A TRUTH-TELLING MECHANISM FOR CONJOINT ANALYSIS

Min Ding*

May 5, 2005
Revised January 1, 2006

* Min Ding is Assistant Professor of Marketing, Smeal College of Business, Pennsylvania State University, University Park, PA 16802-3007; phone: (814) 865-0622; fax: (814) 865-3015; e-mail: mqd9@psu.edu. The author acknowledges constructive comments from Gary Lilien on several previous versions of this article, as well as those from Rajdeep Grewal, Guang Guo, John Liechty, Bill Ross, participants in presentations given at Cornell University and the Marketing Science conference in Atlanta, and the editor and three anonymous reviewers; financial support from the Smeal College of Business, Pennsylvania State University; and assistance from the Laboratory for Economic Management and Auctions (LEMA) at Pennsylvania State University.

A TRUTH-TELLING MECHANISM FOR CONJOINT ANALYSIS

ABSTRACT

This paper specifies, analyzes, and validates a rigorous and practical truth-telling mechanism (game) for conjoint applications. The mechanism requires only one real product variation and has truth telling in conjoint as its Bayesian Nash equilibrium; we show empirically that the mechanism improves purchase prediction substantially compared with a standard conjoint procedure. We also report three empirical generalizations of the bias induced by the hypothetical nature of standard conjoint.

Conjoint analysis, a centerpiece of marketing research (Carroll and Green 1995; Cattin and Wittink 1982; Wittink and Cattin 1989; Wittink, Vriens, and Burhenne 1994), has many substantive applications, such as new product development (e.g., Kohli and Mahajan 1991), pricing (e.g., Mahajan, Green, and Goldberg 1982), segmentation (e.g., Green and Krieger 1991), and positioning (e.g., Green and Krieger 1992). Since the methodology was introduced to marketing some 30 years ago in a seminal paper by Green and Rao (1971) (for reviews, see Green, Krieger, and Wind 2001; Green and Srinivasan 1978, 1990; Hauser and Rao 2004), researchers have been continuously realizing new and major advances in the field, including hierarchical Bayesian estimation (Allenby and Ginter 1995), polyhedral methods (Toubia, Hauser, and Simester 2004; Toubia et al. 2003), and partial conjoint profiles (Bradlow, Hu, and Ho 2004).

Almost without exception, however, conjoint data have been collected in hypothetical settings that offer no consequences for participants' decisions. The economics literature (e.g., Camerer and Hogarth 1999; Diamond and Hausman 1994; List 2001; Smith and Walker 1993) has long warned about the perils of inferring preferences in such hypothetical conditions, because participants are not incentive aligned to report their true preferences, and all strategies (responses) become weakly dominant (i.e., all strategies have the same payoff). Recent marketing research (Ding, Grewal, and Liechty 2005) shows that data collected in such hypothetical settings have much weaker external validity than data collected from incentive-aligned participants. Across two studies, Ding, Grewal, and Leichty (2005) find that the incentive-aligned choice conjoint considerably outperforms the hypothetical choice conjoint in out-of-sample predictions. They also find that participants in the hypothetical condition exhibit lower heterogeneity in price sensitivity and are less price sensitive on average compared with those in the incentive aligned condition.

In light of this literature, it seems critical that conjoint practitioners attempt to incorporate proper incentives into their studies so that participants are motivated to reveal their true preferences. This task, unfortunately, is not trivial for most applications. Standard guidelines in the experimental economics literature (Smith 1976) require that participants be paid for their performance of all tasks, which implies that, for conjoint studies, all product variations a participant is asked to evaluate should be available at the time of study (or shortly after) for the participant to purchase or consume if he or she so chooses. Any missing variations will likely make the conjoint study hypothetical. Although this requirement can be satisfied easily in some product categories (e.g., food, computer configurations, travel packages), conjoint practitioners and researchers often have access to only a few product variations because other variations are either not possible or too expensive to produce/offer, making it unclear how to incentive align conjoint participants using the established guidelines. Existing literature in both marketing and economics is silent on this important issue, which evidently has contributed to the continued practice of conjoint data collection in hypothetical settings.

An ideal solution to this problem should satisfy several criteria. First, it should be theory driven and provide general guidelines for conjoint practice, not be ad hoc. Second, to provide maximum applicability to existing conjoint methodologies, it should not require significant changes to existing conjoint practice, because the undue resistance these changes might provoke among practitioners would limit its potential adoption. Third, the solution should attain maximum applicability to varied product categories and development stages. To achieve this broad applicability, it should demand as few variations of real products as possible. Fourth, the burden of the approach on practitioners should be minimal in terms of both monetary and nonmonetary commitment. Fifth and along the same lines, its burden on participants should be minimal.

We develop a truth-telling mechanism that satisfies these five criteria using the theory of incomplete information games (more specifically, mechanism design theory: for an accessible introduction, see Mas-Colell, Whinston, and Green 1995). We show theoretically that it is in the best interest—namely, the Bayesian Nash equilibrium (BNE)—of a participant in the proposed mechanism to respond truthfully in conjoint. In an empirical study using iPod packages, this mechanism leads to substantially better out-of-sample prediction than a standard (hypothetical) approach. The empirical study also allows us to establish three generalizations related to hypothetical bias (defined as the bias induced by the hypothetical nature of a task), (1) hypothetical bias is present for expensive durable products as well as for frequently purchased inexpensive products, (2) hypothetical bias pattern associated with price sensitivity for expensive durable products is different from that for frequently purchased inexpensive products, and (3) (replicable) hypothetical bias patterns associated with physical features appear to be linked to how likely a physical feature will be used by the owner if purchased. Under hypothetical condition, on average, participants tend to understate their valuation for physical features that are very likely to be used while overstate their valuation for physical features that are unlikely to be used. Participants also are less heterogeneous in their valuations for physical features under hypothetical condition.

The rest of the article is organized as follows: we first specify the mechanism and present theoretical insights. We then describe the empirical study, followed by a synthesis and generalization of the three empirical insights related to hypothetical bias. Finally, we conclude with general discussions and some fruitful directions for further research.

THEORETICAL FRAMEWORK

Mechanism design theory studies problems in which a principal wants to obtain certain private information about agents but agents will not report this private information truthfully unless given the proper incentives by the principal. The objective of a design is to identify, for a specific problem, a mechanism (embedded in an incomplete information game) for which it is in the best interest of the agents to report their true types.[1] A mechanism contains two elements: (1) a message space from which a player can choose to send a message and (2) an outcome function that determines the outcomes for any given combination of messages sent by all players. When combined with types, beliefs, and payoff functions, a mechanism defines an incomplete information game.

The existing mechanism design literature generally studies two topics: social efficiency (e.g., Clarke 1971; Groves 1973; Vickrey 1961) and revenue maximization (e.g., Myerson 1981; Wilson 1993). None of the existing mechanisms, however, can be readily applied to the conjoint context, as existing mechanisms pertain to products/goods that are either currently available (e.g., auction design) or will be available in the near future if participants so choose (e.g., public goods). However, as previously noted, the challenge in motivating conjoint participants is that most product variations are not available. In addition, the existing mechanism design literature has restricted its attention to cases in which the agent's private information (type) is one dimensional (cf. Krishna and Perry 1998), whereas conjoint participants' private information is clearly multidimensional.

Thus, we specify a new, theoretically sound mechanism to motivate truth telling in conjoint analysis, developed through three key steps. First, we reframe conjoint analysis as an incomplete

[1] Unlike most applied game theoretical models, in which the rules of the games are given exogenously and the task of researchers is to identify the behavior (e.g., equilibrium) of such games, mechanism design focuses on the design of the game (rules) itself, and researchers attempt to identify a game structure such that a desirable behavior (e.g., truth telling) is embedded in the game.

information game. Second, we add a purchase task in which a participant could purchase a real product based on his or her conjoint responses. Third, we employ an incentive-compatible (truth-telling) mechanism to help determine whether a participant could purchase the real product and, if so, at what price.

The intuition for the proposed truth-telling mechanism comes from recasting the conjoint study as a game between a participant and the experimenter. From this perspective, a participant's preference structure (partworths) could be considered his or her type. A participant's type is continuous in a multidimensional space, in which the dimensions equal the total number of levels studied in a given conjoint application. The participant's responses to the conjoint questions in turn can be considered his or her strategies in this game. The information in this game is asymmetric: A participant's type is known to him- or herself but not to the experimenter. The experimenter poses various (conjoint) questions to acquire information about the participant's type, but the participant has no incentive to adopt a strategy that is consistent with his or her type because all possible strategies have the same payoff (i.e., there is no consequence regardless of his or her responses).

Once a conjoint study is recast as an incomplete information game, the task of incentivizing conjoint participants becomes equivalent to modifying the game such that the truth-telling strategy becomes the BNE. To do so, a consequence must be appended to the answers to all conjoint questions. We achieve this consequence by adding a task in which a participant could purchase a real product, depending on his or her willingness to pay (WTP), as inferred by the experimenter using the participant's conjoint responses. The intuition behind this modification is follows: According to the definition of conjoint, an experimenter can infer an individual participant's preference structure on the basis of his or her conjoint responses, and this inferred preference structure enables the experimenter to predict the participant's WTP for any variation of the product being studied. As long as the participant does not know the identity of the real product before completing the conjoint study, he or she will be incentivized to respond to all conjoint questions carefully to ensure that the experimenter will infer his or her entire preference structure accurately. As a result, it is possible to append consequences to all conjoint responses using one real product.

Finally, we incorporate the BDM (Becker, DeGroot, and Marschak 1964) procedure to ensure that it is in the best interest of the participant to have the inferred WTP (based on conjoint responses in the modified game) equal his or her true WTP. The BDM procedure, which guarantees it is in the best interest of a participant to state his or her true WTP, has been widely used in economics, though only recently introduced into marketing (Wertenbroch and Skiera 2002). The procedure involves the following steps: (1) a participant states his or her WTP for an item; (2) a number is drawn randomly from a certain (in most cases, uniform) distribution; and (3) the outcome is determined as follows: If the number drawn is higher than the stated WTP, the participant will not be able to purchase the item, but if the number drawn is lower than or equal to the stated WTP, the participant will be able to purchase the item but pay only the randomly drawn number (price). As a result, either overstating or understating WTP will lead to an inferior outcome for the participant, and the participant's optimal strategy is to state his or her true WTP. In our proposed mechanism, we use the BDM procedure (except replacing the stated WTP with inferred WTP) to determine whether a participant could purchase the real product and, if so, at what price.

An alternative incentive-compatible procedure commonly used in the economics literature is the Vickrey auction (Vickrey 1961), in which the highest bidder wins the auction but pays only an amount equal to the second-highest bid. Although theoretically equivalent (BDM could be considered playing against a bidder that randomly determines its bid), a Vickrey auction is inferior to BDM in our context for three reasons: (1) It requires multiple players and thus unnecessarily complicates the mechanism, (2) people participating in auctions against other humans tend to deviate from their true WTP (i.e., derive utility from beating other human bidders), and (3) it substantially reduces truth-telling incentives for those participants who believe their WTP are smaller than those of most other participants. Because these participants are almost certainly going to lose in a Vickrey auction, they have little incentive to think carefully about their conjoint responses.

In Figure 1, we present the complete mechanism graphically. Note that participants are informed about the entire process (game) prior to the start of the conjoint task. The mechanism proceeds as follows: First, each participant completes the standard conjoint task as usual, which can use any type of conjoint methodology (e.g., rating, choice, polyhedral). Second, the experimenter reveals one real product to the participants that they could potentially purchase. Third, after collecting all data, the experimenter estimates the participants' preferences (partworths) using their conjoint responses and infers each participant's WTP for the real product. Fourth, the BDM procedure, using the inferred WTP, determines whether a participant will be able to purchase the real product and, if so, at what price.

Insert Figure 1 Here

The BDM procedure ensures it is in the best interest of a participant to have his or her inferred WTP equal to his or her true WTP, but the BDM procedure itself does not necessarily guarantee, due to errors in conjoint (e.g., design, estimation, quantity and quality of participants' responses), that it will be in the best interest of a participant to respond truthfully in conjoint. We show in Appendix 1 (a formal treatment of the mechanism), truth telling in conjoint indeed represents the unique BNE in this mechanism under very general conditions (including all existing conjoint methodologies). The main result can be stated as Theorem 1:

Theorem 1. *Under the mechanism specified in Figure 1,*

(a) If the variance of inferred WTP for the truth-telling strategy is equal to or smaller than that for non–truth-telling strategies, truth telling in conjoint is the unique BNE;

(b) If the variance of the inferred WTP for the truth-telling strategy is greater than some non–truth-telling strategies, truth telling in conjoint is the unique BNE as long as those non–truth-telling strategies have an expected inferred WTP that is sufficiently different from the true WTP.

Theoretically, all inconsistent strategies in conjoint (i.e., participants are not consistent in their responses to the conjoint questions) should have higher variance than a consistent strategy, whereas all consistent strategies (including truth telling) should have the same variance (as commonly assumed in conjoint theory and practice). As a result, Theorem 1a implies that truth telling is always the unique BNE for the existing conjoint methodologies. Although there is little reason to believe an estimation procedure will produce smaller variance for a non–truth-telling strategy than for the truth-telling strategy, practitioners can take comfort (Theorem 1b) that, even if superior non–truth-telling strategies actually exist for a yet-to-be developed conjoint method and participants are able to discover them, the preference structures represented by these superior non–truth-telling strategies will be similar to those of true preferences.

EMPIRICAL IMPLEMENTATION

The empirical study has two objectives. The first is to provide empirical evidence regarding whether the truth-telling mechanism will lead to improved predictive performance. Furthermore, it aims to test whether lottery incentives (only a certain percentage of participants, selected randomly, receive rewards based on their decision) will work effectively for the mechanism.[2] To achieve this

[2] In existing incentive-aligned research, all participants receive rewards based on their performance or choice. In real conjoint applications, however, practitioners cannot afford to award expensive products to every participant. As a result,

objective, one or more contrast experiment(s) need to be conducted, in which predictive performance for the participants in the standard (hypothetical) conjoint gets compared with that for the participants in the truth-telling mechanism (with lottery incentives).

The second objective is to build on existing literature pertaining to the hypothetical bias observed in conjoint (Ding, Grewal, and Liechty 2005) and explore the generalizability of its key empirical insights. Bass (1995) defined an empirical generalization as “a pattern or regularity that **repeats** over **different** circumstances”. Based on this definition, Ding, Grewal, and Liechty has established two generalizable findings, (1) hypothetical bias exists in inexpensive food products (documented in both Chinese Dinner Special and Snack Combo), and (2) price sensitivity exhibits a specific hypothetical bias pattern related to inexpensive food products[3] (participants in the hypothetical condition have lower heterogeneity in price sensitivity and are less price sensitive on average compared with those in the incentive aligned condition, documented in both Chinese Dinner Special and Snack Combo). Building on these two findings, we want to investigate whether hypothetical bias and the specific price sensitivity bias pattern exist in products that are very different from inexpensive food products. In addition, we want to examine hypothetical bias patterns associated with physical product features. To qualify for empirical generalization, following Bass (1995)’s definition, the physical features to be examined must be present in at least two different products, each is used in a separate contrast experiment, to allow for possible repetition of feature-specific bias patterns in different contexts.

Taken together, these guidelines call for two contrast experiments to test predictive performance. Each experiment uses a different new durable product with a price range of a few

---

it is critical to test the validity of lottery reward structures. Note that the objective here is not to compare the effectiveness of lottery incentives with incentives that reward every participant but rather to examine whether lottery incentives work when it is not possible to provide incentives for every participant.

[3] Price is the only feature present in both Chinese dinner special study and snack combo study, lending itself to possible generalization based on Bass’s definition.

hundred dollars (to explore the pervasiveness of hypothetical bias and the specific bias pattern associated with price sensitivity), and both new products are chosen such that they belong to the same product category and share certain physical features (to examine replicable hypothetical bias patterns for these common physical features). These guidelines also call for identical experimental designs for the two experiments to allow for a clean interpretation of patterns; any changes in the hypothetical bias pattern between the two experiments for the same feature then can be attributed only to the feature itself instead of (artificial) design parameters.

In the rest of this section, we discuss the products, validation task, design, procedure, participants, estimation, and fit and predictive performance. In the following section, we describe the three empirical generalizations related to hypothetical bias based on the estimated partworths. All descriptions apply to both experiments, unless otherwise noted.

Product Category

We selected the products using the following criteria: (1) the potential participants (university students) represent the products' key target market, (2) the products are durables that cost a few hundred dollars, and (3) the products belong to a category in which new products are regularly introduced. Unlike potential implementations of the mechanism (for which only one product variation is needed), the first objective of this empirical study is to demonstrate its validity. Therefore, we use an additional task to demonstrate external validity, in which a participant chooses from many different product variations. As a result and specific to this study, we also needed products for which many variations could be generated readily.

Informal discussions with students at the major U.S. university where the study was conducted revealed that, for durable goods in the price range of a few hundred dollars, they were most interested in digital cameras, digital music players, printers, personal data assistants, and cell telephones. After further investigation of such products, we selected the Apple iPod product category for its overall match with the selection criteria. Most consumers purchase several accessories when they buy this product; Apple's Web site even suggests seven major iPod gift sets (Starter, Teens, College Students, Athletes, Commuters, Travelers, and Gadget Lovers), each of which consists of one version of the iPod and several different accessories (e.g., the Athletes set contains the 1GB iPod Shuffle, armband for the Shuffle, and sport case). Borrowing the gift set concept from Apple, we define the product in the empirical study as an iPod package that consists of a newly launched iPod and several different accessories. To determine the specific attributes and levels for each attribute, we developed an initial list based on the common accessories recommended by Apple's Web site, then selected a subset of this list for the actual experiments following a focus group session with 10 students from the university. The subset (used for both experiments) consists of an iPod (storage size variation), case/holder, headphones, speakers, car audio, power, and warranty. The price levels, according to consultations with the focus group, are realistic yet attractive. The numbers of levels within each attribute are also kept the same in both experiments to allow for identical experimental designs. The final attribute space is $2^2 3^5 4^1$.

Experiment 1 employs the iPod Shuffle, the first new addition to the Apple iPod family after the start of this research project. Experiment 2 uses the next new addition to the iPod family, the iPod Nano, which was launched nine months after the iPod Shuffle. Of the physical accessories in the iPod package (case/holders, headphones, speakers, car audio, and power), half of them (two speakers, two car audio, and one power) appear in both experiments,[4] which allows for general conclusions about the hypothetical bias patterns associated with these five features, if they are replicable. Table 1 summarizes the various aspects of the two experiments.

[4] Half (excluding those come with an iPod, e.g., basic earphone) of the accessories intentionally differed to ensure sufficient variance between the two experiments. Storage size and warranty are specific for each iPod version (Shuffle or Nano) and thus not good candidates for examining replicable hypothetical bias patterns.

Insert Table 1 Here

Validation Task

The purpose of conjoint analysis is to predict a consumer's real-life decisions. As a result, the best metric to judge a new conjoint method is to examine whether it could lead to better predictions for choice decisions similar to those encountered in real life. Because consumers usually are exposed to a large number of options when they make an actual purchase decision (e.g., Best Buy carries more than 20 different digital cameras in its retail stores at any given time), a good validation task for conjoint methods should not only be well constructed (e.g., an orthogonal fractional factorial design) but also include enough options to make it commensurate to real-life choices for the product category under study. A validation task with only three or four options may be too artificially easy to discriminate the performance among different conjoint methods[5]. In the case of the iPod package, for example, four profiles chosen from a space of $2^2 3^5 4^1$ would likely be so different from one another that potential hypothetical bias probably would not affect which of the four options a participant will rank as the most preferred. That is, small validation tasks are not appropriate for establishing the usefulness of a new method like the truth-telling mechanism. We follow this guideline in constructing the validation task[6].

Design

In each experiment, we employed two conditions: one that corresponds to the standard (hypothetical) choice conjoint (control) and one for the truth-telling mechanism. The truth-telling mechanism includes five parts: introduction, conjoint task, purchasing task, external validity task, and a brief survey. The control condition does not include the purchasing task.

---

[5] This is analogous to measuring computer performance. A superior computer will excel in demanding tasks, but most likely will not be noticeably different in handling simple word processing tasks.

[6] Note this is consistent with the validation task used in the relevant literature (Ding, Grewal, and Liechty 2005).

The introduction contains both experimental instructions and a detailed description of the iPod and those accessories used in the experiment. The detailed descriptions (including pictures), which are identical for both conditions in each experiment, were reproduced from Apple's Web site. However, the experimental instructions differ between the two conditions. Instructions for the control condition mimic instructions in standard choice conjoint studies, except that participants were asked to select one package from a list at the end of the study (external validity task) and a randomly selected winner (from every 40–50 participants) would be given his or her chosen option (which includes not purchasing any package from the list), plus the difference between a certain amount of cash ($250 in Experiment 1 and $320 in Experiment 2) and the price of that option. All participants in the control received $10.

The instructions for the truth-telling mechanism follow the corresponding theoretical mechanism. Participants were told that their responses in the conjoint task would be used to infer their WTP for a specific product X, that they would know the identity of X after completing the conjoint task, and that they would participate in an external validity task, as described previously. All participants received $10, and a winner (from every 40–50 participants) would be randomly selected at the end of each experiment. For each winner, a coin toss would determine whether he or she would receive X on the basis of the inferred WTP using the BDM procedure or the option chosen in the external validity task. If X were chosen, a price (x) would be drawn from a uniform distribution that includes all reasonable valuations for an iPod package. If x were less than or equal to the inferred WTP, the winner would receive X at price x, plus the difference between a certain amount of cash ($250 in Experiment 1 and $320 in Experiment 2) and x. In contrast, he or she would receive the cash and no X if x were higher than the inferred WTP. Finally, if the coin toss resulted in the external validity task, the winner would receive the option selected in the external validity task, plus the difference between the cash and the price of that option.

The instructions for the remaining three parts (conjoint task, purchasing task, external validity task) are straightforward, following the practice in the field, participants in the control condition were urged to "image that you were asked to choose RIGHT HERE and RIGHT NOW" in the conjoint task. The product variations (profiles) used in these three parts were generated by SAS experimental design macros to ensure design objectivity, which indicates that a 72-profile design is the most efficient design for the attribute space ($2^2 3^5 4^1$). We then used SAS to generate 72 profiles, with 20 additional nonduplicate profiles for the purchasing and external validity tasks. We divided the 72 profiles into 24 groups using the random sequence generated by SAS (with a few rearrangements to ensure there was no dominant profile in any given group) and used the groups as the 24 choice tasks (after adding an option of not purchasing to each group) in the conjoint task. Of the 20 additional profiles, we eliminated four profiles that would dominate (packages with the higher-end iPod at the lowest price) or be dominated (packages with the lower-end iPod at the highest price). The remaining 16 profiles appeared in the external validity task, along with the option of not purchasing any of the 16 profiles. In the experiments, we used one of the four eliminated profiles as the real product (without price) for the purchasing task in the truth-telling mechanism.

Procedure

Experiment 1 was conducted one month after the iPod Shuffle launch, and Experiment 2 took place one month after the iPod Nano launch (approximately nine months after the iPod Shuffle). At the start of each experiment, the participants received their respective introductions and were urged to familiarize themselves with the descriptions of the iPod and its accessories. They were told to keep the introduction during the entire experiment so that they could refer back to the instructions and descriptions as needed. Participants in the control condition then completed the conjoint task, followed by the external validity task, and finally a brief survey. Participants in the truth-telling mechanism viewed the configuration of the real product in the purchasing task between the conjoint task and the external validity task. Experimenters collected completed responses for each task before distributing material for the next task.

<u>Participants</u>

Participants in both experiments were recruited from the same undergraduate and graduate student population at a major U.S. university. To ensure participants were potential buyers of the new product, the recruiting e-mail and advertisement explicitly stated that students should not participate in the study if they had no interest in purchasing a digital music player. This selective recruiting intends to mimic what practitioners do in the field when they routinely screen participants before conducting new product research (e.g., conjoint). A total of 49 students participated in Experiment 1 and were randomly assigned to the control (24) and truth-telling mechanism (25); a total of 117 students participated in Experiment 2 and were randomly assigned to the control (58) and truth-telling mechanism (59). None of the participants in Experiment 2 had participated in Experiment 1.

The participants in Experiment 2, on average, appeared to be more knowledgeable about the product category than those in Experiment 1 (Table 1). For example, 26% of participants in Experiment 2 already owned a version of the iPod compared with 10% in Experiment 1. This difference is expected given the fast-diffusion of iPod and the nine month time lag between the two experiments. This difference allows us to examine whether hypothetical bias depends on participants' knowledge and familiarity with the product category under study.

Estimation

We use a random effects hierarchical Bayesian multinomial logit model for estimation, similar to that specified by Allenby and Ginter (1995) and Allenby, Arora, and Ginter (1998). The probability that the ith participant chooses the jth alternative from the tth choice set is given by

$$\Pr\left(z_{it}=j\right)=\frac{\exp\left\{\beta_i^T d_{itj}\right\}}{\sum_{\ell}\exp\left\{\beta_i^T d_{it\ell}\right\}}, \tag{1}$$

where z_{it} is the choice made by the ith participant in the tth choice set, d_{itl} describes the lth option in the tth choice set evaluated by the ith participant, and β_i is a vector of partworths for the ith participant. We assume, a priori,

$$\beta_i \sim Normal\left(\overline{\beta},\Lambda\right) \tag{2}$$

and vague conjugate priors for $\overline{\beta}$ and Λ. The hierarchical Bayesian approach enables us to estimate individual-level partworth parameters (β_i), average partworth parameters $(\overline{\beta})$, and the partworth heterogeneity (Λ). Inferences were made after we ensured that the convergence properties of the Markov chain Monte Carlo (MCMC) analysis were met. In addition, we tested a range of different prior values to ensure that the results were invariant to the prior specification. For participants in the truth-telling mechanism, we infer the WTP of each participant by calculating his or her utility for the iPod package in the purchasing task (without price), and then dividing this utility by his or her coefficient for price.

Fit and Predictive Performance

The estimated partworths for Experiment 1 appear in Table 2. To assess the in-sample goodness of fit for Experiment 1, we calculated the percentage of times the model correctly identified the choice in each of the 24 tasks in the conjoint experiment for each participant. The averages are identical between the two conditions (78%). However, the predictive performance for the external validity task improves quite dramatically for the truth-telling mechanism in Experiment 1 (Table 3). That is, we can correctly predict the choices of 36% of the participants in the truth-telling mechanism compared with only 17% in the control condition (p = .085) and 6% with a naïve prediction (1 of 17). We also calculated the percentage of choices in the external validity task that agrees with one of the top two predicted options to measure the sensitivity of the predictive performance. With this criterion, we can correctly predict 64% of choices in the truth-telling mechanism compared with 38% in the control condition (p = .043).

<u>Insert Tables 2 and 3 Here</u>

The estimated partworths for Experiment 2 appear in Table 4. The in-sample fits are almost identical to those in Experiment 1—78% and 79% for the control and truth-telling mechanism, respectively. The out-of-sample predictions are also consistent with those in Experiment 1 (Table 5). We can correctly predict the choices of 34% of participants in the truth-telling mechanism compared with 21% in the control condition (p = .067). With the top two predicted options, we can correctly predict 56% of choices in the truth-telling mechanism compared with 40% in the control condition (p = .047).

<u>Insert Tables 4 and 5 Here</u>

The improvement in out-of-sample predictive performances in both experiments provides empirical validation for the proposed truth-telling mechanism. Furthermore, this improvement in performance is achieved using a cost-effective lottery incentive structure, thus removing a practical hurdle to implementing the truth-telling mechanism for expensive products.

EMPIRICAL GENERALIZATION OF HYPOTHETICAL BIAS

In this section, we discuss three empirical generalizations regarding hypothetical bias, namely, the prevalence of hypothetical bias, the hypothetical bias pattern associated with price sensitivity, and the hypothetical bias patterns associated with physical features.

Prevalence of Hypothetical Bias

Judging from the superior predictive performance under the truth-telling mechanism in both experiments, hypothetical bias indeed exists for expensive durable goods such as iPods. This finding generalizes the previous evidence based on inexpensive food (Ding, Grewal, and Liechty 2005), as iPods and food represent poles on two key product dimensions (price, frequency of purchase). Furthermore, we find that hypothetical bias does not seem to disappear as people become more knowledgeable about a product. Experiment 1 was conducted at a time when an iPod was still a novelty, whereas Experiment 2 was conducted when an iPod had become a must-have student accessory. The student population at the university, on average, had gotten much more knowledgeable about the iPod by the time of Experiment 2, as is reflected in the difference between the participants in the two experiments (Table 1). Nevertheless, hypothetical bias was observed among participants in both experiments.

Characterization of Hypothetical Bias Patterns

Following Ding, Grewal, and Liechty (2005), we characterize hypothetical bias on the basis of the mean and variance of the estimated partworths for price or a physical feature. We first test whether there are statistical differences ($p<.1$) between the partworths for participants in the control and truth-telling mechanism, with regard to their mean (t-test) and variance (F-test). We then examine whether a pattern (denoted as **higher**, **same**, and **lower** – the mean/variance is statistically higher, equivalent to, or lower in the hypothetical condition than in the truth-telling mechanism) observed in Experiment 1 is replicated in Experiment 2.

These results are presented in Table 6. The two metrics (mean and variance) of hypothetical bias prove remarkably consistent (replicable) between the two experiments, allowing us to draw empirical generalizations (Bass 1995). Of the twelve patterns (mean and variance for price and five physical features) observed in Experiment 1, nine are replicated in Experiment 2. We hypothesize that the difference in participants (participants in Experiment 2 are more knowledgeable about the product category than those in Experiment 1) has likely led to the change of the remaining three patterns. We discuss the patterns associated with price and the five physical features in the next two subsections, respectively.

Insert Table 6 Here

Hypothetical Bias Pattern Associated with Price Sensitivity

The hypothetical bias pattern associated with price sensitivity for iPOD packages is characterized by equal mean and higher variance in both experiments. These results differ from those previously reported. Economics (Diamond and Hausman 1994; List 2001) and marketing (Ding, Grewal, and Liechty 2005) literature have shown that people, on average, are less price sensitive in hypothetical conditions because they appear to discount their budget constraints. We hypothesize that this difference in the mean patterns is likely due to experimental design (e.g., price intervals used). A participant who is not paying close attention to his or her budget constraints likely will ignore small price differences (cf. Ding, Grewal, and Liechty [2005], in which the price difference is only $1 between the two closest price levels and the maximum difference is $2). In contrast, participants are much less likely to ignore the difference in price in this study (maximum difference is $90, minimum difference is $30). As a result, we do not observe any mean price sensitivity bias for iPOD packages.

Ding, Grewal, and Leichty (2005) also report that participants in hypothetical conditions have less heterogeneous price sensitivities for Chinese Dinner Special and Snack Combo, contrary to what we observe for iPod packages. We hypothesize that, in the hypothetical condition, participants paid too little attention to the price for inexpensive items (thus appear more homogeneous) and too much attention to the price for expensive items (thus appear more heterogeneous).

Based on the empirical results from this study and Ding, Grewal, and Liechty (2005), we draw two conclusions: (1) Conjoint design may affect whether certain bias patterns of price (e.g., difference in mean) will be observable, and (2) hypothetical bias pattern associated with price sensitivity for expensive durable products is different from that for frequently purchased inexpensive products.

Hypothetical Bias Patterns Associated with Physical Features

Bass (1995) stated that an empirical generalization could only be made if a pattern is replicated in different circumstances. Based on this guideline, the empirical study was designed such that five physical features (Monster iSpeaker, Creative Speaker, Power adapter, Sony cassette adapter, and Belkin FM transmitter) were present in both iPod Shuffle package (Experiment 1) and iPod Nano package (Experiment 2), allowing us to make general statements about hypothetical bias patterns associated with each of the five physical features. These patterns appear to belong to three distinct classes, which we termed PF1, PF2, and PF3.

Three physical features (Monster iSpeaker, Creative Speaker, and Power adapter) are characterized by the same type of hypothetical bias (Class PF1) -- lower means and lower variances.

With the exception of the mean pattern for Power adapter (same mean in Experiment 2), all these patterns observed in Experiment 1 were replicated in Experiment 2. In general, the physical features in Class PF1 appear to be nonessential to the core product (iPod). Once purchase, however, they are likely to be used by the owner (speakers could be used to share music with others and power adapter allows greater flexibility in charging an iPod).

The hypothetical bias pattern associated with Sony cassette adapter is characterized by higher mean and lower variance (Class PF2), except the difference in mean disappeared in Experiment 2. Similar to physical features in Class PF1, cassette adapter also appears to be nonessential to the core product (iPod). But unlike those in Class PF1, it is much less likely to be used if purchased (cassette adapter is only useful if the owner’s car has a cassette player, but many cars now only have CD players).

The hypothetical bias pattern associated with Belkin FM transmitter is characterized by an equal mean and lower variance in Experiment 1 (Class PF3). While the mean pattern was replicated in Experiment 2, its variance pattern in Experiment 2 contrasts with that in Experiment 1. Upon close examination, FM transmitter appears to represent a middle ground between the features in Class PF1 and features in Class PF2, in terms of how likely it will be used if purchased. Although participants in the study are less likely to use FM transmitter than speakers/power adapter (most students do not drive a car, at least not regularly), FM transmitter is likely to be more useful than a cassette adapter as almost all cars are equipped with an FM radio.

The empirical evidences presented above with regard to physical features appear to support the following two conclusions, (1) the hypothetical bias patterns associated with physical features are replicable and qualify for empirical generalization (Bass 1995), and (2) these hypothetical bias patterns are linked to how likely a physical feature will be used by the owner if purchased. On average, participants tend to understate their valuation for physical features that are very likely to be used while overstate their valuation for physical features that are unlikely to be used. They also are less heterogeneous in their valuations for physical features under hypothetical condition.

GENERAL DISCUSSION

Building on the mechanism design literature, this article specifies a truth-telling mechanism that embeds standard conjoint studies in an incomplete information game and proves that it is the BNE for participants to reveal their true preferences in conjoint studies. In addition to its rigorous theoretical foundation, this mechanism contains several desirable features that will facilitate its adoption among practitioners. It does not require any changes in existing conjoint methodologies and could be used for all of them (e.g., rating, choice, polyhedral). As a result, a practitioner still could rely on his or her expertise in any specific conjoint methodology and perform the same data collection and analysis. Equally important, this mechanism removes the onerous burden of requiring that all product variations be available at the time of the experiment (as is required by existing incentive-alignment guidelines), such that only one product variation is needed at the time of the conjoint study. In terms of additional effort, the only major nonfinancial burden it imposes on practitioners is to calculate each participant's WTP for a product variation using the conjoint results after experiment. The additional financial burden is also limited; practitioners will need to provide real products as the prize of random drawings but only to the extent that the expected value of this random drawing for each participant is higher than his or her opportunity cost (e.g., a study of a $200 television may require a 1 in 10 chance of winning, whereas a study of a $2000 refrigerator may only require a 1 in 100 chance of winning). Finally, this mechanism does not impose any additional burdens on participants, other than having to read extended (but easy-to-understand) instructions.

The empirical tests conducted using the iPod Shuffle and iPod Nano packages demonstrate the superior external validity of the truth-telling mechanism and show that such improvement can be achieved using lottery incentives, which reduces the financial cost associated with the implementation of this mechanism. The empirical study also generalizes three important insights with regard to hypothetical bias: Bias exists not only in inexpensive and frequently purchased product categories but also in expensive durable product categories; hypothetical bias pattern associated with price sensitivity for expensive durable products is different from that for frequently purchased inexpensive products; and participants are less heterogeneous in their valuations for physical features under hypothetical condition and, on average, tend to understate their valuation for physical features that are very likely to be used while overstate their valuation for physical features that are unlikely to be used.

Given its sound theoretical foundation and empirical support, this mechanism should lead incentive-aligned conjoint studies to become field tested and then, perhaps become standard practice which will provide greater external validity.

However, many promising areas still deserve continued, active research. First, this general mechanism may not be effective for special applications, such as business-to-business (B2B) or inexpensive products, and alternative tailor-made mechanisms should address each special situation. In the B2B context, in which a purchasing agent (not a final user) is the conjoint participant, a mechanism that involves three players (experimenter, agent, and final user) might be needed. For inexpensive products (e.g., a box of cereal that retails for $2.99), this mechanism may not generate a great enough incentive for some participants to tell the truth, because the potential penalty inherently is limited by the maximum value of the product. Second, following the guidelines of experimental economics, the winner in the mechanism will be endowed with a certain amount of money (equal to or slightly larger than the highest price used in a study), the majority of which will be used to purchase the real product in the study. However, people may behave differently than they would in real life because of this potential gain. It would be interesting to investigate how participants behave if they have to pay for the product using their own money. Third, with regard to the effectiveness of the lottery incentive structure, it would be interesting to conduct controlled experiments and compare the relative effectiveness of the lottery incentive with an incentive structure that rewards every participant, though the latter is not financial feasible for expensive products such as an iPod. Fourth, it appears that differences in the partworth means between the hypothetical condition and the truth-telling mechanism might be reduced as participants become more knowledgeable (and conceivably, have more precise preference structure) about the products (e.g., Power adapter in PF1 and Sony cassette adapter in PF2), though this observation does not necessarily indicate less hypothetical bias overall but only less difference between the means. Further research is needed to study this important question explicitly.

In conclusion, it is critical for practitioners to conduct incentive-aligned conjoint applications because of their greater external validity. We hope the truth-telling mechanism described herein will enable practitioners to do just that.

TABLE 1. COMPARING EXPERIMENTS 1 AND 2

		Experiment 1	Experiment 2
Product Package	iPod	Shuffle 512 MB Shuffle 1 GB	Nano 2 GB Nano 4 GB
	Case/holder	None Armband for Shuffle Sports case	None Armband for Nano Incase leather folio
	Headphones	Basic Earphone for Shuffle** Nike Vapor sport bud Nike Duro behind-the-head	Basic Earphone for Nano** Apple lanyard Sony Fontopia
	Speakers	None Monster iSpeaker* Creative Speaker*	
	Car audio	None Sony cassette adapter* Belkin FM transmitter*	
	Power	USB (built in Shuffle)** Battery pack Power adapter*	USB (built in Nano)** Tekkeon myPower
	Warranty	Basic for Shuffle** Extended for Shuffle	Basic for Nano** Extended for Nano
	Price	$129, $159, $189, $219	$209, $239, $269, $299
Design		Same	
Procedure		Same	
Time of Experiment		1 month after Shuffle launch	1 month after Nano launch, 9 months after Experiment 1
Participants	Number	24 (control), 25 (truth telling)	58 (control), 59 (truth telling)
	% Own MP3 Player	45%	55%
	% Own an iPod	10%	26%
Validation Task		Same	
Estimation		Same	

* These five physical features are present in both experiments.

** These features come with iPods.

TABLE 2. PARAMETER ESTIMATES FOR EXPERIMENT 1

Attribute	Level	Control (hypothetical)		Truth-Telling Mechanism	
		Mean[5]	Hetero.[6]	Mean	Hetero.
INTERCEPT		6.19 (.87)	1.41 (1.43)	5.22 (.66)	1.35 (1.35)
STORAGE	Base: 512 MB				
	1 GB	3.44 (.43)	2.32 (1.25)	2.86 (.31)	1.13 (.68)
CASE/HOLDER	Base: None				
	Armband for Shuffle	1.73 (.36)	1.64 (.87)	.60 (.28)	.90 (.54)
	Sports case	.91 (.28)	.58 (.39)	.74 (.25)	.63 (.37)
HEADPHONES	Base: Apple[1]				
	Apple + Nike Vapor[2]	.72 (.30)	.83 (.48)	.65 (.27)	.61 (.39)
	Apple + Nike Duro[3]	.39 (.29)	.39 (.24)	.35 (.28)	.60 (.36)
SPEAKERS	Base: None				
	Monster iSpeaker	1.17 (.27)	.39 (.24)	1.56 (.27)	.53 (.36)
	Creative speaker	1.59 (.27)	.41 (.25)	1.87 (.29)	.60 (.36)
CAR AUDIO	Base: None				
	Sony cassette adapter	.72 (.26)	.45 (.29)	.24 (.29)	.94 (.54)
	Belkin FM transmitter	1.89 (.30)	.92 (.60)	2.01 (.35)	1.66 (.97)
POWER	Base: USB				
	USB + Battery pack	.06 (.25)	.36 (.23)	.38 (.24)	.42 (.26)
	USB + Power adapter	.17 (.24)	.33 (.19)	.61 (.24)	.46 (.26)
WARRANTY	Base: Basic				
	Extended	.69 (.27)	.82 (.43)	.07 (.22)	.38 (.22)
PRICE[4]		–5.87 (.66)	6.12 (2.83)	–5.43 (.47)	1.96 (1.25)

[1]Basic Apple earphone that comes with any iPod Shuffle purchase.
[2]Nike Vapor sport bud headphones.
[3]Nike Duro behind-the-head headphones.
[4]Price is divided by 100 before estimation for ease of presentation.
[5]Posterior mean and standard deviation of $\overline{\beta}$
[6]Posterior mean and standard deviation of diagonal of Λ

TABLE 3. PREDICTIVE PERFORMANCE FOR EXPERIMENT 1

Condition	Total Number	Actual Choice Matches the Top Predicted Option		Actual Choice Matches One of the Top Two Predicted Options	
		Number Correct*	Percentage	Number Correct*	Percentage
Control	24	4	17%	9	38%
Truth-telling mechanism	25	9	36%	16	64%

*Bootstrap is used to obtain the probability of observing a difference between the two samples that is at least this extreme if they were drawn from the same population. This probability is .085 and .043 for the top and the top two predicted options, respectively.

TABLE 4. PARAMETER ESTIMATES FOR EXPERIMENT 2

Attribute	Level	Control (hypothetical)		Truth-Telling Mechanism	
		Mean[6]	Hetero.[7]	Mean	Hetero.
INTERCEPT		9.54 (1.11)	44.81 (12.2)	9.56 (.50)	1.34 (.91)
STORAGE	Base: 2 GB				
	4 GB	3.01 (.25)	2.26 (.76)	2.96 (.25)	2.69 (.78)
CASE/HOLDER	Base: None				
	Armband for Nano	.83 (.17)	.55 (.26)	1.18 (.19)	.68 (.28)
	Incase leather folio	.80 (.15)	.41 (.20)	.79 (.15)	.35 (.17)
HEADPHONES	Base: Apple[1]				
	Apple + lanyard[2]	.44 (.15)	.25 (.10)	.51 (.16)	.46 (.19)
	Apple + Sony[3]	.57 (.17)	.45 (.19)	.60 (.15)	.38 (.20)
SPEAKERS	Base: None				
	Monster iSpeaker	1.00 (.17)	.70 (.30)	1.26 (.21)	1.60 (.56)
	Creative speaker	1.40 (.18)	.68 (.28)	1.90 (.19)	1.39 (.51)
CAR AUDIO	Base: None				
	Sony cassette adapter	.31 (.17)	.79 (.31)	.32 (.21)	1.22 (.45)
	Belkin FM transmitter	1.43 (.17)	.58 (.26)	1.56 (.15)	.35 (.16)
POWER	Base: USB				
	USB + Tekkeon[4]	.80 (.16)	.52 (.22)	.79 (.20)	.54 (.21)
	USB + Power adapter	.79 (.14)	.22 (.09)	.81 (.18)	.33 (.16)
WARRANTY	Base: Basic				
	Extended	.61 (.16)	.62 (.23)	.51 (.15)	.63 (.23)
PRICE[5]		–5.21 (.42)	6.85 (1.78)	–5.63 (.25)	1.37 (.52)

[1] Basic Apple earphone that comes with any iPod Nano purchase.
[2] Apple Nano lanyard headphones.
[3] Sony Fontopia earphones.
[4] Tekkeon myPower for iPod Nano
[5] Price is divided by 100 before estimation for ease of presentation.
[6] Posterior mean and standard deviation of $\overline{\beta}$
[7] Posterior mean and standard deviation of diagonal of Λ

TABLE 5. PREDICTIVE PERFORMANCE FOR EXPERIMENT 2

Condition	Total Number	Actual Choice Matches the Top Predicted Option		Actual Choice Matches One of the Top Two Predicted Options	
		Number Correct*	Percentage	Number Correct*	Percentage
Control	58	12	21%	23	40%
Truth-telling mechanism	59	20	34%	33	56%

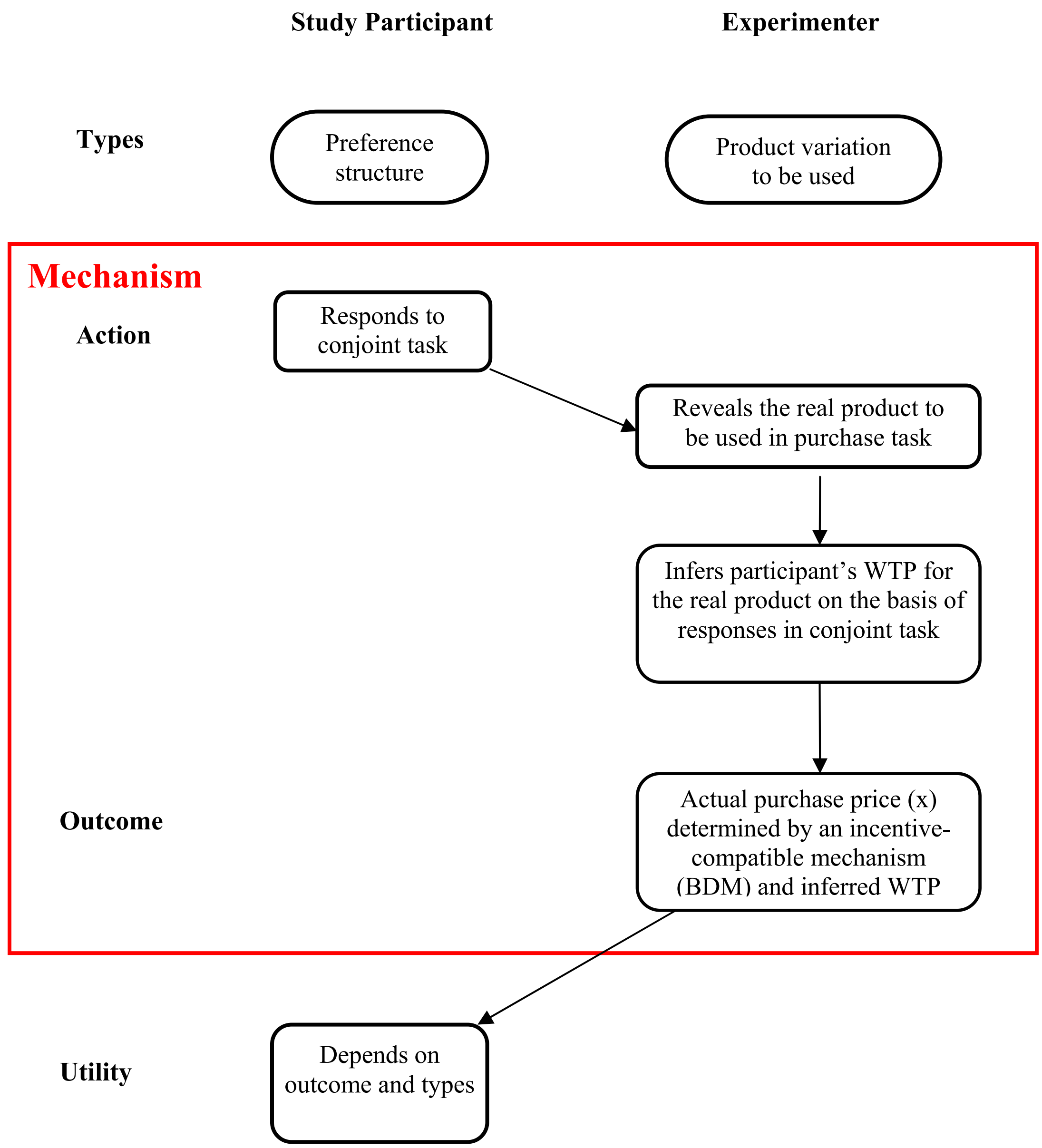
*Bootstrap is used to obtain the probability of observing a difference between the two samples that is at least this extreme if they were drawn from the same population. This probability is .067 and .047 for the top and the top two predicted options, respectively.

TABLE 6. PATTERNS OF HYPOTHETICAL BIAS

	Class	Feature	Mean		Variance	
			Experiment 1	Experiment 2	Experiment 1	Experiment 2
Price		Price	↔	↔	↑	↑
Physical Features	PF1	Monster iSpeaker	↓	↓	↓	↓
		Creative Speaker	↓	↓	↓	↓
		Power adapter	↓*	↔	↓	↓
	PF2	Sony cassette adapter	↑*	↔	↓	↓
	PF3	Belkin FM transmitter	↔	↔	↓*	↑

↔: The mean (or variance) in the control (hypothetical) group is statistically equivalent to that in the truth-telling group.
↑: The mean (or variance) in the control (hypothetical) group is significantly higher ($p<.1$) than that in the truth-telling group.
↓: The mean (or variance) in the control (hypothetical) group is significantly lower ($p<.1$) than that in the truth-telling group.
*: Patterns in Experiment 1 that were not replicated in Experiment 2.

FIGURE 1. A GRAPHIC REPRESENTATION OF THE TRUTH-TELLING MECHANISM



WTP: Willingness-to-pay.
BDM: Becker, DeGroot, and Marschak Procedure (1964)

REFERENCES

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1998), "On the Heterogeneity of Demand," *Journal of Marketing Research*, 35 (August), 384-389.

——— and James L. Ginter (1995), "Using Extremes to Design Products and Segment Markets," *Journal of Marketing Research*, 32 (November), 392-403.

Bass, F.M. (1995), "Empirical Generalizations and Marketing Science: A Personal View", Marketing Science, 14(3), G6-G19.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak (1964), "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 9 (July), 226-232.

Bradlow, Eric T., Ye Hu, and Teck-Hua Ho (2004), "A Learning-Based Model for Imputing Missing Levels in Partial Conjoint Profiles," *Journal of Marketing Research*, 41 (4) 369-381.

Camerer, Colin F. and Robin M. Hogarth (1999), "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," *Journal of Risk and Uncertainty*, 19 (1-3), 7-42.

Carroll, J. Douglas and Paul E. Green (1995), "Psychometric Methods in Marketing Research: Part I, Conjoint Analysis," *Journal of Marketing Research*, 32 (November), 385-391.

Cattin, Philippe and Dick R. Wittink (1982), "Commercial Use of Conjoint Analysis: A Survey," *Journal of Marketing*, 46 (July), 44-53.

Clarke, E. (1971) "Multipart Pricing of Public Goods," *Public Choice*, 2**,** 19-33.

Diamond, P.A. and J.A. Hausman (1994), "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives*, 8 (4), 45-64.

Ding, M., R. Grewal, and J. Liechty (2005), "Incentive-Aligned Conjoint Analysis," *Journal of Marketing Research*, 42 (1), 67-82.

Green, Paul E. and Abba M. Krieger (1991), "Segmenting Markets with Conjoint Analysis," *Journal of Marketing*, 55 (October), 20-31.

——— and ——— (1992), "An Application of Product Positioning Model to Pharmaceutical Products," *Marketing Science*, 11 (2), 117-132.

———, ———, and Yoram Wind (2001), "Thirty Years of Conjoint Analysis: Reflections and Prospects," *Interfaces*, 31 (3), S56-S73.

——— and Vithala R. Rao (1971), "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research*, 8 (August), 355-363.

——— and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5 (September), 103-123.

——— and ——— (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (October), 3-19.

Groves, T. (1973), "Incentives in Teams," *Econometrica*, 41, 617-631.

Hauser, John R. and Vithala R. Rao (2004), "Conjoint Analysis, Related Modeling, and Applications" in *Advances in Marketing Research: Progress and Prospects*, Jerry Wind, ed. Kluwer Academic Publishers.

Kohli, R. and V. Mahajan (1991), "A Reservation-Price Model for Optimal Pricing of Multiattribute Products in Conjoint Analysis," *Journal of Marketing Research*, 28 (3), 347-354.

Krishna, V. and M. Perry (1998), "Efficient Mechanism Design," mimeo, Penn State University and Hebrew University of Jerusalem (April).

List, J.A. (2001), "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards," *American Economic Review*, 91 (5), 1498-1507.

Mahajan, V., P.E. Green, and S.M. Goldberg (1982), "A Conjoint Model for Measuring Self- and Cross-Price/Demand Relationships," *Journal of Marketing Research*, 19 (August), 334-342.

Mas-Colell, A., M. Whinston, and J. Green (1995), *Microeconomic Theory*. Oxford: Oxford University Press.

Myerson, R. (1981), "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58-73.

Smith, Vernon L. (1976), "Experimental Economics: Induced Value Theory," *American Economic Review*, 66 (2), 274-79.

——— and James M. Walker (1993), "Monetary Rewards and Decision Cost in Experimental Economics," *Economic-Inquiry*, 31 (2), 245-61.

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

———, Duncan I. Simester, John R. Hauser, and Ely Dahan (2003), "Fast Polyhedral Adaptive Conjoint Estimation," *Marketing Science*, 22 (3), 273-303.

Vickrey, W. (1961), "Counterspeculation, Auctions and Competitive Sealed Tenders," *Journal of Finance*, 16, 8-37.

Wertenbroch, Klaus and Bernd Skiera (2002), "Measuring Consumers' Willingness to Pay at the Point of Purchase," *Journal of Marketing Research*, 39 (May), 228-241.

Wilson, R. (1993). *Nonlinear Pricing*. Oxford: Oxford University Press.

Wittink, Dick R. and Philippe Cattin (1989), "Commercial Use of Conjoint Analysis: An Update," *Journal of Marketing*, 53 (July), 91-96.

———, Marco Vriens, and Wim Burhenne (1994), "Commercial Use of Conjoint Analysis in Europe: Results and Critical Reflections," *International Journal of Research in Marketing*, 11 (1), 41-52.

APPENDIX 1. FORMAL SPECIFICATION OF THE TRUTH-TELLING MECHANISM

This appendix formally specifies the proposed incomplete information game, and describes its key theoretical properties. Adopting the structure used by Mas-Colell, Whinston, and Green (1995), we define the incomplete information game by (D1) types, (D2) probability, (D3) mechanism, and (D4) payoff (utility) functions. Specifically,

D1 *Each participant has an N_1-dimensional type $t_1 = (t_1(1), t_1(2), \cdots, t_1(N_1)) \in \mathbf{R}^{N_1}$, where N_1 is the total number of levels across all attributes. A specific participant type corresponds to a specific preference structure. The experimenter has an N_2-dimensional type $t_2 = (t_2(1), t_2(2), \cdots, t_2(N_2)) \in \mathbf{R}^{N_2}$, where N_2 is the total number of attributes. A specific experimenter type corresponds to a specific product variation (profile). T_1, T_2 denote the sets of possible types for the participant and experimenter, respectively.*[7]

D2 *Each player knows his or her own type but not the other player's type. The participant knows the probability distribution of the experimenter's type, denoted as $p_1(t_2)$.*

D3 *The mechanism $\Gamma = (A_1, A_2, o(\cdot))$ has possible strategy sets (A_1, A_2) and an outcome function $o : A_1 \times A_2 \to Z$.*

The strategy space for the participant (A_1) is all possible combinations of answers the participant could provide to all tasks in the conjoint study, which may or may not reflect his or her true type t_1. The strategy space for the experimenter (A_2) is all possible real products, and the experimenter's action is simply to reveal the real product (type t_2). The outcome function (o) is the BDM procedure, in which a random price (x) is drawn from a uniform distribution and

[7] Alternatively, the total dimensions of participant's type could be interpreted as the sum of the attribute levels of more than one product categories, and the experimenter's type is a combination of products from these categories.

compared with the inferred WTP (w^i). If x is equal to or less than w^i, the participant receives the product but only pays x. The participant cannot purchase the product if x is greater than w^i:

D4 *The participant's utility* $eu_1(\cdot)$ *is determined by the outcome function* $o: A_1 \times A_2 \to Z$*, his or her type* t_1*, and the experimenter's type* t_2.

The participant's payoff (utility) $eu_1(\cdot)$ is the expected increase in utility if he or she were able to purchase the product at the randomly drawn price x. This expected utility could be obtained by considering three levels of expectations over (1) the experimenter's type t_2 (the identity of the real product is not revealed at the time of conjoint), (2) the inferred WTP (estimation is unbiased but has error due to design, estimation method, quantity or quality of participants), and (3) x. Using the characteristics of variance and the expected value of continuous distributions, the expected utility for a given participant can be obtained, which is stated as Lemma A1:

Lemma A1. Assuming the purchase price (x) is randomly drawn from a uniform distribution, $x \in [\underline{c}, \overline{c}]$ and $w^i \in [\underline{c}, \overline{c}]$, the expected utility for a type t_1 participant who chooses strategy a_1 in the conjoint study is

$$eu_1 = \frac{E\left[\left(W(t_2) - \underline{c}\right)^2\right] - E\left[\left(W(t_2) - m(t_2)\right)^2\right] - E\left[v(t_2)\right]}{2(\overline{c} - \underline{c})}, \qquad \text{(A1)}$$

where m and v are the mean and variance of w^i, respectively, and W is the true WTP.

The main result follows directly from Lemma A1:

Theorem A1. Truth telling in a conjoint study is the unique BNE if and only if,

(a) $E\left[V(t_2)\right] \le E\left[v(t_2)\right]$ for any participant strategy a_1, or (A2)

(b) $E\left[\left(m(t_2) - W(t_2)\right)^2\right] \ge E\left[V(t_2)\right] - E\left[v(t_2)\right]$ for any strategies whose $E\left[v(t_2)\right] < E\left[V(t_2)\right]$ (A3)

where V is the variance of w^i if the participant chooses to tell the truth in the conjoint study.