Greedoid-Based
Non-compensatory Consideration-then-Choice Inference

Michael Yee

Ely Dahan

John R. Hauser

James Orlin

January 2005

Michael Yee is a graduate student at the Operations Research Center, E40-149, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 253-3601, myee@mit.edu.

Ely Dahan is an Assistant Professor of Marketing at the Anderson School, University of California at Los Angeles, 110 Westwood Plaza, B-514, Los Angeles, CA 90095, (310) 206-4170, fax (310) 206-7422, edahan@ucla.edu.

John R. Hauser is the Kirin Professor of Marketing, Sloan School of Management, Massachusetts Institute of Technology, E56-314, 38 Memorial Drive, Cambridge, MA 02142, (617) 253-2929, fax (617) 253-7597, jhauser@mit.edu.

James Orlin is the Edward Pennell Brooks Professor of Operations Research and Co-director, MIT Operations Research Center, Massachusetts Institute of Technology, E40-417, 77 Massachusetts Avenue, Cambridge, MA 02139, (617) 253-6606, jorlin@mit.edu.

This research was supported by the MIT Sloan School of Management, the Center for Innovation in Product Development at MIT, the Operations Research Center at MIT, and by the Office of naval Research, contract N00014-98-1-0317. This paper may be downloaded from http://mitsloan.mit.edu/vc. We wish to thank Ashvini Thammaiah who assisted in the development and pretesting of the web-based questionnaire. The study benefited from comments by numerous pretests at the MIT Operations Research Center and the Kappa Alpha Theta Sorority at MIT. SmartPhone images produced by R. Blank. Our thanks to Theodoros Evgeniou, Shane Frederick, Steven Gaskin, Rejeev Kohli, and Olivier Toubia for insightful comments on an earlier draft. This paper has benefited from presentations at the MIT Operations Research Center, the Center for Product Development, the MIT Marketing Group Seminar, Michigan‘s Ross School of Business Seminar Series, the 2004 Marketing Science Conference in Rotterdam, The Netherlands, the AMA’s 2004 Advanced Research Techniques Forum, Whistler, B.C., the 2004 Explor Award Case Study Showcase in New Orleans, LA, and the 2004 Management Roundtable Voice-of-the-Customer Conference, Boston MA. The final three names are listed alphabetically. Contributions were extensive and synergistic.

Greedoid-Based
Non-compensatory Consideration-then-Choice Inference

Abstract

We propose practical methods to test whether respondents use non-compensatory processes and, if so, to infer the details of those processes from either consideration-then-choice or full-rank tasks. Inference is a non-trivial combinatorial problem which has hitherto been computationally infeasible for practical problems. "Greedoid languages" provide a structure and theory to transform this problem and decrease estimation time by a factor of approximately 10^9 for practical 16-aspect problems. Monte Carlo experiments suggest that it is feasible to infer, albeit with noise, the process that respondents use to evaluate product profiles.

We test Greedoid inference empirically by asking 481 respondents to evaluate SmartPhones. The data suggest that most respondents use a lexicographic strategy in this category. Interestingly, allowing respondents to sort profiles by features does not induce more lexicography. Two additional experiments use the Greedoid inference engine to (1) replicate the result that non-compensatory processes become more likely as the number of products increases and (2) examine the impact of pruning profiles adaptively so that the only feasible rank-orders are those that are consistent with linear combinations of partworths.

Greedoid methods also provide a non-compensatory conjoint-like method to forecast consumer response and to find the minimum feature levels required to enter a product category. This method is particularly useful in categories where consumers are presented with large numbers of potential choices. Not only do Greedoid estimates appear to predict better (in most cases) than purely compensatory estimates, but the consideration-then-choice task is perceived by respondents as more enjoyable, more accurate, and more interesting. It also saves time and increases completion rates – both of which translate directly into cost savings.

Keywords: Lexicography, Non-compensatory decision rules, Choice heuristics, Optimization methods in marketing, Conjoint Analysis

1. Two-Stage Consideration/Choice Decision Processes

Personal digital assistants (PDAs) are popular productivity enhancing tools. On the website of our university's approved supplier, GovConnection, we find 97 PDAs from which we can choose. Local retailers also have moderately broad product lines: 21 at Circuit City, 25 at Staples, 27 at Microcenter, and 33 at CompUSA. On websites and in retail stores sellers provide tools to simplify PDA choice for consumers. For example, Staples, CompUSA, and Microcenter structure the choice sets (e.g., Palm-OS vs. Windows CE) while all retailers and websites encourage consumers to self-organize the choice sets by operating system, brand, price, or other features. See Figure 1. This retail environment is consistent with (but not proof of) a two-stage decision process in which consumers first judge which products to consider, perhaps by a heuristic judgment process. Consumers might then choose from the set of considered products. Processing rules might be either compensatory or non-compensatory.

Figure 1
PDAs Available from CompUSA

There is ample evidence in the psychology, consumer-behavior, and marketing-science literatures that two-stage consideration-then-choice decision processes describe, at least paramorphically, real decision processes in a variety of circumstances (e.g., Bettman and Park 1980; Einhorn and Hogath 1981; Gensch 1987; Hauser and Wernerfelt 1990; Montgomery and Svenson 1976; Payne 1976; Payne, Bettman and Johnson 1993; Roberts and Lattin 1991). Furthermore, many consumers simplify the first-stage judgment with a heuristic process (e.g., Bett-

man, Luce and Payne 1998; Bröder 2000; Einhorn 1970; Gigerenzer and Goldstein 1996; Martignon and Hoffrage 2002). Heuristic simplification is more likely when the initial set of products is large (e.g., Johnson and Meyer 1984; Payne, Bettman and Johnson 1993).

Two-stage decision processes are important and interesting scientifically and managerially. Firstly, researchers in consumer behavior are interested in studying when (if) such two-stage processes are used. By using a variety of methodologies ranging from verbal process tracing to information display mechanisms (e.g., Mouselab), researchers have studied how consumers adapt and/or construct decision processes based on the characteristics of the decision environment (see Payne, Bettman and Johnson 1993 for a review). Some researchers suggest that two-stage heuristic processes might even be appropriately robust strategies for consumers and that such strategies are based in evolution (Chase, Hertwig and Gigerenzer 1998; Dawes 1979; Dawkins 1998). From a policy perspective, the popularity of web-based virtual advisors suggests that consumers find value in heuristics that help them search categories that contain many brand-feature combinations.

Secondly, if consumers use two-stage processes, then data obtained directly by observing a two-stage process may more accurately reflect their behavior because it is consistent with the consumers' actual decision processes. Further, a two-stage data collection task may by more enjoyable and less effortful for respondents (e.g., Malhotra 1986; Oppewal, Louviere and Timmermans 1994; Srinivasan and Park 1997). A less-time-consuming and more-enjoyable task might mean shorter questionnaires (less cost) and may encourage more respondents to complete the task (fewer non-response issues).

Finally, the identification of heuristic processes is important managerially. Not only does it inform retailers and web-site designers how to organize their shelves and websites, but it suggests to product designers which features to include and to advertisers how they might present products (and choices) to consumers. For example, in the automobile market consumers consider fewer than 5% of the make-model combinations available (5-10 out of 350+, Urban and Hauser 2004). In consumer packaged goods knowing the consideration set can explain up to 80% of the uncertainty in choice (Hauser 1978). A big part of the battle for sales is won if a manufacturer can design products that consumers consider. In advertising, knowledge of consumers' heuristics might suggest which features to stress in positioning or in unique selling propositions (e.g., Batra, Myers and Aaker 1996; Ogilvy 1985; Reeves 1961).

In this paper we explore new methodologies to study two-stage consideration/choice decision processes. These methodologies build upon a structured theory called Greedoid languages and require that we solve combinatorial computational problems significantly more efficiently than as reported in the extant literature. We demonstrate with simulation and empirical tests how the methodologies can be used to infer the underlying decision process by observing consideration, then choice from a set of product profiles. The proposed web-based data-collection and inference methods appear to be accurate while reducing respondent burden relative to common benchmarks.

The paper proceeds as follows. In the next section we review briefly the literature on two-stage processes and in a subsequent section we describe the respondents' tasks. We then present the Greedoid-based methods for non-compensatory inference and discuss the compensatory inference methods to which they are compared. The inference and data-collection methods are tested with a 4x2x4 simulation experiment – four generating models that vary from fully compensatory to fully non-compensatory, crossed with low vs. high respondent errors, crossed with four inference models. We then test the methods empirically in a 2x2 experiment in which 339 respondents choose from a 32 SmartPhones chosen as a fractional factorial from a $4^3 2^4$ design. Respondents (a) either sort all 32 profiles as in standard full-profile conjoint analysis or use the simplified two-stage task and (b) either are allowed to presort the profiles or not. Finally, as examples of the use of Greedoid-based inference, we analyze two additional experiments with 142 respondents: (1) whether the use of non-compensatory processes is affected by the number of profiles in the task (16 vs. 32) and (2) whether we can induce compensatory processing by pruning the profiles so that the only feasible rank-orders are those that are consistent with linear combinations of partworths. The first experiment demonstrates how the new methods can be used to replicate previous findings. The second experiment is explores new issues.

2. Brief Review of Two-Stage, Potentially Non-compensatory Decision Processes

We consider decision processes in which products are represented by their features and consumers must decide which product to purchase or consume. While the process by which consumers encode products into features can be complex and important (Einhorn and Hogarth 1981), this topic is beyond the scope of this paper. Our scope includes situations in which such encoding is feasible and reasonably descriptive of consumer decision processes. For practical applications we might use voice-of-the-customer methods to identify a representative set of fea-

tures (e.g., Griffin and Hauser 1993, Zaltman 1997). When a feature is binary, it is called an aspect (e.g., Tversky 1972). Multi-level features can be considered collections of aspects that are related (Verizon vs. Cingular vs. Nextel vs. Sprint for SmartPhone service providers). A profile is the feature description of a product.

Non-compensatory Processes

In a compensatory process high levels on some aspects can compensate for low levels on other aspects. In a non-compensatory process high levels on some aspects may not compensate for low levels on other aspects. One well-known process is a lexicographic process: consumers evaluate profiles first by one feature, then another until a judgment or choice is made (Fishburn 1974; Nakamura 2002). For example, for 32 SmartPhones, a consumer might rank the 8 Verizon phones first, the 8 Cingular phones next, then the 8 Nextel phones, then the 8 Sprint phones.

The literature rarely distinguishes whether the lexicographic ranking is with respect to features or aspects, but the implied rank orders of profiles depends on this distinction. For example, ranking first by "Verizon then Cingular then Nextel then Sprint" is a rank by the feature of "service provider." However, the consumer might rank more generally. The consumer might first rank by "Verizon vs. not Verizon," then by "flip phones vs. non-flip phones," and only then by "Cingular vs. not Cingular."[1] When ranking by "service provider," a Cingular non-flip phone might be ranked <u>before</u> a Nextel flip phone. If the ranking allows the consumer to evaluate "flip vs. non-flip" between evaluations of "Verizon vs. non-Verizon" and "Cingular vs. non-Cingular," a Cingular non-flip phone would be ranked <u>after</u> a Nextel flip phone. When the consumer processes all aspects associated with a feature at the same time, we call this process "lexicographic by feature (LBF)." We call the less restrictive process that allows the consumer to split aspects "lexicographic by aspect (LBA)." Note that LBF is a special case of LBA.

There are two important special cases of LBA. For multi-aspect features, consumers can either eliminate profiles by aspects, e.g., first eliminate all Sprint phones, or they can accept profiles by aspects, e.g., rank first all Verizon phones. For multi-aspect features, an elimination strategy may yield a different profiles ordering than an acceptance strategy. Following Tversky (1972) we call the former "elimination by aspects (EBA)." We call the latter "acceptance by aspects (ABA)."[2] Because LBA allows aspects to be either acceptance or elimination criteria,

[1] Naturally, within the category of "Verizon" SmartPhones there are no "Cingular" SmartPhones.
[2] Kohli and Jedidi (2004) analyze ABA under the name "binary lexicographic model."

LBA = EBA ∪ ABA. The distinction among processes is only necessary for multi-aspect features. If all features are binary (aspects) then LBA = LBF = EBA = ABA. (Tversky defines EBA as a random process in which the probability that an aspect is chosen is proportional to its measure. In this paper we follow Johnson, Meyer and Ghose (1989), Montgomery and Svenson (1976), Payne, Bettman and Johnson (1988), and Thorngate (1980), and use EBA to refer to a deterministic process in which an aspect order is given.)

We might wish to impose some constraints on processing strategies. For example, for a "price" feature, we might want to assume that, when ranking by price, lower prices are preferred to higher prices. Such constraints are allowable, but not required, in the methods we develop. Finally, LBA, LBA, EBA, and ABA can each be used either to rank profiles or to classify profiles into considered vs. not-considered profiles.

Compensatory Processes

Many authors represent a compensatory process as an arithmetic rule in which each aspect receives a weight and consumers sum the weights associated with the aspects in a profile to form "utility." Consumers then choose the product with the highest "utility." However, this arithmetic rule also produces lexicographic (and other non-compensatory) processes if the weights are sufficiently extreme. For example, if the aspects are ordered by importance and if the nth smallest aspect weight is proportional to 2^{1-n}, then an additive model produces a lexicographic process (Jedidi, Kohli and DeSarbo 1996; Kohli and Jedidi 2004; Olshavsky and Acito 1980). For this paper we call a process compensatory if the weights are less extreme, that is, if the weighted process is constrained so that it is not equivalent to a non-compensatory process.

Constructive Consideration and Choice Processes

Research suggests that consumer decision processes are contingent on many context effects including the range of aspects, correlation among aspects, base-rate information, reference points, the size of the choice set, the relevance of the decision, and the difficulty of comparison (see review in Payne, Bettman and Johnson 1993). For example, we expect most respondents to use a non-compensatory process when the number of products is large as in our PDA example or our empirical test with 32 SmartPhones profiles (e.g., Johnson and Meyer 1984). We expect the more respondents to use a compensatory process with fewer profiles – say 16 SmartPhones.

There are many hypotheses to explain observed non-compensatory processes. Some hypotheses rely on a cost-benefit tradeoff between effort and decision accuracy (e.g., Shugan

1980). Other hypotheses suggest that non-compensatory processes are inherently more robust and are best suited for judgment tasks when preferences are to be constructed, when consumers are uncertain about the valuations of the aspects, or when emotion and "rationality" are integrated (e.g., Luce, Payne and Bettman 1999; Martignon and Hoffrage 2002). From our perspective, we seek to infer the net result of any (unobserved) process construction. To the extent that data collection approximates the essential characteristics of real choice environments, the methods should infer the processes appropriate for those environments. Fortunately, our methods are readily extendable to other respondent tasks. We leave such exploration to future research.

Existing Methods to Infer Non-compensatory Processes

Because decision processes might be contingent, measurement of consumer decision processes attempts to approximate the conditions under which real consumers make decisions. Many measurement examples in the marketing science literature are consistent with non-compensatory or hierarchical decision processes. For example, both Srinivasan and Wyner's (1988) Casemap and Johnson's (1991) Adaptive Conjoint Analysis (ACA) include steps in which respondents are asked to eliminate unacceptable levels. Because this task is often difficult for respondents (Green, Krieger and Bansal 1988; Klein 1988), other researchers have attempted to infer the elimination process in a single estimation step (DeSarbo, et. al. 1996; Gilbride and Allenby 2004; Gensch 1987; Gensch and Soofi 1995; Jedidi and Kohli 2004; Jedidi, Kohli and DeSarbo 1996; Kim 2004; Roberts and Lattin 1991; Swait 2001). For example, Gilbride and Allenby (2004, p. 399) use hierarchical Bayes methods to analyze choice-based conjoint data to infer screening rules for cameras. They estimate that 58% of the respondents screen on a single feature, 33% on two features, 2% on three features, and 8% use fully compensatory processes.

In psychology, Bröder (2000) analyzes choices among two profiles described by four aspects. He compares the fit of an unconstrained additive model to two additive models: (1) a model with the aspects constrained to 2^{1-n} (non-compensatory) and (2) an additive model with equal weights (Dawes' 1979 model). In one experiment 28% of the forty respondents use a non-compensatory process while none follow Dawes' model. The remainder, 72%, could not be classified. Bröder's method is feasible for a small number of aspects – with four aspects the ratio of the largest-to-smallest partworth in a non-compensatory model is based on 2^3 ? 8:1. For the sixteen aspects, as in our simulation and empirical experiments, the range of partworths in a non-compensatory model would be at least 2^{15} ? 32,768:1, a ratio that puts severe strains on any sta-

tistical regression-like procedure.[3] We would like to develop a procedure that is more robust and can classify more than 28% of the respondents.

Kohli and Jedidi (2004) propose a greedy heuristic to identify a linear representation of lexicographic processes from metric conjoint data.[4] They modify Bröder's procedure by computing the number of violated pairs between predicted and observed rank orders for both linear and 2^{1-n} models. Their *t*-statistics suggest that a 2^{1-n} representation is not significantly different from an unconstrained linear model for 67% of the 69 respondents who evaluated profiles with five features (eleven aspects).

Our approach differs from methods that infer a (possibly) non-compensatory consideration process with a latent construct and from methods that ask directly for unacceptable levels. We use a measurement process chosen to approximate how consumers make decisions in real environments. We ask respondents to identify directly which product profiles they would consider and then obtain full rankings only among the remaining profiles. The measurement process is similar to that employed by Malhotra (1986), which, in turn, is related to that used by Bettman and Park (1980). (However, Malhotra analyzed only the second-stage process which he modeled as compensatory.) With these data we attempt to identify (1) the best non-compensatory process and (2) the best compensatory process to explain each respondent's consideration-then-rank (or rank-only) data. Our approach differs from Kohli and Jedidi (2004) along a number of dimensions including respondent task, inference algorithms, and focus. Nonetheless, these parallel independent studies suggest many opportunities to apply discrete optimization methods to infer non-compensatory processes.

3. The Respondent's Task

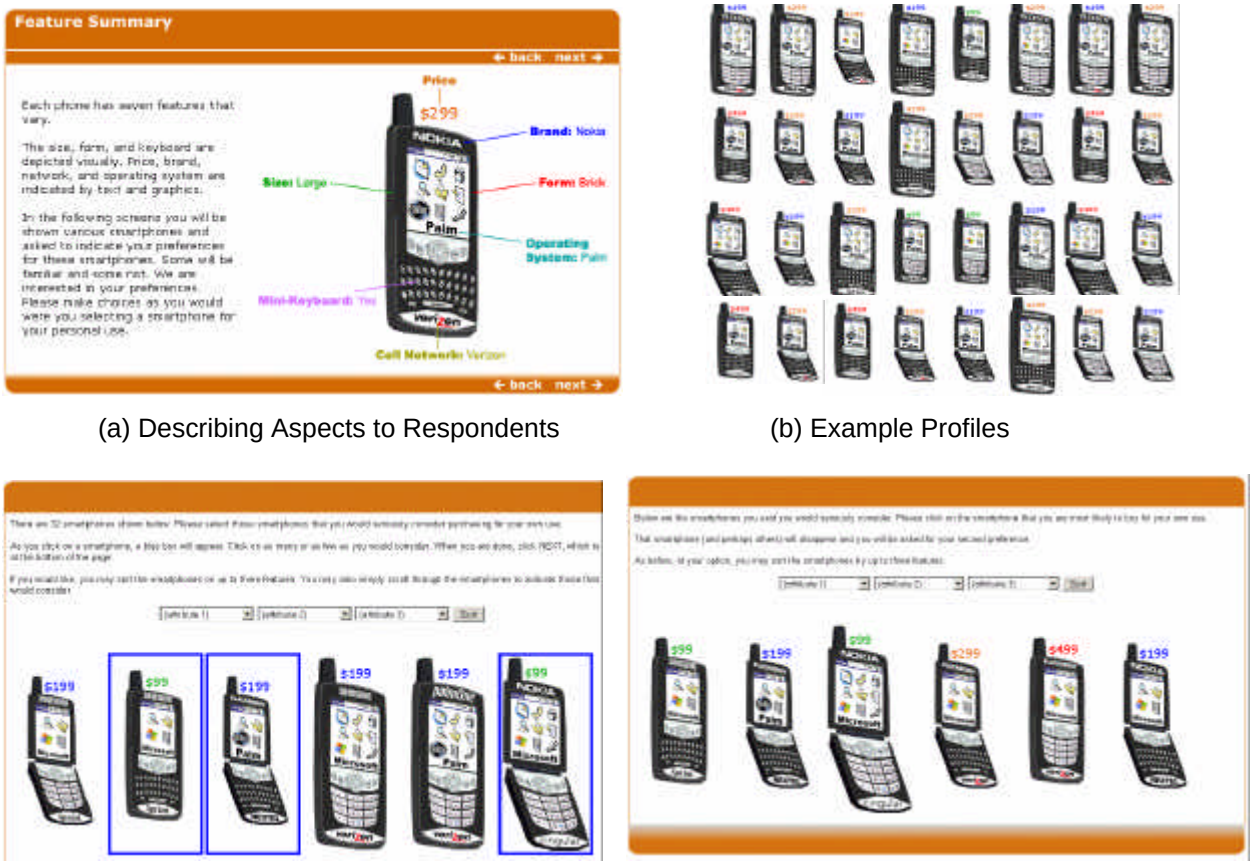
It is easier to understand the theory if we first review the respondents task as illustrated with an example from our empirical experiments.[5] Respondents are first introduced to the product category and the seven features (sixteen aspects). Figure 2a is one of many screens. Respondents are then presented with SmartPhone profiles (Figure 2b). Respondents in the (two-stage)

[3] The ratio is not as severe if we allow partial lexicographic orders. However, Bröder's method cannot handle partial orders without first solving the combinatorial problems we describe later.
[4] Both approaches were developed independently. We became aware of one another's approaches after all empirical work had been completed and papers written. Perhaps future research might modify each of the methods so that they might be compared on the same respondent task.
[5] Greedoid analysis can be extended to tasks such as those used by Bröder (2000), Gigerenzer and Goldstein (1996), or Gilbride and Allenby (2004).

consideration-then-choice cells simply click on those profiles they would seriously consider – part of the screen is shown in Figure 2c. These respondents then see only their considered profiles; they are asked to rank them by successively clicking on the profile they would choose from the offered set (Figure 2d). That profile disappears and they choose again until all considered profiles are chosen. Respondents in the one-stage cells skip the consideration task. In Figure 2 we illustrate an additional twist. Some, but not all, respondents were allowed to presort the profiles in either or both tasks. In a later section we describe the full experimental protocol (incentives, filler tasks, holdout measures, etc.) and provide statistics such as response rates.

A few comments are in order. Firstly, respondents are <u>not</u> asked to indicate unacceptable feature levels (aspects) directly. Any unacceptable levels (aspects) are inferred from the data. Secondly, the measurement tasks themselves do not assume that the judgment (consideration) and decision (choice) processes are compensatory or that they are non-compensatory. This, too, is inferred from the data.

Figure 2
Illustrative Respondent's Task



(a) Describing Aspects to Respondents

(b) Example Profiles

(c) Consideration Judgment Task

(d) Choice from Consideration Set

Inferring Non-compensatory Processes with Greedoid Languages

For ease of exposition we consider lexicographic orderings generated by acceptance-by-aspects (ABA). Elimination-by-aspects (EBA) can be inferred with the same algorithm by redefining aspects by their negation; lexicographic-by-features (LBF) can be inferred by imposing constraints on the aspect orders; and lexicographic-by-aspects (LBA = EBA $\cup$ ABA) can be inferred with only a slight modification in the algorithms.

To infer the best lexicographic representation of the consideration-then-rank and the full-rank process we develop a procedure to identify the aspect order that maximizes fit (minimizes errors) on some metric. Unfortunately, as Martignon and Hoffrage (2002, Theorem 2, p. 39) prove, this problem is NP-hard and suggest exhaustive enumeration. For example, they sought to determine the aspect order (state capital, soccer team in the national league, etc.) that best explains the relative populations of two German cities. Because their problem had nine aspects, they needed to search 9! orderings – "a UNIX machine" took two days to find the best ordering. Their problem is a relatively small problem. For our $4^3 2^4$ empirical problem we need to check 16! aspects orders for ABA and 13! orders for LBF.[6] Because 13! = 17,160*9! – their algorithm would have taken almost 100 years per respondent for LBF. Because 16! = 3,360*13!, their algorithm would have taken over 300 millennia for ABA, EBA, or LBA. Although faster computers help, it is clear that we must find a more efficient algorithm if we are to develop a practical method to infer non-compensatory processes.

We address computational efficiency in two steps. We first demonstrate that a partial lexicographic ordering forms a "Greedoid language." This enables us to use established results to find the appropriate lexicographic ordering, if one exists, much more efficiently than existing methods. Because the ordering need only be partial, we can handle consideration data and/or partial rank data.

A perfect lexicographic ordering of aspects is extremely rare. With 32 profiles there are 32! rank orders, but only 16! aspect orders. Thus, the chances that an arbitrary profile order is consistent with an aspect order is less than 8.0×10^{-23}. Any respondent errors are likely to cause the data to be inconsistent with an aspect order.[7] To infer non-compensatory processes in the

[6] In a $4^3 2^4$ design, there are $3 \times 4 + 4 \times 1 = 16$ aspects. If we group aspects by features, we can select one aspect of each multi-level feature as the base aspect. This leaves 13 independent aspects.

[7] Interestingly, less that 1/10th of 1% of the profile orderings are consistent with linear combination of aspect measures. This includes both compensatory and non-compensatory linear combinations.

presence of response errors, we prove that a dynamic program can find the best ordering relative to a commonly-used goodness-of-fit metric. The dynamic programming algorithm substantially reduces computation and makes it feasible to identify the best lexicographic ordering for large samples of respondents and moderately large numbers of aspects. For 16 aspects, the dynamic program need only evaluate (the order of) 2^{16} = 65,536 subsets of aspects, comparable to approximately 0.0000003% of the 16! orderings.[8] We begin with notation and definitions.

Partial Orders and Consistency

Let $L = (L_1, \ldots, L_a)$ be an ordered subset of N aspects where $a \leq N$. For a given profile, P, we let $L_i(P)$ be 1 if profile P contains aspect L_i and 0 otherwise. We write $P \succ_L P'$ if $L_i(P) = 1$ and $L_i(P') = 0$ where i is the first index for which $L_i(P) \neq L_i(P')$. Accordingly, for each totally ordered set L of aspects, there is a unique induced order of preferences, but the converse is not true. In some cases, there may be many different orders of aspects that lead to the same order of profiles. This is particularly true for the consideration task – if a respondent will only consider Verizon SmartPhones that flip open, then both the order (Verizon, flip) and the order (flip, Verizon) are consistent with the respondent's consideration process.

Suppose that X is a partial order of profiles for a respondent. For example, X might define which profiles are in a consideration set and which are not or X might define a rank-order within the consideration set. We write $P \succ_X P'$ if the respondent prefers Profile P to Profile P'. We say that an ordered subset L of aspects is *inconsistent* with a partial order X of profiles if there are profiles P and P' such that $P \succ_X P'$ and $P' \succ_L P$. Otherwise, we say that L and X are *consistent*. If L is an ordered subset of aspects, and if $e \notin L$ is an aspect, then (L, e) is the ordered subset of aspects obtained by appending aspect e to the end of L. Let $L \backslash Y$ denote the set L with all elements of Y deleted. For example, if L = (Verizon, flip, Palm operating system), e = (Nokia, under $299 dollars) and Y = (Palm operating system), then (L, e) = (Verizon, flip, Palm operating system, Nokia, under $299 dollars) and $L \backslash Y$ = (Verizon, flip).

[8] With 13 aspects, the dynamic program's running time is comparable to evaluating approximately 0.00013% of the aspect orders. If we allow aspects to have either orientation, e.g., a respondent might prefer "small to large" or "large to small," then the dynamic program runs on the order of $2*2^{16}$ and exhaustive enumeration the order of $2^{16}16!$. The savings are even more dramatic.

Greedoid Languages

Greedoid languages were developed by Korte and Lovasz (1985) to study conditions under which a greedy algorithm can solve optimization problems. They have proven useful in sequencing and allocation problems (e.g., Niño-Mora 2000). We believe that this is the first application in marketing or consumer behavior.

Let E be a set of aspects, let G be a collection of ordered subsets of E, let $\varnothing$ be the null set, and denote the number of elements in a set L by $|L|$. We say that G is a Greedoid language if the following conditions are satisfied:

1. $\varnothing \in G$
2. If $L \in G$, and if element $e \in E$ is the last element of L, then $L \backslash e \in G$.
3. If $L \in G$ and if $L' \in G$ and if $|L'| > |L|$, then there is an element e in $L' \backslash L$ such that $\hat{L} = (L, e)$, and $\hat{L} \in$ G.

In the appendix we demonstrate that non-compensatory consideration and rank rules form a Greedoid language. Corollary 1 follows from Proposition 1 because Algorithm 1 is a greedy algorithm that either finds a consistent aspect order, $L \in G$, of maximum length or terminates early.[9]

Proposition 1. *Let E be the set of aspects, and let X be a partial order on the profiles. Let G be the collection of ordered subsets of E that are lexico-consistent with X. Then G is a Greedoid language.*

Corollary 1*. Algorithm 1 determines whether there exists a lexicographic ordering consistent with X and, if an ordering exists, finds the ordering(s).*

Algorithm 1. (for determining if X is lexicographic)

```
Begin
    L = ∅;
    while L is not a complete order of the aspect set E do
    begin
        if there is no aspect e of E\L such that (L, e) is consistent with X, then
            quit because X is not lexicographic;
        else choose an aspect e of E\L for which (L, e) is consistent with X, and
            replace L by (L, e);
    end
end
```

[9] Kohli and Jedidi (2004) use a greedy algorithm on permutation matrices to identify a metric representation of a lexicographic ordering when there is no response error. They apply their algorithm to metric data.

Finding the "Best" Lexicographic Description

Even if real respondents are attempting to use a non-compensatory rule, they might make mistakes. Alternatively, the non-compensatory rule might only approximate their decision process. In either case, we seek to find the lexicographic ordering that best describes their partial order, X. As a measure of fit between L and X, we define "closeness" as the number of inconsistencies (violated pairs) between the profile order induced by L and that observed in X.[10] We now develop an algorithm to find the "closest" lexicographic ordering to X. Proposition 2 implies Algorithm 2 which is a dynamic program on the set of all subsets of the aspects. Because this set has dimensionality 2^N, and because 2^N is substantially less than $N!$, Algorithm 2 is feasible when exhaustive enumeration is not. In the appendix we prove Proposition 2 by showing that the marginal inconsistencies induced by adding a new aspect to an existing ordering depends only on that aspect and not on the prior orderings. This implies a (forward) recursive structure which, in turn, implies dynamic programming. Dynamic programming implies that the solution to Algorithm 2 is optimal. This dynamic program is similar to the Held and Karp's (1962) classic solution to the traveling-salesman problem.

> ***Proposition 2.*** *Let L and L' be two different permutations of the subset E' of aspects, and let e be any aspect not in E'. Then the number of inconsistencies directly caused by e in (L, e) is the same as the number of inconsistencies caused by e in (L', e).*
>
> ***Corollary 2****. Algorithm 2 identifies the best lexicographic description of a profile order, X.*

[10] Minimizing violated pairs is equivalent to maximizing Kendall's tau (1975) where $t = 1 - 2*$(fraction violated).

Algorithm 2. (for finding L that is least inconsistent with X).

```
Begin
    J(∅) = 0;
        for k = 1 to |E|
            for all (unordered) subsets, S ⊆ E of size k
                for all i ∈ S
                    c(S\i, i) = number of inconsistencies caused by aspect i fol-
                    lowing set S\i
                next i
                J(S) = min_{i∈S} [J(S \ {i}) + c(S \ {i}, i)]
                L(S) is the ordering of aspects in S yielding J(S) [retained]
            next S
        next k
end
```

When the algorithm terminates, $J(E)$ is the minimum number of inconsistencies between the respondent's profile ordering, X, relative to aspect set E. $L(E)$ is the best lexicographic orders, which may or may not be unique. Algorithm 2 applies directly to either acceptance-by-aspects or elimination-by-aspects. Fortunately, for lexicographic-by-aspects the number of steps in the algorithm only doubles. Specifically, in the innermost loop of Algorithm 2 we need only check both i and its negation. We call Algorithm 2 a Greedoid-based dynamic program.

We have programmed Algorithm 2 in Java running on an IBM 1.7 GHz laptop. For a 16-aspect problem, the run time is approximately 1.85 seconds. Relative to Martignon and Hoffrage (2002), some savings are due to Algorithm 2 and some are due to faster computers and programming languages. We project that Martignon and Hoffrage's exhaustive enumeration would take 14 years for a 16-aspect EBA problem with Java on the same computer.[11]

There is one further result that enables us to decrease the running time of the algorithm for larger problems. Proposition 3 enables the algorithm to begin with any maximally-consistent ordered subset of aspects and then use the dynamic program on the remaining aspects.

Proposition 3. *Suppose that L is an ordering of a subset of aspects that is consistent with the preferences of X. Then there is an optimal ordering of aspects that begins with the order L.*

[11] For LBA each lexicographic ordering includes all orientations (ABA or EBA). The number of solutions to check exhaustively would be $2^{16}16!$. Exhaustively enumerating LBA would take over 900 millennia. Exhaustively enumerating a nine-aspect problem would take 9 seconds for EBA (or ABA), but 1.3 hours for LBA.

4. Estimation of Compensatory Consideration-and-Rank Processes

To evaluate the adequacy of non-compensatory descriptions, we must also estimate the best compensatory description(s). Fortunately, estimation of the best set of aspect weights, $\vec{w}_c$, is well-studied. Many methods have been proposed and tested including monotonic regression, linear programming methods, hierarchical Bayes logistic models, and Bayesian hybrids (e.g., Rossi and Allenby 2003; Srinivasan and Shocker 1973; Toubia, et. al. 2004). We expect that some methods will be stressed more than others if the process is truly non-compensatory over all 16 aspects requiring a 32,768:1 ratio in the largest to smallest partworths. We select two methods that have proven accurate and robust for individual respondents and that are readily adapted to handle the constraints as defined below – linear programming (e.g., LINMAP) and analytic-center estimation. Because both provide similar estimates and fit statistics, we report only the analytic-center estimates.[12] We leave to future research other methods such as hierarchical Bayes shrinkage estimates.[13]

Scientifically, we seek to identify respondents who use a non-compensatory process and those that use a compensatory process. Because an unconstrained estimation might yield aspect partworths that are consistent with a non-compensatory (e.g., 2^{1-n}) process, we address this scientific question by constraining the linear-programming and analytic-center (AC) estimates. We use a generalization of Bröder's tests that imposes a parametric constraint on the space of feasible partworths. If w_{ic} is the partworth of the ith aspect for respondent c, then, for some $q \geq 1$:

compensatory constraints: $$w_{ic} \leq q w_{\ell c} \quad \textit{for all } \ell \neq i$$

The compensatory constraints enable us to estimate a series of models, each progressively less constrained. If $q = 1$, the compensatory constraints impose Dawes' equal-weights model. If $q \to \infty$, the model is unconstrained and all compensatory and non-compensatory models are allowed. Intermediate q (between 1 and 2^N) assure that some aspects compensate for the loss of other aspects. Because there is no prior theory to establish the best q, we report results for $q = 1$, 4, and ∞. Simulations suggest that these results bracket most compensatory models.

[12] On average, LINMAP differs from AC by 4/10^{ths} of 1% on fit and less than 1/10th of 1% on holdout violated pairs.
[13] Shrinkage estimates use data from all respondents to inform estimates for each respondent. To date, Greedoid estimators have not been extended to shrinkage estimates. In this paper, we compare "apples with apples" – compensatory and non-compensatory estimates that are focused that the level on an individual respondent.

We estimate a compensatory model subject to the linear constraints. Specifically, describe each profile, *j*, by a vector, $\vec{p}_j$, with elements 0 or 1 to indicate the presence or absence of an aspect. Denote respondent *c*'s partworths by the vector, $\vec{w}_c$. Then following constraints define a feasible region for the estimates. The last constraint is a normalization constraint.

$p_i^{'}\vec{w}_c \geq p_j^{'}\vec{w}_c$ for all *i* in the consideration set and *j* not in the consideration set
$\vec{p}_j^{'}\vec{w}_c \geq \vec{p}_k^{'}\vec{w}_c$ when *j* precedes *k* in a choice (rank) order
$w_{ic} \leq q w_{\ell c}$ *for all* $\ell \neq i$
$0 \leq w_{ic} \leq 1$ *for all i*

When $\vec{w}$'s can be found that satisfy the constraints, we use the analytic center of the feasible region as the estimates of $\vec{w}_c$. When there are no feasible $\vec{w}$'s we minimize the maximum error among the consideration and rank-order constraints. LINMAP is similar; it minimizes the average error rather than the maximum error.

5. Monte Carlo Experiments

Before we analyze empirical data, we test the accuracy of our measurement instrument. In particular, when we generate data with a non-compensatory process, we would like LBA models identified by the Greedoid-based dynamic program to predict better than the best compensatory estimates. When the generating model is compensatory, we would like the opposite to hold. For simplicity of exposition, we report Monte Carlo results for rank-orders only. We obtain qualitatively similar results for consideration-then-rank synthetic data.

For our generating model, we modify a functional form proposed by Einhorn (1970). We first define a set of generating weights, $w_n = 2^{1-n}$ for n = 1 to N. We then select each synthetic respondent *c*'s true partworths as follows: $w_{nc} = (w_n)^m = 2^{m(1-n)}$ for the nth smallest partworth. Following Einhorn, m = 0 implies Dawes' model and m = 1 implies a minimally lexicographic model. Setting $0 < m < 1$ generates a compensatory model. By setting m = 0, 1/15, 2/15, 4/15, 8/15, and 16/15 we generate range of models that are successively less compensatory. We then generate 1,000 synthetic respondents for each m as follows where u_{jc} is respondent *c*'s true utility for profile *j*.

1. For each *m*, generate $\vec{w}_c$, normalize so w_{nc}'s sum to 1.0. $\vec{w}_c \Rightarrow u_{jc} = \vec{p}'_j \vec{w}_c$.
2. For each *c*, add error to the true utility: $\tilde{u}_{jc} = u_{jc} + \tilde{e}_j$ *where* $\tilde{e}_j \sim N(0, e)$, e = 0.2, 0.4.
3. Given $\{\tilde{u}_{kc}\}$, generate a rank order of 32 cards for respondent *c*. Repeat for all *m*.

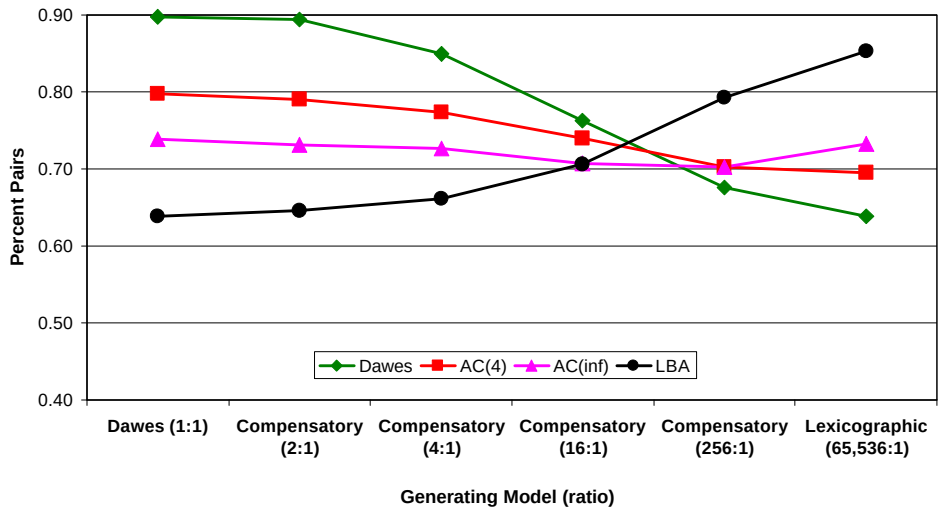
For each respondent, we then use either the Greedoid-based dynamic program to estimate an LBA aspect order or an analytic-center method to estimate partworths. Estimated partworths imply a rank-order of 32 profiles from a $4^3 2^4$ design. The fit statistic is the percent of ordered pairs of profiles from the estimated model that are consistent with the true model. The results are shown in Figure 3. For ease of interpretation and comparison with the compensatory constraints, we label the horizontal axis with the ratio of the largest to the smallest partworth. For example, m = 2/15 implies a ratio of 4:1.

There are many implications of Figure 3. First, we examine the ability of fit statistics to identify a respondents' processing strategy. If the respondent is truly lexicographic, Greedoid-based LBA estimation does much better than any of the compensatory estimates (RHS of Figure 3). If the respondent is truly compensatory, then any of the compensatory estimates do much better than the Greedoid-based LBA estimates (LHS of Figure 3). All but the smallest difference in Figure 3 is significant. (Details from the authors.) Thus, the assignment of respondents to processing strategies by best-fit should do well for respondents who are lexicographic and for respondents whose compensatory partworths are not too extreme (less than 16:1). Some ties will occur for partworths more extreme than 16:1. Interpolating Figure 3, this will occur for linear processes with ratios in the range of 32:1 (low error) and 256:1 (high error).

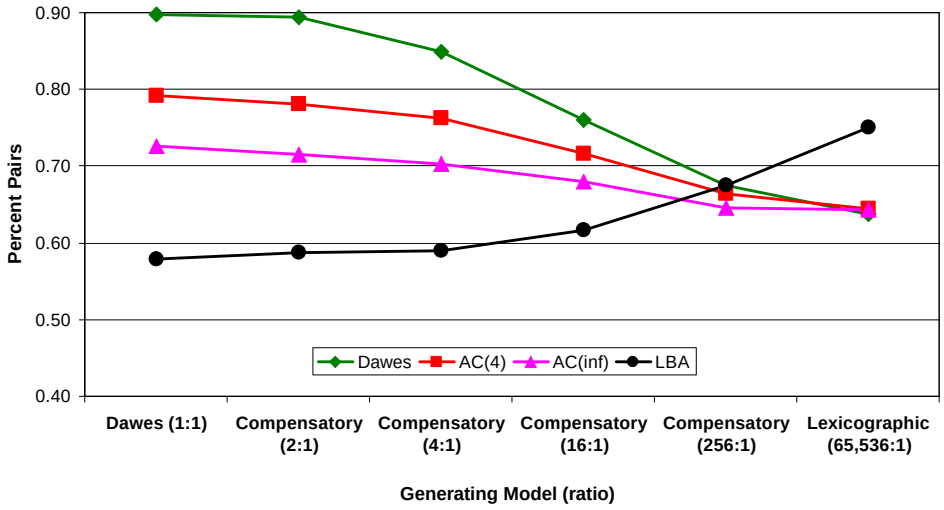
Figure 3 is also interesting in terms of robust estimators. As suggested by Dawes (1979), Dawes and Corrigan (1974), and Hagerty and Srinivasan (1991), a robust equal-partworth model does surprisingly well for low *m*. Compensatory constraints can provide useful information, e.g., $w_{ic} \leq 4 w_{\ell c}$ for AC(4) when the maximum ratio is truly 4:1. Compensatory constraints also appear to make estimation more robust. As suggested in the machine learning literature, simpler models guard against over-fitting the data (Mitchell 1997). Even though linear weights can reproduce lexicographic processes, Greedoid-based estimates appear to be less-sensitive to error and better able to identify lexicographic processes than linear estimation with either analytic-center or linear-programming methods. These results are consistent with Martignon and Hoffrage

(2002, p. 31) who suggest that lexicographic processes "tend to be robust (and) … generalize well to new, unknown data.

Figure 3
Results of the Monte Carlo Experiments



(a) Low Response Error in Simulations



(b) High Response Error in Simulations

Robustness may be driven by the number of aspects. In simulations, the differences between LBA and unconstrained linear models (AC(∞)) were more pronounced with 16 aspects than with 10 aspects (simulation results available from the authors). This robustness suggests that Greedoid-based inference should be better able to identify lexicographic processes than Bröder's (2000) regression-based methods when the number of aspects are large. Bettman, Luce

and Payne (1998, p. 200) suggest that there are conflicting results on whether increasing the number of aspects induces non-compensatory processes. Greedoid approaches might be well-suited to address this research controversy.

In summary, the Monte Carlo experiments suggest that Greedoid-based estimation should be able to identify the respondents' lexicographic process when respondents are truly lexicographic. Furthermore, classifying respondents by "best fit" (LBA vs. AC(4)) provides a viable means to classify, albeit with noise, respondents as compensatory vs. non-compensatory.

6. SmartPhone Empirical Study

To test the Greedoid method with real respondents we invited respondents to complete a web-based questionnaire about SmartPhones. The respondents were students drawn from the undergraduate and graduate programs at two universities. To the best of our knowledge, they were unaware of the Greedoid methods or the purpose of our study. As an incentive to participate, they were offered a 1-in-10 chance of winning a laptop bag worth $100, yielding a 63% response rate. Pretests in related contexts suggested that the choice of SmartPhones might include non-compensatory screening for some consumers and thus represented an interesting category for a first test of Greedoid methods.

The survey consisted of six phases. The first three phases are as described in Figure 2 – respondents were introduced to the category and SmartPhone features, indicated which SmartPhones they would consider (in half the cells), and successively chose SmartPhones in order to rank their considered products (or rank all products depending on cell). Respondents then completed a mini-IQ test to cleanse memory – a task which pretests suggested was engaging and challenging. Following this filler task, respondents completed a holdout task in which they were presented four SmartPhones chosen randomly from another 32-profile fractional factorial design. The final task was a short set of questions about the survey itself – data which we use to compare task difficulty.

For the holdout task, respondents indicated which profiles they would consider and then ranked all four profiles. They were not constrained to rank considered profiles above non-considered profiles – a small fraction did not. To avoid unwanted correlation due to common measurement methods, the holdout rank task used a different interface. Respondents used their pointing device to shuffle the profiles into a rank order as one might sort slides in PowerPoint. Pretests suggested that respondents understood this task and found it different than the task in

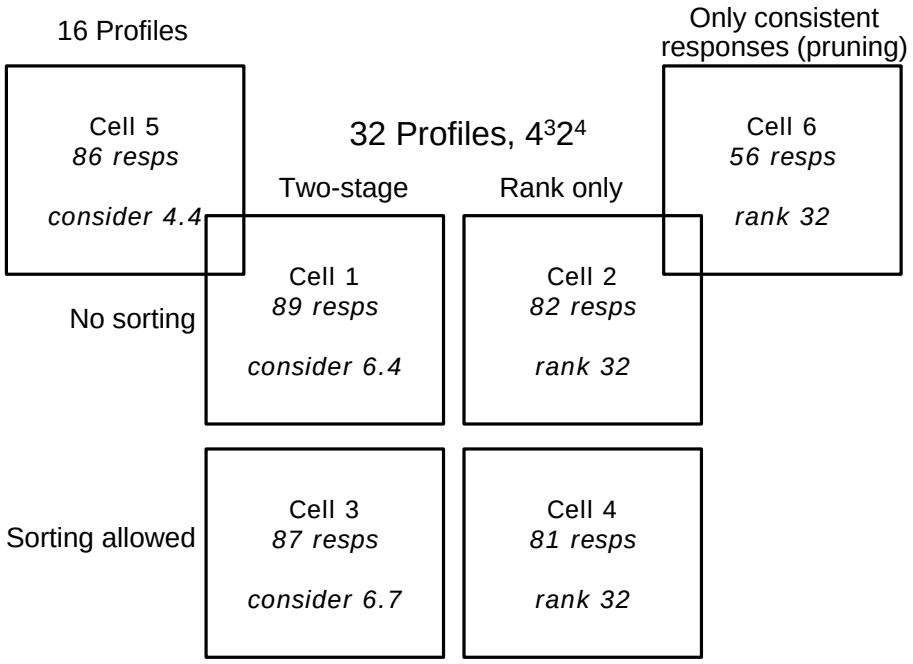
Figure 2d.

The survey was programmed in Php and debugged through a series of pretests with 56 respondents chosen from the target population. At the end of the pretests, all technical glitches were removed. Respondents understood the tasks and found them realistic.

Experimental Design

Respondents were assigned randomly to one of six experimental cells as indicated in Figure 4. The basic experimental design is a 2x2 design in which respondents complete either a two-stage or a one-stage task and are given the opportunity to presort profiles or not. (In the two-stage sort cell, respondents could sort prior to consideration and prior to choice. Respondents in the sort cells could re-sort as often as they liked.)

Figure 4
SmartPhone Experimental Design



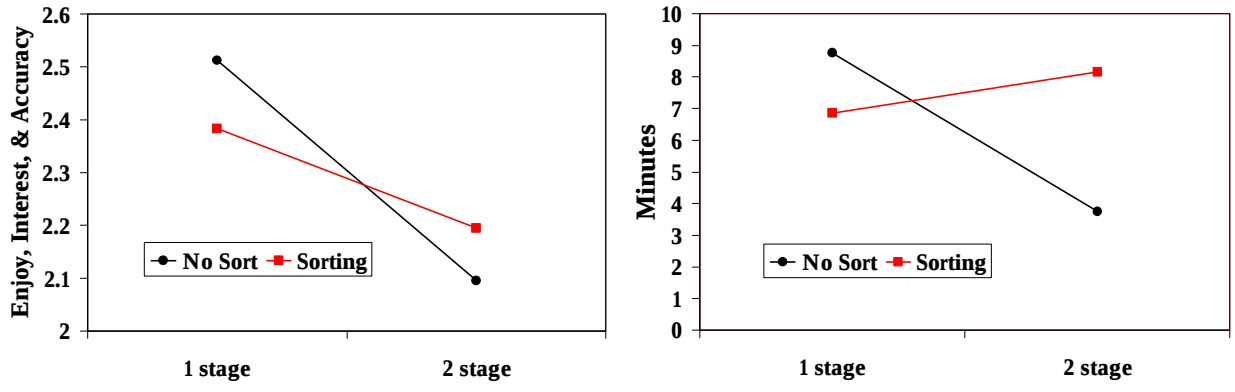
Cells 5 and 6 enable us to examine two additional issues. Cell 5 is a no-sort, two-stage task as in Cell 1, but with 16 profiles rather than 32 profiles. This comparison is interesting because dominant theories suggest that respondents are more likely to use non-compensatory processes when the number of profiles is large (e.g., Bettman, Luce, and Payne 1998). In Cell 6 (pruning), respondents complete a no-sort, full-rank task as in Cell 2, but, as they make choices,

we remove any profile that is not consistent with an additive partworth model.[14] We include this task to test whether or not it affects respondents' evaluative processes. Cell 6 also represents an alternative way to reduce the respondents' task while obtaining a full rank order. We make no predictions as to whether pruning will increase or decrease the use of non-compensatory processes or whether it will increase or decrease predictive accuracy (enforced consistency vs. response-error propagation). We begin with Cells 1, 2, 3, and 4 and return to Cells 5 and 6 later.

Basic Experimental Results

Ease of use. Before examining fit and predictions, we examine how respondents perceive a two-stage consideration-then-choice task. We hypothesized that a two-stage process would be perceived as more enjoyable, more accurate, and more engaging. We also expected that it would save time. The results are reported in Figures 5a and 5b. We have oriented both axes such that down is better. We see first that, in the base condition of no sorting, the two-stage task is seen as significantly more enjoyable, accurate, and engaging.[15] Sorting increases enjoyment, accuracy, and interest for the one-stage task but not for the two-stage task, however neither difference is significant.[16] For the no-sort cell, two-stage data collection saves substantial time (3.75 minutes compared to 8.75 minutes, $t = 2.8$, $p = 0.01$). Such time savings could reduce data collection costs substantially. More respondents completed the no-sort, two-stage task, 94%, than the no-sort one-stage task, 86%, although the difference is marginally significant ($t = 1.7$, $p = 0.10$).

Figure 5
Task Difficulty (less is better on both graphs)



(a) Attitudes Toward Task Difficulty (b) Times to Complete the Tasks

[14] Cell 6 has fewer respondents because we improved the profile pruning algorithm during data collection. Any respondents who answered the survey prior to the improvement in the algorithm were removed from the study.
[15] Comparing one-stage to two-stage, $t = 2.2$, $p = 0.03$ for no-sort cells; $t = 0.9$, $p = 0.37$ for the sort cells.
[16] Comparing sort to no-sort, $t = 0.8$, $p = 0.45$ for the two-stage cells; $t = 0.6$, $p = 0.53$ for the one-stage cells.

However, contrary to expectations, sorting did not reinforce these attitude, time, and completion-rate advantages. In fact, sorting appears to mitigate some advantages.[17] It appears that sorting was effortful and was perceived as such.

Use of non-compensatory processes. We next examine classifications of respondents as using either a compensatory or a non-compensatory process. We also test whether sorting induced the use of a non-compensatory process. To address these questions, we examine the number of respondents for whom the Greedoid fit was better than the compensatory fit (AC(4)).[18] The results strongly suggest that respondents were using non-compensatory consideration and choice processes. For the one-stage rank task, 99% of the respondents were classified as using a non-compensatory process. This decreased to 89% for the two-stage, consideration-then-rank task; the remainder were classified as tied (10%) or compensatory (1%). Our results do not necessarily signal that there is less LBA processing in the consideration task, but rather lower precision in the two-stage task – two-stage cells are based on fewer observations than one-stage cells.[19] Overall, the vast majority of respondents appear to use a non-compensatory process.

Interestingly, counter to expectations, sorting did not induce a lexicographic process. There were virtually no differences between the no-sort and sort cells – 99% vs. 100% for the one-stage task and 92% vs. 87% for the two-stage task.[20] Nor did sorting induce respondents to consider substantially more profiles – on average 6.4 profiles were considered in the no-sort cell and 6.7 profiles in the sort cell (t = 0.4, p = 0.66).

Fit and prediction. We examine the predictive ability of the Greedoid method in many ways. We first compare the fit and prediction of a non-compensatory (LBA) process to a compensatory process (AC(4)) for all respondents in Cells 1-4. We later relax the restrictions on the compensatory model and consider AC(∞). On average, LBA fits significantly better and predicts holdouts significantly better on the violated-pairs metric. This performance also holds when we use a metric, hit rate, that was not used in estimation. See Table 1. We also considered a model

[17] The two- vs. one-stage time difference is not significant for the sort cells (t = 0.4, p = 0.70). The no-sort vs. sort time difference is not significant for the one-stage cells (t = 1.3, p = 0.19), but is significant at the 0.10 level for the two-stage cells (t = 1.8, p = 0.07). No other completion rate differences were significant.

[18] In Section 8 we address the question of mixed processes, e.g., non-compensatory consideration and compensatory choice processes.

[19] The two-stage cells yield, on average, only 40% of the pairs observed in one-stage cells. The average respondent considered about 7 profiles yielding (6x7/2) + (7x(32-7) = 196 pairs. A full rank yields (31x32/2) = 496 pairs.

[20] Neither difference is significant. Because of the small numbers of observations in the compensatory cells, we used the Fisher exact test, p = 0.50 for the one-stage cells and p = 0.21 for the two-stage cells.

in which we assign respondents to their best-fit model. This mixed model does better on both fit and holdout predictions, however, none are significantly better because so few respondents were assigned to a compensatory model.

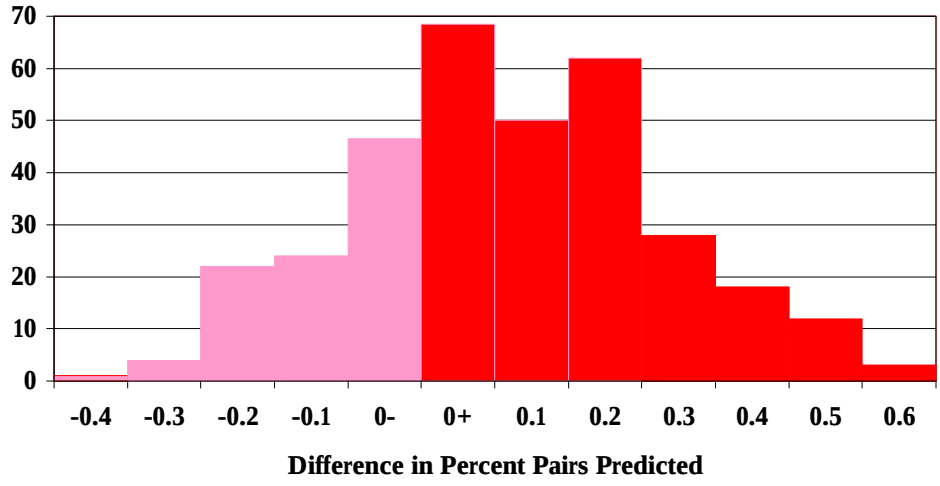
Table 1
Comparison of Fit and Prediction

	Lexicographic by Aspects	Compensatory	Best of L, C
Fit (percent pairs)	0.955*	0.697	0.956*
Holdout (percent pairs)	0.745*	0.626	0.749*
Holdout hit rate	0.597*	0.382	0.602*

* Significantly better than compensatory at the 0.05 level

We next examine the data by individual respondent. Figure 6 portrays the ability of LBA and AC(4) to predict holdout pairs for each respondent. In particular, we plot a histogram of respondents versus the difference in predictive ability between the two models. Positive numbers (shown in dark red) indicate the respondents for whom LBA predicts best. The mass of the distribution favors a non-compensatory process.

Figure 6
Histogram of Comparative Predictive Ability



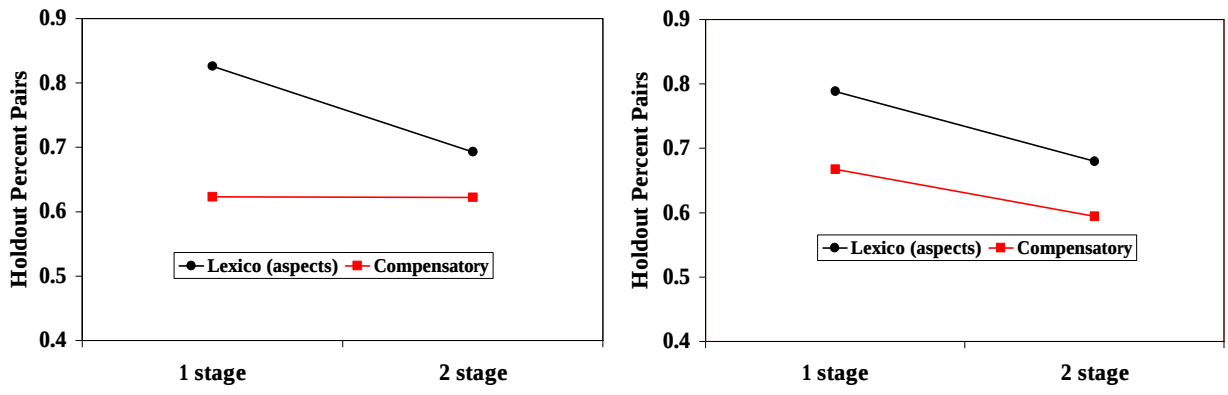
Because classification is based on fit, we examine holdout predictions by respondent classes. On holdout sets, LBA does significantly better than AC(4) for respondents classified as

lexicographic. The reverse is true for respondents classified as compensatory. The fit classification seems to apply to holdout predictions.

We next examine holdout predictions by experimental cell. See Figure 7. (Calibration fit and hold-out hit rates provide similar qualitative interpretations.) In all four cells, the non-compensatory model predicts significantly better than the compensatory model.[21]

In most cases, the predictive ability decreases when a two-stage rather than a one-stage task is used, most likely because there are fewer paired comparisons with which to fit the model.[22] On the other hand, sorting does not appear to have a significant effect on predictive ability.[23]

Figure 7
Predictive Ability by Cell, Non-compensatory vs. Compensatory Processes



(a) No Sorting Allowed (b) Sorting Allowed

<u>Summary of Basic Experimental Results</u>. The basic experiments suggest that:

- most respondents use non-compensatory processes (for SmartPhones),
- a non-compensatory model fits and predicts better than a compensatory model,
- fit classifications appear to apply for holdout predictions,
- the basic two-stage task is much faster, is perceived much better by respondents, and is completed in higher percentages than the one-stage full-rank task, however,

[21] t = 3.3, 8.1, 3.9, and 8.0, respectively, for Cells 1-4. All significant at 0.001 or greater.
[22] Comparing two-stage to one-stage cells, for LBA, t = 5.8, p = 0.00 for the no-sort cells and t = 4.4, p = 0.00 for the sort cells. For AC(4), t = 0.02, p = 0.98 for the no-sort cells and t = 2.6, p = 0.01 for the sort cells.
[23] Comparing no-sort to sort, for LBA, t = 1.4, p = 0.16 for one-stage and t = 0.6, p = 0.57 for two-stage. For AC(4), t = 1.6, p = 0.14 for one-stage and t = 1.0, p = 0.33 for two-stage.

  - the two-stage task yields fewer paired-comparison observations which leads to reduced predictive ability and a reduced ability to classify respondents – researchers must make cost-benefit tradeoffs in deciding among tasks,
  - the time, attitude, and completion-rate advantages of a two-stage task are mitigated if the respondent is allowed to sort profiles,
- allowing respondents to sort profiles does not induce lexicographic processing and has little impact on predictive ability.

7. Using Greedoids to Study Behavior: Two Experiments

One of our goals was to develop an inference method that could be used to study, in a more-or-less natural task, how environmental variables influence respondents' consideration-and-choice processes. Using Cells 1 and 2 of the basic experiment as baselines, we analyze two additional experiments as implemented by Cells 5 and 6. (Review Figure 4.) The first experiment replicates one of the most accepted findings of constructive processes – more choice alternatives induce greater use of non-compensatory processes. The second experiment is new – we modify the choice sets adaptively to be consistent with additive processing rules.

The Impact of More Profiles on Respondents' Consideration-and-Choice Processes

Behavioral theory suggests that respondents are more likely to use a non-compensatory process if they are asked to evaluate more profiles (e.g., Bettman, Luce, and Payne 1998; Bröder 2000; Johnson, Meyer and Ghose 1989; Lohse and Johnson 1996; Payne, Bettman and Johnson 1993). Behavioral theory also suggests that respondents will include fewer profiles in their consideration sets, in part, because fewer are likely to satisfy their evaluative criteria. However, the theory is mute on the number of profiles because that number will depend upon the specific features of the profiles, the specific respondent criteria, and the unobserved cost of evaluation (e.g., Hauser and Wernerfelt 1990, Roberts and Lattin 1991).

To test these theories we compare Cells 1 and 5. In both cells respondents completed a two-stage task with no sorting allowed, but in Cell 1 they evaluated 32 profiles and in Cell 5 they evaluated 16 profiles. In both cells the profiles were chosen from (different) fractional factorials of the $4^3 2^4$ design. Consistent with behavioral theory we identified proportionally more respondents as using non-compensatory processes in the cell with more profiles – 92% for 32 profiles and 72% for 16 profiles ($t = 2.5$, $p = 0.02$). Although the number of profiles reduced the use of non-compensatory processes, non-compensatory processes remained in the majority for the 16-

profile cell. Also consistent with theory, respondents considered, on average, fewer profiles in the 16-profile cells (4.4 vs. 6.4, $t = 4.6$, $p = 0.00$).

Although it was not our goal to test the predictive ability of the models in the two cells, it is interesting to examine the impact of the number of profiles on the tradeoff between task difficulty and predictive ability. Interestingly, there was a significant increase in enjoyment and interest ($t = 2.0$, $p = 0.05$), but no significant decrease in task time (3.4 vs. 3.75 minutes, $t = 0.5$, $p = 0.64$). There was also no significant reduction in predictive ability for either non-compensatory ($t = 0.03$, $p = 0.98$) or compensatory models ($t = 0.1$, $p = 0.89$). Perhaps 16 profiles provide sufficient information. These data suggest that, for the consideration-then-choice task, the researcher who is interested in predictive ability is free to choose the number of profiles (16 vs. 32) that appears more realistic to the respondent.

The Impact of Pruning on Respondents' Consideration-and-Choice Processes

In our second experiment we examine whether pruning profiles to be consistent with an additive process would cause respondents to appear more or less lexicographic. Recall that only a tiny fraction (10^{-23}) of all possible profile orders are consistent with a pure lexicographic process, that lexicographic processes can be represented by extremely-chosen additive partworths, and that only $1/10^{th}$ of 1% of all profile orders are consistent with an additive-partworth representation.

Profiles are pruned as follows. Respondents are <u>not</u> constrained in their first choice. However, each respondent's first choice imposes a set of linear constraints on that respondent's partworths. For some of the remaining profiles there may be no values of the partworths that could predict those profiles as the respondent's second choice. We remove those profiles from the respondent's second choice set. These profiles may reappear in later choice sets. We continue this pruning process until we know the rank of all profiles.

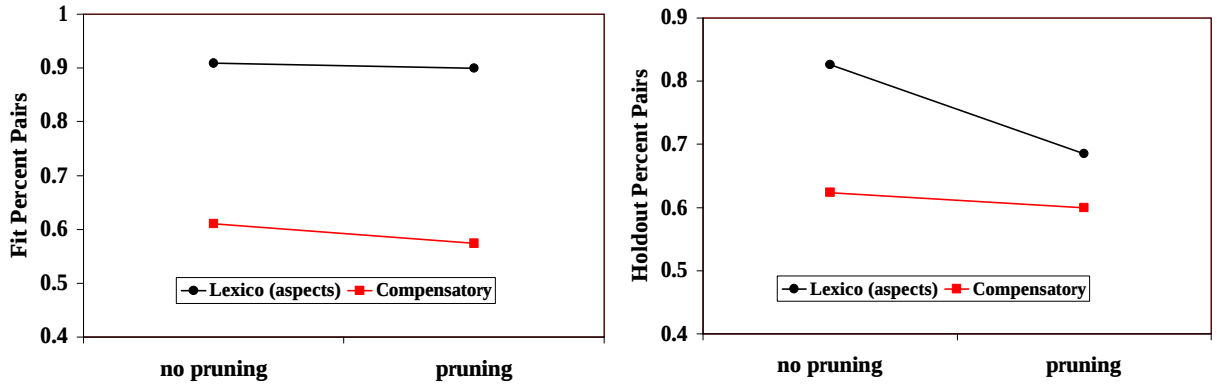
We implement pruning by solving a linear-programming feasibility problem for each unranked profile and we remove those profiles for which there is no feasible solution for additive partworths. The linear program runs sufficiently fast; respondents notice no undue delay. The savings from pruning can be substantial. In our experiments respondents saw, on average, 73% fewer profiles, 32% fewer choice sets, completed the task in 5.9 rather than 8.8 minutes ($t = 1.6$, $p = 0.12$), and saw the task as more enjoyable and interesting, albeit not significantly so ($t = 1.7$, $p = 0.10$). Predictions about predictive ability are mixed – pruning is an easier task (+), helps re-

spondents avoid "errors" (+), is sensitive to errors in early choices (–), and provides less data with which to estimate respondents' choice processes or partworths (–).

In our experiment, Cell 6 differs from Cell 2 only on pruning – both are full-rank, no-sort cells. Because 99% of the respondents were identified as non-compensatory in Cell 2 and 100% were identified as non-compensatory in Cell 6, we can not distinguish the cells on this measure (p = 0.6 with Fisher's exact test). We can, however, examine the impact of pruning on fit and prediction. As indicated in Figure 8, pruning had no significant effect on fit or the predictive ability of the compensatory model, but did decrease significantly the predictive ability of the non-compensatory model (LBA, 83% without pruning, 69% with pruning, t = 4.8, p = 0.00).

The results are interesting. In traditional conjoint analysis, which assumes an additive partworth model, pruning saves considerable time and effort with little loss of predictive ability. However, this loss is only significant when the respondents use non-compensatory processes (as identified by Greedoid methods). Phrased another way, pruning is a promising method for full-profile conjoint analysis in product categories that are likely to be described by compensatory features (or aspects). Pruning does not appear promising if the features are non-compensatory.

Figure 8
The Impact of Pruning on Fit and Prediction



(a) Fit – Percent Pairs Predicted Correctly (b) Holdout Predictions

Summary – Two Experiments

The two additional experiments demonstrate how Greedoid methods can be used to study the impact of the choice environment on constructed consideration and choice processes. The number-of-profiles experiment suggests that the Greedoid method is consistent with established theory and, hence, has face validity. The advantage of the Greedoid method is that it can infer consumer decision processes directly from the decisions that consumers make – we do not need

verbal protocols or information display boards that might, themselves, influence consumer processing strategies. However, as this experiment itself demonstrates, researcher decisions, such as the number of profiles, can influence the decision processes that consumers construct.

The second experiment investigates pruning – an induction that we have not seen in the literature. In our data, pruning does not appear to discourage lexicographic processing. It might be a viable method to reduce respondent burden if the features of the profiles are compensatory, however, it appears to hurt predictions if the features are non-compensatory. This phenomenon bears further study because most adaptive questioning algorithms select questions based on an assumed linear model (Abernethy and Evgeniou 2004; Johnson 1991; Toubia, et. al. 2004).

8. Alternative Models of Consideration or Choice Processes

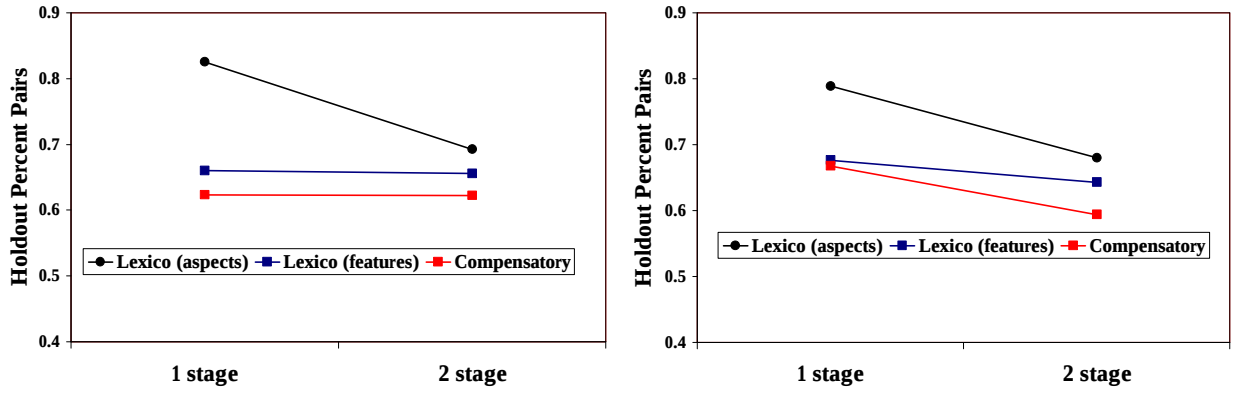
In this section we analyze two additional models of consideration-and-choice processes – lexicographic-by-features (LBF) and an unconstrained linear process ((AC(∞)). Other processes, such as elimination-by-aspects (EBA), acceptance-by-aspects (ABA), and alternative constraints on compensatory processes (AC(2) and AC(10)), provide no surprises. Details are available from the authors.

Lexicographic-By-Features (LBF)

We implement LBF by constraining aspects within a feature to appear together in the aspect ordering of LBA. Because LBF is a special case of LBA, the fit statistics will be lower for LBF by the principle of optimality. However, if respondents are really processing profiles by features rather than aspects, an LBF model may be more robust for holdout data. The results are shown in Figure 9. The holdout predictions are significantly lower for LBF in most cases suggesting that the respondents are processing profiles by aspects rather than features.[24] For comparison, we repeated the compensatory predictions in Figure 9. LBF predicts better than a compensatory model (AC(4)), but the differences are not significant.[25] EBA and ABA (not shown) are close to LBA and do better than LBF in Figure 9.

[24] t = 1.8, 7.1, 2.3, and 4.5 and p = 0.07, 0.00, 0.02, and 0.00, respectively, for Cells 1-4.
[25] t = 1.2, 1.1, 1.8, and 0.3 and p = 0.24, 0.29, 0.07, and 0.74, respectively, for Cells 1-4.

Figure 9
Lexicographic-by-Features vs. Lexicographic-by-Aspects

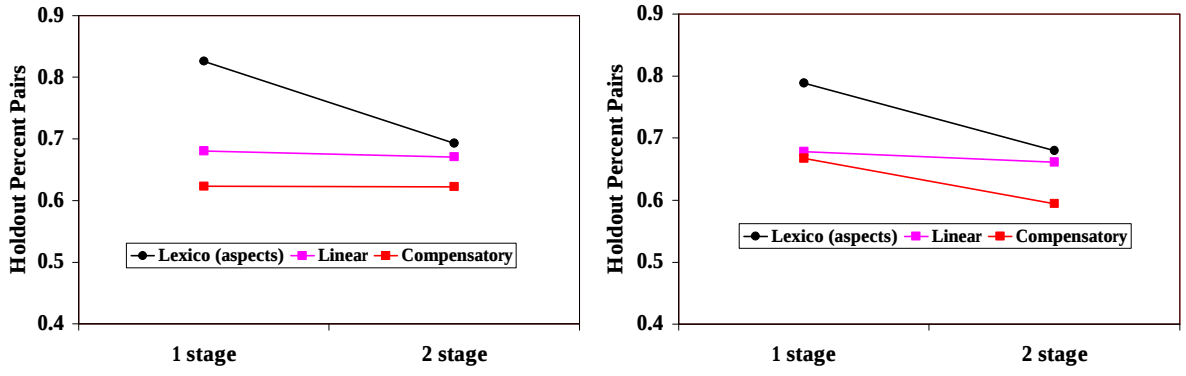


(a) No Sorting Allowed (b) Sorting Allowed

Unconstrained Linear Models

In theory, unconstrained linear models (AC(∞)) can represent lexicographic processes. However, these representations are not unique and the extreme nature of lexico-consistent part-worths may make it difficult statistically to obtain empirical estimates.[26] On the other hand, the Greedoid representation is simple and, potentially, robust. Simple, robust representations often predict well. This is the case in our data, Figure 10. The Greedoid representations out-predict unconstrained linear models, significantly so for the one-stage cells.[27] AC(10) is bounded by the lexicographic and linear models in Figure 10. AC(2) does not predict as well as AC(4).

Figure 10
Estimates from an Unconstrained Linear Model



(a) No Sorting Allowed (b) Sorting Allowed

[26] For example, the representation is not unique: $2^{m(1-n)}$, works for all $m \geq 1$. The set of lexico-consistent partworths is not convex, but has positive measure (uncountably infinite).
[27] The differences are significant for the one-stage cells (t = 7.4, 6.9), but not for the two-stage cells (t = 1.4, 1.1).

Mixed Models

In the two-stage models analyzed thus far, we assumed that respondents use the same processing strategy for both the consideration and choice stages. Some researchers suggest that consumers use a non-compensatory decision process for the consideration stage and a compensatory decision process for the choice stage (e.g., Payne 1976). Greedoid methods can be used in either stage. However, in our data, the small (natural) size of the consideration sets does not yield sufficiently many ranked pairs to obtain precise estimates for choice within the consideration set. Even when we parse the fit statistics by stage, LBA fits significantly better than AC(4) in both stages. We leave full exploration of mixed models to future data collection in which more information is collected about choices within consideration sets.

9. Managerial Implications

The average consideration set for our sample was approximately 6-7 profiles. Reducing a set of 32 profiles to 6-7 profiles requires that a consumer process relatively few aspects lexicographically. Table 7 lists the six aspects that were used most often and indicates whether they tended to be used to accept profiles or reject profiles (second column), the percent of consumers who used that aspect as one of the first three aspects in a lexicographic order (third column), and the percent who used that aspect as the first aspect in a lexicographic order (fourth column).[28]

It is interesting that, in our data, price aspects were often, but not always, rejection criteria, while all other aspects were acceptance criteria. (This is true for aspects not shown in Table 2.) We do not know if this generalizes to other categories. Furthermore, although "high-price" was the top lexicographic aspect in our study, this may be a consequence of the category or our student sample. We do not expect price to be the top lexicographic aspect in all categories nor do we feel that this result affected the basic scientific and methodological findings about lexicographic processing or predictive ability.

Table 7 has a number of implications. Firstly, for our student sample, there are clear price segments – almost half the sample rejected high-priced SmartPhones. Secondly, the aspects of "flip" and "small" are each used by about 30% of the respondents. For this sample, any manufacturer will lose considerable market share if it does not include SmartPhones that are small and flip. The keyboard aspect is interesting. One segment, 17.3%, seem to require a keyboard and another segment, 7.5% (not shown), seem to reject all keyboard SmartPhones. On this

[28] Column 3 sums to 300% across all aspects.

aspect, a manufacturer would be best advised to offer both SmartPhones with keyboards and SmartPhones without keyboards. Finally, brand, service provider, and operating system are not high in the lexicographic ordering. They are more likely to be more compensatory.

Table 2
Top Lexicographic Aspects for SmartPhones (for our sample)

Aspect	ABA or EBA	Affect Consideration	Top Aspect
Price – $499	Reject	49.2%	26.1%
Flip	Accept	32.0%	10.4%
Small	Accept	29.4%	10.0%
Price – $299	Reject	19.8%	4.2%
Keyboard	Accept	17.3%	7.5%
Price – $99	Accept	14.5%	4.8%

10. Summary, Conclusions, and Future Research

In this paper we propose a method to test whether respondents use non-compensatory processes and, if so, to infer the details of those processes from either consideration-then-choice or full-rank tasks. Inference is a non-trivial combinatorial problem which has hitherto been too time-consuming to solve Greedoids provide a structure and theory to transform this $N!$ problem into a 2^N problem, which, for practical problems, decreases running time by a factor the order of 10^9. Monte Carlo simulation experiments suggest that it is feasible to infer, albeit with noise, the process that respondents use to evaluate profiles.

We tested Greedoid methods empirically by asking respondents to evaluate SmartPhones. The data suggest that the Greedoid methods can identify the respondents' processing strategies and that the Greedoid estimates can predict better than compensatory models developed from the same data. We demonstrate how Greedoid methods might be used to study constructed judgment and decision processes by replicating one well-known result and by investigating a new induction.

Greedoid methods also provide a practical conjoint-like method to forecast consumer response to new products or changes in the features of existing products. This method is particularly useful in product categories where consumers are presented with large numbers of potential choices. Not only do the Greedoid estimates appear to predict better (in most cases) than com-

pensatory estimates, but the consideration-then-choice task is perceived by respondents as more enjoyable, more accurate, and more interesting. It also saves time and increases completion rates – both of which translate directly into cost savings. While combining estimation method and task still requires careful tradeoffs by the researcher, the new methods (combined with a simpler task) provide greater capabilities.

In summary, based on the simulations and the SmartPhone data:

Methodological

- it is feasible to infer non-compensatory processes with a Greedoid dynamic program,
- a consideration-then-choice task reduces task time, increases completion rates, and improves perceived enjoyment, accuracy, and interest, however, at some loss in predictive ability,
- enabling respondents to sort profiles by aspects is seen as more difficult and time-consuming, but does not increase predictive ability,
- non-compensatory estimates predict holdouts better than compensatory estimates,
- Greedoid methods appear to be more robust – they predict better than additive models even though additive models can represent lexicographic processes.

Consumer Behavior

- for SmartPhones, most respondents appear to use non-compensatory processes,
- respondents appear to process <u>aspects</u> rather than <u>features</u> lexicographically,
- enabling respondents to sort profiles <u>does not</u> increase their tendency to use non-compensatory processes,
- increasing the number of profiles <u>does</u> increase respondents' tendency to use non-compensatory processes,
- pruning profiles to force consistency with an additive representation <u>does not</u> affect the ability to fit the process <u>nor does</u> it change the ability of compensatory estimates to predict holdouts, but it <u>does</u> decrease the ability of non-compensatory estimates to predict holdouts,

Managerial (for our sample and category)

- Price is often used as a rejection criterion. Non-price aspects seem to be used as acceptance criteria.

- “Small” and “Flip” are key acceptance criteria for SmartPhones.

Future Directions

Greedoid methods provide a powerful tool to study consumer behavior. Although we used rank and consideration-then-rank tasks, Greedoid methods can be extended to the stated-choice task that is popular in choice-based conjoint analysis. We might also investigate two-stage processing rules that are non-compensatory for consideration and compensatory for choice. Researchers can use the Greedoid inference engine to investigate many impacts of consumers’ constructive judgment and decision processes – manipulations that might be too intrusive if implemented by verbal protocols or information display tasks.

Methodologically, the exact dynamic program is still exponential, 2^N, in the number of aspects. We handled 16 aspects in 1.85 seconds. At this rate Greedoid inference can be used to evaluate up to 21 aspects in under a minute. However, we can handle much larger problems if we seek to concentrate on the first few lexicographic aspects in a respondents LBA process. Because the theory applies to partial orders, we can stop the (forward induction) dynamic program after M aspects yielding a running time proportional to ${}_N\mathrm{C}_M$. For example, we could identify the top five out of fifty aspects in approximately one minute. Partial-order Greedoid inference can be used to identify satisficing processes in which some aspect levels are considered as equivalent by respondents. Other heuristics might also be used (Kohli and Jedidi 2004; Kohli, Krishnamurthi, and Jedidi 2003).

Finally, SmartPhones were chosen as a first test because the dominant features were likely to be non-compensatory for some respondents. It will be interesting to see how Greedoid methods perform in product categories that have features that are more compensatory in nature.

References

Abernethy, Jacob, Theodoros Evgeniou, and Jean-Philippe Vert (2004), "An Optimization Framework for Adaptive Questionnaire Design, INSEAD, Fontainebleau, FR.

Batra, Rajeev, John G. Myers, and David A. Aaker (1996), *Advertising Management, 5E*, (Englewood Cliffs, NJ: Prentice-Hall, Inc.).

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25, 3, (December), 187-217.

----- and L. W. Park (1980), "Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Processes: A Protocol Analysis," *Journal of Consumer Research*, 7, 234-248.

Bröder, Arndt (2000), "Assessing the Empirical Validity of the "Take the Best" Heuristic as a Model of Human Probabilistic Inference," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 5, 1332-1346.

Chase, Valerie M., Ralph Hertwig, and Gerd Gigerenzer (1998), "Visions of Rationality," *Trends in Cognitive Science*s, 2, 6, (June), 206-214.

Dawes, R. M. (1979), "The Robust Beauty of Improper Linear Models in Decision Making," *American Psychologist*, 34, 571-582.

Dawkins, Richard (1998), *Unweaving the Rainbow: Science, Delusion, and the Appetite for Wonder*, (Boston, MA: Houghton Mifflin Company).

DeSarbo, Wayne S., Donald R. Lehmann, Greg Carpenter, and I. Sinha (1996), "A Stochastic Multidimensional Unfolding Approach for Representing Phased Decision Outcomes, " *Psychometrica*, 61 (September), 485-508.

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

----- and Robin M. Hogarth (1981), "Behavioral Decision Theory: Processes of Judgment and Choice," *Annual Review of Psychology*, 32, 52-88.

Fishburn, Peter C. (1974), "Lexicographic Orders, Utilities and Decision Rules: A Survey," *Management Science*, 20, 11, (Theory, July), 1442-1471.

Frederick, Shane (2004), "Cognitive Impulsivity and Decision Making." Working Paper, MIT Sloan School of Management, Cambridge, MA 02142. Under review *Journal of Economic Perspectives*.

Gensch, Dennis H. (1987), "A Two-stage Disaggregate Attribute Choice Model," *Marketing Science*, 6, (Summer), 223-231.

----- and Ehsan S. Soofi (1995), "Information-Theoretic Estimation of Individual Consideration Sets," *International Journal of Research in Marketing*, 12, (May), 25-38.

Gigerenzer, Gerd and Daniel G. Goldstein (1996), “Reasoning the Fast and Frugal Way: Models of Bounded Rationality,” *Psychological Review*, 1003, 4, 650-669.

Gilbride, Timothy and Greg M. Allenby (2004), “A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules,” *Marketing Science*, 23, 3, (Summer), 391-406.

Green, Paul E., Abba M. Krieger, and Pradeep Bansal (1988), “Completely Unacceptable Levels in Conjoint Analysis: A Cautionary Note,” *Journal of Marketing Research*, 25, (August), 293-300.

Griffin, Abbie and John R. Hauser (1993), "The Voice of the Customer," *Marketing Science*, vol. 12, No. 1, (Winter), 1-27.

Hagerty, M. R., and V. Srinivasan (1991), “Comparing the Predictive Powers of Alternative Multiple-Regression Models,” *Psychometrika*, 56, 1, (March), 77-85.

Hauser, John R. (1978), "Testing the Accuracy, Usefulness and Significance of Probabilistic Models: An Information Theoretic Approach," *Operations Research*, Vol. 26, No. 3, (May-June), 406-421.

----- and Birger Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, Vol. 16, (March), 393-408.

Held, M. and R. M. Karp (1962), “A Dynamic Programming Approach to Sequencing Problems,” *SIAM Journal of Applied Mathematics*, 10, 196-210.

Jedidi, Kamel and Rajeev Kohli (2004), “Probabilistic Subset-Conjunctive Models for Heterogeneous Consumers,” Working Paper, Graduate School of Business, Columbia University (November).

-----, -----, and Wayne S. DeSarbo (1996), “Consideration Sets in Conjoint Analysis,” *Journal of Marketing Research*, 33, (August), 364-372.

Johnson, Eric J. and Robert J. Meyer (1984), “Compensatory Choice Models of Noncompensatory Processes: The Effect of Varying Context,” *Journal of Consumer Research*, 11, 1, (June), 528-541.

-----, -----, and Sanjoy Ghose (1989), “When Choice Models Fail: Compensatory Models in Negatively Correlated Environments,” *Journal of Marketing Research*, 26, (August), 255-290.

Johnson, Richard (1991), “Comment on `Adaptive Conjoint Analysis: Some Caveats and Suggestions,” *Journal of Marketing Research*, 28, (May), 223-225.

Kendall Maurice G. (1975), *Rank Correlation Methods*, (London, England: Charles Griffin & Company, Ltd.)

Kim, Jin Gyo (2204), “Dynamic Heterogeneous Choice Heuristics: A Bayesian Hidden Markov Mixture Model Approach,” Working Paper, MIT Sloan School of Management, Cambridge, MA.

Klein, Noreen M. (1988), “Assessing Unacceptable Attribute Levels in Conjoint Analysis,” *Advances in Consumer Research* vol. XIV, pp. 154-158.

Kohli, Rajeev and Kamel Jedidi (2004), “Representation and Inference of Lexicographic Preference Models and Their Variants,” Working Paper, Columbia University, New York, NY, December.

-----, Ramesh Krishnamurthi, and Kamel Jedidi (2003), "Subset-Conjunctive Rules for Breast Cancer Diagnosis," Graduate School of Business, Columbia University, (May).

Korte, B. and L. Lovasz (1985), "Basis Graphs of Greedoid and Two-connectivity," *Mathematical Programming Study*, 24, 158-165.

Lohse, Gerald J. and Eric J. Johnson (1996), "A Comparison of Two Process Tracing Methods for Choice Tasks," *Organizational Behavior and Human Decision Processes*, 68, 1, (October), 28-43.

Luce, Mary Frances, John W. Payne, and James R. Bettman (1999), "Emotional Trade-off Difficulty and Choice," *Journal of Marketing Research*, 36, 143-159.

Malhotra, Naresh (1986), "An Approach to the Measurement of Consumer Preferences Using Limited Information," *Journal of Marketing Research*, 23, (February), 33-40.

Martignon, Laura and Ulrich Hoffrage (2002), "Fast, Frugal, and Fit: Simple Heuristics for Paired Comparisons," *Theory and Decision*, 52, 29-71.

Mitchell, Tom M. (1997), *Machine Learning*, (Boston, MA: WCB McGraw-Hill).

Montgomery, H. and O. Svenson (1976), "On Decision Rules and Information Processing Strategies for Choices among Multiattribute Alternatives," *Scandinavian Journal of Psychology*, 17, 283-291.

Nakamura, Yutaka (2002), "Lexicographic Quasilinear Utility," *Journal of Mathematical Economics*, 37, 157-178.

Niño-Mora, José (2000), "On Certain Greedoid Polyhedra, Partially Indexable Scheduling Problems, and Extended Restless Bandit Allocation Indices," Economics Working Papers 456, Department of Economics and Business, Universitat Pompeu Fabra.

Ogilvy, David (1985), *Ogilvy on Advertising*, (New York, NY: Vintage Books).

Olshavsky, Richard W. and Franklin Acito (1980), "An Information Processing Probe into Conjoint Analysis," *Decision Sciences*, 11, (July), 451-470.

Oppewal, Harmen, Jordan J. Louviere, and Harry J. P. Timmermans (1994), "Modeling Hierarchical Conjoint Processes with Integrated Choice Experiments," *Journal of Marketing Research*, 31, (February), 92-105.

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

-----, James R. Bettman, and Eric J. Johnson (1988), "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 534-552.

-----, -----, and Eric J. Johnson (1993), *The Adaptive Decision Maker*, (Cambridge, UK: Cambridge University Press).

Reeves, Rosser (1961), *Reality in Advertising*, (New York, NY: Knopf Publishing).

Roberts, John H. and James M. Lattin (1991), “Development and Testing of a Model of Consideration Set Composition,” *Journal of Marketing Research*, 28, (November), 429-440.

Rossi, Peter E and Greg M. Allenby (2003), “Bayesian Statistics and Marketing,” *Marketing Science*, 22, 3, (Summer), 304-328.

Shugan, Steven M. (1980), “The Cost of Thinking,” *Journal of Consumer Research*, 7, 2, (September), 99-111.

Srinivasan V. and Chan Su Park (1997), “Surprising Robustness of the Self-Explicated Approach to Customer Preference Structure Measurement,” *Journal of Marketing Research*, 34, (May), 286-291.

----- and Allan Shocker (1973), “Linear programming Techniques for Multidimensional Analysis of Preferences, *Psychometrika*, 38, 3, (September), 337-369.

----- and Gordon A. Wyner (1988), “Casemap: Computer-Assisted Self-Explication of Multiattributed Preferences,” in W. Henry, M. Menasco, and K. Takada, Eds, *Handbook on New Product Development and Testing*, (Lexington, MA: D. C. Heath), 91-112.

Swait, Joffre (2001), “A Noncompensatory Choice Model Incorporating Cutoffs,” *Transportation Research*, 35, (Part B), 903-928.

Thorngate, W. (1980), “Efficient Decision Heuristics,” *Behavioral Science*, 25, (May), 219-225.

Toubia, Olivier, John R. Hauser, and Duncan Simester (2004), “Polyhedral Methods for Adaptive Choice-based Conjoint Analysis,” *Journal of Marketing Research*, 41, 1, (February), 116-131.

Tversky, Amos (1972), "Elimination by aspects: A theory of choice," *Psychological Review*, 79, 281-299.

Urban, Glen L. and John R. Hauser (2004), “’Listening-In’ to Find and Explore New Combinations of Customer Needs,” *Journal of Marketing*, 68, (April), 72-87.

Zaltman, Gerald (1997), “Rethinking Market Research: Putting People Back In,” *Journal of Marketing Research*, 23, (November), 424-437.

Appendix: Proofs of the Formal Propositions

Proposition 1. *Let E be the set of aspects, and let X be a partial order on the profiles. Let G be the collection of ordered subsets of E that are lexico-consistent with X. Then G is a Greedoid language.*

Proof. We show that Greedoid properties (2) and (3) hold for collection *G*. Property (1) is implied by (2). Property (2): Lexico consistent means that there is no pair of profiles *P* and P' with $P \succ_L P'$ and $P' \succ_X P$. So, if *L* is consistent with *X*, then *L**e* is consistent with *X* since the relations with respect to *L**e* are a subset of the relations with respect to *L*. Property (3): Let *e* be the first aspect in L' such that $e \notin L$. Such an *e* is guaranteed to exist since $|L'| > |L|$. We show that (*L*, *e*) ∈ *G* via a contradiction. Suppose that there are profiles *P* and P' with $P \succ_{L,e} P'$ and $P' \succ_X P$. Since *L* and *X* are consistent, it follows that *P* and P' are unrelated with respect to *L*, and thus $P \succ_e P'$. Let L'' be aspects in L' prior to *e*. Then $L'' \subseteq L$, so *P* and P' are unrelated in L''. It follows that $P \succ_{L'',e} P'$ and thus $P \succ_L P'$, contradicting that L' is consistent with *X*. We conclude that Property 3 is true.

Proposition 2. *Let L and L' be two different permutations of the subset E' of aspects, and let e be any aspect not in E'. Then the number of inconsistencies directly caused by e in (L, e) is the same as the number of inconsistencies caused by e in (L', e).*

Proof. Suppose that for profiles *P* and P', $P \succ_X P'$. The aspect *e* causes an inconsistency with respect to the profiles *P* and P' in (*L*, *e*) if and only if the following conditions hold: (i) profiles *P* and P' are undifferentiated by the aspects in *L*, and (ii) $P' \succ_e P$. These are the same conditions under which *e* causes an inconsistency with respect to *P* and P' in (*L'*, *e*).

Proposition 3. *Suppose that L is an ordering of a subset of aspects that is consistent with the preferences of X. Then there is an optimal ordering of aspects that begins with the order L.*

Proof. Suppose that L' is an optimal ordering of aspects; that is, it is the one that minimizes the number of inconsistencies with respect to *X*. Let L'' be obtained from L' by moving the aspects of *L* to the front of the order. We will show that any inconsistency with respect to L'' is also an inconsistency with respect to L', thus showing that L'' is at least as optimal as L'. Suppose that for profiles *P* and P', $P \succ_X P'$, and $P' \succ_{L''} P$. Let *e* be the first aspect in L'' that differentiates *P* and P'. Since *L* is consistent with *X*, it follows that $e \notin L$. Therefore, *e* is also the first aspect in L' that differentiates *P* and P', and so $P' \succ_{L'} P$. It follows that *P* and P' also cause an inconsistency with respect to L', proving the proposition.