1 − Pr(Target)). Differences in marginal means between the two levels of each attribute were then computed to obtain measures of the relative salience of each attribute for each level of set size (e.g., Anderson 1982; Green 1974).

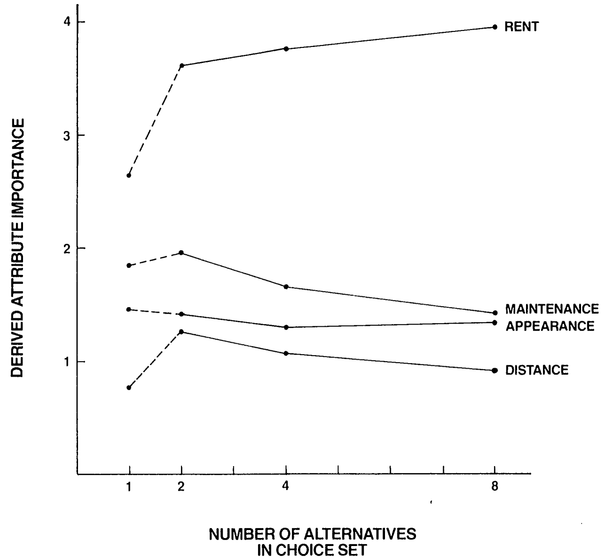
A plot of the derived measures of attribute importance with set size provided initial support for the research hypothesis (Figure C). In particular, as hypothesized, the distribution of experimental effects becomes increasingly skewed with increases in the size of the choice set. Over the three "choice" conditions (set size of 2, 4, or 8), the importance of "rent" monotonically increased, while those of "distance" and "maintenance" monotonically decreased. The only exception was "appearance," which did not demonstrate a monotonic change in salience.

We can also compare the plot of changes in revealed attribute salience (Figure C) and the plot of changes in attribute use reported by subjects in the protocol task (Figure B). Although these results are based on the responses of two different groups of subjects, in both cases the distribution of attribute saliencies and use became more skewed with increases in set size, with rent increasing in importance and maintenance and appearance decreasing in importance. The exception was distance, which decreased in revealed salience with increases in set size in the choice task, while being referred to more frequently by subjects in the protocol task.

Statistical Analysis. To provide a statistical test of the above results, the discrete choices were subjected to a binary logit analysis. The model we estimated centered on the probability of subjects choosing the target alternative (t) in a given choice set j. Letting $P(t|j)$ be this probability, we estimated a model of the following general form:

FIGURE C

ATTRIBUTE IMPORTANCE AS A FUNCTION OF SET SIZE[a]



[a]Attribute importance is defined as the absolute difference in logit-transformed marginal means (choice proportions) for each attribute.

$$P(t|j) = \frac{1}{(F(X_t|j))} \tag{11}$$

The function $F(X_t|j)$ was a linear combination of the following 63 terms:

1. The value of the target on each of the four attributes, expressed as a deviation from the choice set mean (not including the target).[8]
2. Three estimatable two-way attribute interactions (each with "rent").
3. Three variables reflecting linear, quadratic, and cubic trends in "choice set size."
4. 21 two- and three-way interactions between the "set size" trends and each attribute main effect and two-way interaction.
5. 32 terms reflecting the effect of "replication" (a main effect and 31 interactions).

Choice set size trends were coded as orthogonal polynomials, and each attribute was expressed as a deviation from the choice set mean to minimize that error variance which might have been associated with random differences

[7]To investigate the effects of this transformation on parameter estimates and model fit, the judgmental data were analyzed using both the original ratings data (transformed to a probability scale) and the 1–0 binary scale. The results of both analyses were comparable. In particular, both yielded nearly identical fits (pseudo R^2's of 0.269 in both cases) and, correspondingly, similar parameter estimates.

[8]For choice sets of size one (the judgment condition), the choice set mean was uniformly set equal to the grand choice set mean.

in the way each treatment combination was presented. Because each attribute main effect was represented as a continuous variable, each difference vector was centered about its respective mean to maximize orthogonality with each two-way attribute interaction vector (Robson 1959).

The 64 parameters in Equation 11 were estimated via maximum likelihood, using the software package CRAWTRAN (Avery 1982). The results can be summarized as follows (see Table 2):

1. All four attribute main effects were highly significant, as was the rent × distance interaction. This interaction implied that the underlying "average" judgment rule had a nonadditive—or perhaps a noncompensatory—component. As suggested from the graphical analyses, rent emerged as the most salient predictor.
2. There was a significant decrease in the relative frequency of choices of the target as set size increased. As one would expect, the larger the set size, the lower the mean probability of choosing the target.
3. There was no systematic effect due to "replication," implying that subjects did not vary decision strategies across replications in the experiment.
4. Set size systematically interacted with only one apartment attribute: rent.

The significant set size × rent interaction lends statistical support to our earlier observations that the revealed importance of rent increased with increases in the size of the choice set. Moreover, associated with this change were significant linear and quadratic trends. The directions of the trend coefficients implied that:

1. As set size increased, the slope or effect of rent became increasingly negative (the linear set size × rent trend was negative).
2. As set size increased, this increase decelerated (the quadratic set size × rent trend was positive).

The fact that none of the other set size × attribute interactions were significant implied that the importance of the other three apartment attributes was relatively constant across the four set sizes. Although the graphic analysis suggested a monotonic decrease over the ($N = 2, 8$) interval for distance and maintenance, the overall effect was not statistically significant.

In sum, the data lent support to the basic hypothesis that increases in set size affects the distribution of derived attribute effects. Similar to the protocol results, attribute weights became polarized, with rent increasing in importance and other attributes decreasing in importance with larger set sizes.

Results Related to Hypothesis 4: Changes in Predictive Ability Across Set Sizes. This analysis focuses on the robustness of the compensatory choice model in prediction. The analysis centered on two related issues:

1. The fit of a compensatory model as set size increased (Hypothesis 4).
2. The predictive ability of such models across tasks—that is, the ability of a compensatory model derived in one context (e.g., judgment) to predict behavior in another (e.g., choice).

We addressed both these issues in a single analysis.

Group-level binary logit models were estimated for each of the four choice set-size conditions of the experiment, and were then used to predict observed choices in all four set sizes.[9] To provide a basis of comparison, the predictive

TABLE 2

LOGIT ESTIMATION RESULTS

	Coefficient	T-statistic	Pr(T)
Attribute main effects			
Rent	−.028	−33.833	.0000
Distance	−.079	−13.217	.0000
Maintenance	.530	16.29	.0000
Appearance	.433	13.49	.0000
Set size effects			
Linear trend	−.245	−10.215	.0000
Quadratic trend	.340	−6.519	.0000
Cubic trend	.154	6.834	.0000
Interactions with rent (R)			
R × Distance	−.000	−.498	.6181
R × Maintenance	−.002	−3.736	.0186
R × Appearance	−.001	−1.962	.0497
Interactions with set size (N)			
a. Linear trends (NL)			
NL × Rent (R)	−.002	−4.611	.0004
NL × Distance (D)	−.003	−1.012	.3114
NL × Maintenance (M)	−.016	−1.034	.3011
NL × Appearance (A)	−.003	−.190	.8493
NL × R × D	.000	1.007	.3138
NL × R × M	−.000	−1.488	.1366
NL × R × A	−.000	−.925	.3536
b. Quadratic trends (NQ)			
NQ × Rent (R)	.002	2.746	.0060
NQ × Distance (D)	.010	1.673	.0944
NQ × Maintenance (M)	−.036	−1.095	.2736
NQ × Appearance (A)	.046	1.455	.1456
NQ × R × D	.000	2.214	.0268
NQ × R × M	.000	.635	.5253
NQ × R × A	.001	1.281	.2001
c. Cubic trends (NC)			
NC × Rent (R)	−.001	−2.171	.0299
NC × Distance (D)	−.003	−1.236	.2164
NC × Maintenance	.006	.399	.6898
NC × Appearance (A)	.015	1.122	.2617
NC × R × D	.000	.097	.9228
NC × R × M	−.000	−1.399	.1616
NC × R × A	−.000	−.921	.3569

NOTE: Fit indices: Residual R-squared: .462; Log-likelihood R-squared: .405; Chi-squared test of model (54 df): 3230.00.

[9]Our method of cross-prediction was to regress the predicted choice frequency of the target under a model derived in one set size against the observed choices in another. This linear transformation provided an optimal adjustment of predictions for differences in mean choice frequencies between conditions (model intercepts).

ability of a naïve, equal-weight model was also examined. Predictive ability was assessed using three indices of fit:

1. The "predictive efficiency" (pseudo R^2 or ρ^2) of each model.
2. The proportion of correct predictions of the target alternatives.
3. The root mean squared error of predicted choices.

Changes in Predictive Ability across Set Sizes. Results relating to the prediction analysis are summarized in Table 3. The research hypothesis that overall fit should decrease with increases in set size was tested by noting:

1. The predictive ability of the model derived for each set-size condition (the diagonal fit indices in Table 3).
2. The "average predictability" of choices in each condition, which was derived by averaging fit indices across models (the column means in Table 3).

Neither measure supported the hypothesis. Rather than predictive abilities decreasing with increases in set size, we observed, if anything, an increase. For example, in terms of proportion of correct predictions and root mean square error, choices made from sets of size eight displayed the highest predictability. In the task where we most expected subjects to employ noncompensatory decision strategies—the choice among eight alternatives—the compensatory model actually displayed its highest predictive ability.

Mirroring this result, note that transformed choices observed in the judgment condition (which our theory suggests will be best predicted by a compensatory model) were the least predictable. While the overall stability of fit indices within the choice tasks may be somewhat perplexing, the decreases in the judgment conditions may have a rather straightforward explanation: responses in the judgment conditions were fundamentally more variable due to between-subject differences in rating scale usage. Thus the low predictive ability of judgment vis-à-vis choices may reflect differences in the total between-subject error in the experimental conditions.

To provide a statistical test of fit variation, the root-squared error for each observation was subjected to an analysis of variance. This analysis contained two factors, paralleling those displayed in Table 3:

1. The set size of the data used for estimation.
2. The set size of the data being predicted.

This analysis tested the significance of the changes in predictive validity shown in Table 3.

The two main effects were significant at $p < 0.0001$, suggesting that:

1. The various choice set sizes significantly varied in their predictability.
2. Some models provided better levels of prediction than others across set sizes.

Finally, a significant interaction between the two ($p < 0.001$) indicated there was significant variation in the ability of the models to predict choices across tasks. Hence, the analysis suggested that the variation in predictive validity shown in Table 3, although small in absolute terms and counter to the hypothesized direction, was greater than that expected by chance.

Cross-Task Robustness. One of the most impressive aspects of the results was the lack of differences in predictive ability across contexts. Although the cross-task variation in predictive ability of the derived models was statistically significant, the absolute numerical difference was small. Differences in "success rates" for predictions by the models across choice contexts ranged from 6 percent for the judgment data to 2.4 percent for choice sets of size two. Hence, if one were willing to accept 2.4 percent difference in success rates as an acceptable error, one could draw the reasonable conclusion that the models were robust to estimation context.

We should add that even the naïve (equal-weight) model performed reasonably well in cross-task prediction. For example, the average success rate in predictions of the target using the equal-weight model (where 50 percent was chance) was 74.5 percent, compared to an average of 80.1 percent for the best differential weight model.

DISCUSSION

In recent years, the literature in consumer choice has been strongly influenced by two different paradigms: algebraic modeling and process tracing. Algebraic modeling has tended to dominate applied work in disaggregate consumer demand forecasting, while process tracing has dominated research focusing on the cognitive processes underlying choice. Their differences, however, have been more than either their fields of application or methodologies. Each appears to be rooted in fundamentally differing views of how we make choices. The paradigm of algebraic modeling is based on the assumption that it is possible to recover "context-free" judgment policies. Research in process tracing, however, suggests that such policies may not exist: choice strategies often appear to be contingent upon the set of alternatives being evaluated.

The purpose of this research was to explore some of the implications of this apparent paradox. In particular, we examined the impact that variations in processing strategies across contexts might have on the parameters and fit of an algebraic (compensatory logit) model of choice.

We began with a theoretical analysis of the types of specification errors which may arise when representing a noncompensatory choice strategy with a compensatory choice model. This analysis suggested that:

1. The relative revealed importance of an attribute should be related to the relative frequency of its use during a sequential choice process.
2. While the compensatory model may correlate well with data generated by a noncompensatory model, the relative magnitude of this correlation should decrease with increases in the extent of underlying contingent processing.

TABLE 3

FIT INDICES

Set size used for estimation	Fit index	Set size used for prediction 1	2	4	8	Row means (average model fit)
1	ρ^2	.269	.410	.400	.470	.387
	% correct[a]	72.8	81.3	80.2	81.7	79.0
	MSE	.341	.306	.306	.316	.318
2	ρ^2	.254	.438	.425	.488	.401
	% correct	73.1	83.2	81.1	81.5	79.7
	MSE	.326	.242	.248	.330	.299
4	ρ^2	.248	.436	.425	.486	.398
	% correct	73.1	83.7	81.3	81.0	79.8
	MSE	.316	.285	.252	.249	.276
8	ρ^2	.251	.432	.420	.490	.398
	% correct	73.1	83.6	81.3	82.4	80.1
	MSE	.317	.308	.246	.228	.275
Column means (average set-size fit)	ρ^2	.256	.429	.417	.484	
	% correct	73.0	82.9	81.0	81.6	
	MSE	.325	.285	.276	.281	

Equal-weight model

Fit index	Set size 1	2	4	8	Mean
ρ^2	.216	.292	.277	.356	.285
% correct	68.9	75.1	74.0	79.8	74.5
MSE	.408	.381	.399	.406	.399

[a]In this analysis, random assignment would predict the target correctly 50 percent of the time.

These theoretical results were then tested in the context of a controlled laboratory experiment which focused on preferences for hypothetical apartments displayed in choice sets of varying sizes. The results provided mixed support for the research hypotheses. As predicted, there was an increase in the tendency for subjects to use elimination strategies when faced with larger set sizes. Associated with these changes in processing rules were concomitant changes in attribute usage and revealed attribute importance. In particular, the most important apartment attribute—rent—was mentioned more frequently and increased in revealed salience with larger set sizes, while other attributes (maintenance and appearance) gradually decreased in overall importance.

A major prediction not supported by the data was a decrease in the predictive ability of a compensatory model with increases in set size. In terms of root mean squared error and proportion of correct predictions, there was actually a slight increase in the ability of the compensatory logit model to characterize observed choice data with increases in set size.

There are several reasons why this hypothesis may not have held. Our theoretical analysis addressed only error resulting from the mismatch of model to the underlying process, and not error due to other sources. Any other stochastic error was assumed to be homogeneous across set sizes. This assumption may not be reasonable, as other sources of error may well change with set size. For example, individuals may apply choice rules less consistently in one set size than in another, or they may disagree more on attribute importance in one set size than in another. This latter effect could explain the apparent increase in the fit of the logit model across levels of set size; while larger set sizes tend to increase the use of elimination strategies, they may also lead individuals to adopt the same attribute for screening, in this case, rent.

Thus any decrease in fit due to an inappropriate model may be masked by an increase in fit due to greater homogeneity in decision processes. Note, however, that although individual analyses would eliminate this one source of error, it would not eliminate others, such as the inconsistent application of a decision rule. Clearly, the approximation error is more difficult to test unambiguously than it might initially appear.

One intriguing aspect of the results was the uniformly high levels of explanation provided by the compensatory (logit) model. Although significant parameter variation was observed across set sizes, the absolute magnitude of this variation was relatively small and did not involve a reversal of the rank order of importance. Because of this, derived models were remarkably accurate in prediction across different set sizes. For example, the model that was derived on the basis of judgments predicted choice behavior nearly as well as the models which were explicitly estimated on choices.

An apparent implication of this result is additional credibility for methods of consumer forecasting that rely on the ability of compensatory models of decision making derived in one context—such as judgments—to predict choices made in another—for example, choice simulators (e.g., Green et al. 1980, 1981). The fact that the "true" underlying decision process may be different for judgments and choices in a given context may well have little effect on cross-task predictive ability.

Such optimism, however, should be tempered by considering the generalizability of the current results. Recall that in the current experiment, apartment attributes were uncorrelated. The robustness of compensatory models given uncorrelated attributes is well known, so this experimental result might have been anticipated (e.g., Einhorn et al. 1979). Yet in the real world, this will not always be the case. Like "price" and "quality," one might expect many attributes to be negatively correlated. This negative correlation increases differences in the choices made by compensatory and noncompensatory processes (Einhorn et al. 1979; Newman 1977).

Moreover, information search costs in the experiment were small when compared to many nonlaboratory situations; all relevant information was presented on a single page. In actual decision situations, search can be much more costly, involving extensive investment of time and effort. Under these conditions, we may expect a greater degree of contingent processing, with a subsequent decrease in the ability of a compensatory approximation to predict choices.

The primary contribution of this research is the demonstration that compensatory approximations may be transferable under certain conditions. It then becomes an interesting empirical question for future research to identify the boundaries surrounding transferability. Although there are a number of reasons for suspecting that the robustness of models reported here may not extend into real world markets, this remains an empirical issue.

In closing, we should note that the conditions for a model failure can be detected, in part, through the use of multiple methods for representing decision processes. The use of noncompensatory strategies has proven to be a difficult phenomenon to detect within a linear models framework, yet it is readily apparent from even a casual examination of verbal protocols. Similarly, the robustness of compensatory approximations for prediction could be easily established at a group level only by a revealed preference model. Hence, we suggest that the use of two techniques and the combination of individual and group analyses has yielded greater insight than was possible by any single analysis. Further use of such parallel analyses could help in understanding the effect of changes in choice context on the prediction and explanation of consumer choice.

[*Received April 1983. Revised January 1984.*]

REFERENCES

Anderson, Norman H. (1982), *Methods of Information Integration Theory*, New York: Academic Press.

Avery, Robert E. (1982), "Limited Dependent Variable Program CRAWTRAN," working paper, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA.

Batsell, Randy R. (1980), "Consumer Resource Allocation Models at the Individual Level," *Journal of Consumer Research*, 7 (June), 78–87.

Bettman, James R. (1979), *An Information Processing Theory of Consumer Choice*, Reading, MA: Addison-Wesley.

——— and Jacob Jacoby (1976), "Patterns of Processing in Consumer Information Acquisition," in *Advances in Consumer Research*, Vol. 3, ed. Beverlee B. Anderson, Ann Arbor, MI: Association for Consumer Research, 315–320.

——— and C. Whan Park (1980), "Effects of Prior Knowledge and Experience on Consumer Decision Processes: A Protocol Analysis," *Journal of Consumer Research*, 7 (December), 234–248.

Billings, Robert and Stephan Marcus (1983), "Measures of Compensatory and Noncompensatory Models of Decision Behavior: Process Tracing versus Policy Capturing," *Organizational Behavior and Human Performance*, 31 (August), 331–352.

Birnbaum, Michael F. (1973), "The Devil Rides Again: Correlation of An Index of Fit," *Psychological Bulletin*, 79 (2), 239–242.

Currim, Imran (1982), "Predictive Testing of Consumer Choice Models Not Subject to Independence of Irrelevant Alternatives," *Journal of Marketing Research*, 19 (May), 208–222.

Curry, David J., Jordan J. Louviere, and Michael Augustine (1981), "Comment on the Sensitivity of Brand Choice Simulators to Attribute Importance Weights," *Decision Sciences*, 12 (July), 502–516.

Dawes, Robin M. and Bernard Corrigan (1974), "Linear Models in Decision Making," *Psychological Bulletin*, 81 (2), 95–106.

Einhorn, Hillel J. (1970), "The Use of Nonlinear, Noncompensatory Models in Decision Making," *Psychological Bulletin*, 73 (3), 221–230.

———, Donald N. Kleinmuntz, and Benjamin Kleinmuntz (1979), "Linear Regression and Process-Tracing Models of Judgment," *Psychological Review*, 86 (5), 464–485.

Green, Paul E. (1974), "On the Design of Choice Experiments Involving Multifactor Alternatives," *Journal of Consumer Research*, 1 (September), 61–68.

———, J. Douglas Carroll, and Steven M. Goldberg (1981), "A General Approach to Product Design Optimization Via Conjoint Analysis," *Journal of Marketing*, 45 (3), 17–37.

———, Wayne S. DeSarbo, and Paul Kedia (1980), "On the Sensitivity of Brand-Choice Simulators to Attribute Importance Weights," *Decision Sciences*, 11 (July), 439–450.

Grether, David and Louis Wilde (1984), "An Analysis of Conjunctive Choice: Theory and Experiments," *Journal of Consumer Research*, 10 (March), 373–385.

Hahn, G. J. and S. S. Shapiro (1966), "A Catalog and Computer Program for the Design and Analysis of Orthogonal Symmetric and Asymmetric Fractional Factorial Experiments," Technical Report 66-C-165, General Electric Research and Development Center, Schenectady, New York.

Hensher, David A. and Lester Johnson (1981), *Applied Discrete Choice Modeling*, Landon-Croom-Helm/New York: John Wiley.

Johnson, Eric J. and J. Edward Russo (1984), "Product Familiarity and Learning New Information," *Journal of Consumer Research*, 11 (June), 542–550.

Keeney, Ralph L. and Howard Raiffa (1976), *Decision with Multiple Objectives: Preferences and Value Tradeoffs*, New York: John Wiley.

Lichtenstein, Sara and Paul Slovic (1971), "Reversal of Preference Between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology*, 89 (1), 46–55.

Louviere, Jordan J. (1979), "Modeling Individual Residential Preferences: A Totally Disaggregate Approach," *Transportation Research*, 13 (A), 374–384.

——— and George Woodworth (1983), "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350–367.

McFadden, Daniel (1981), "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, eds. Charles F. Manski and Daniel McFadden, Cambridge, MA: MIT Press, 198–272.

Meyer, Robert J. and Thomas C. Eagle (1982), "Context-Induced Parameter Instability in a Disaggregate-Stochastic Model of Store Choice," *Journal of Marketing Research*, 19 (February), 61–71.

Nakanishi, Masao and Lee G. Cooper (1982), "Simplified Estimation Procedures for MCI Models," *Marketing Science*, 1 (3), 314–322.

Newell, Allan and Herbert S. Simon (1972), *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.

Newman, J. Robert (1977), "Differential Weighting in Multiattribute Utility Measurement: Where It Should and Where It Does Make a Difference," *Organizational Behavior and Human Performance*, 20 (December), 312–325.

Olshavsky, Richard W. (1979), "Task Complexity and Contingent Processing in Decision Making: A Replication and Extension," *Organizational Behavior and Human Performance*, 24 (December), 300–316.

——— and Frank Acito (1980), "An Information Processing Probe into Conjoint Analysis," *Decision Sciences*, 11 (July), 451–470.

Payne, John W. (1976), "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance*, 16 (August), 366–387.

——— (1982), "Contingent Decision Behavior," *Psychological Bulletin*, 92 (September), 382–402.

Robson, Donald S. (1959), "A Simple Method for Constructing Orthogonal Polynomials When the Independent Variable is Unequally Spaced," *Biometrics*, June, 187–191.

Russo, Jay E. and Barbara A. Dosher (1983), "Strategies for Multiattribute Binary Choice," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9 (October), 676–696.

Svenson, Ola (1979), "Process Descriptions of Decision Making," *Organizational Behavior and Human Performance*, 23 (1), 86–112.

Tversky, Amos (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79 (July), 281–299.

Urban, Glen L. and John R. Hauser (1980), *Decision and Marketing of New Products*, Englewood Cliffs, NJ: Prentice-Hall.

Wright, Peter L. and Frederick Barbour (1977), "Phased Decision Strategies: Sequels to an Initial Screening," in *North-Holland/TIMS Studies in the Management Science, Volume 6: Multiple Criteria Decision Making*, eds. M. K. Star and Milan Zoleny, Amsterdam: North Holland, 91–109.