An Incentive-Aligned Mechanism for Conjoint Analysis
Author(s): Min Ding
Source: *Journal of Marketing Research*, Vol. 44, No. 2 (May, 2007), pp. 214-223
Published by: American Marketing Association
Stable URL: http://www.jstor.org/stable/30162469
Accessed: 17-07-2015 18:13 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Marketing Association is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Marketing Research*.
http://www.jstor.org

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

MIN DING*

This article 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, thus making it possible to incentive align participants in most conjoint applications. Using the iPod package as the context, the author shows empirically that the mechanism substantially improves purchase prediction compared with a standard conjoint procedure.

An Incentive-Aligned Mechanism for Conjoint Analysis

Conjoint analysis is a centerpiece of marketing research. Since the methodology was introduced to marketing approximately 30 years ago in a seminal article by Green and Rao (1971), 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) has long warned about the perils of inferring preferences in such hypothetical conditions, because participants are not incentive aligned to report their true preferences. Recently in marketing research, Ding, Grewal, and Liechty (2005) showed that data collected in such hypothetical settings have weaker external validity than data collected from incentive-aligned participants. They find that the incentive-aligned choice conjoint outperforms the hypothetical choice conjoint in out-of-sample predictions.

*Min Ding is Assistant Professor of Marketing, Smeal College of Business, Pennsylvania State University (e-mail: minding@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, Massachusetts Institute of Technology, and the Marketing Science conference in Atlanta; and the three anonymous *JMR* reviewers. The author also acknowledges financial support from the Smeal College of Business, Pennsylvania State University, and assistance from the Laboratory for Economic Management and Auctions at Pennsylvania State University. This article was processed under Russell Winer, and William Moore served as guest editor.

To read and contribute to reader and author dialogue on JMR, *visit http://www.marketingpower.com/jmrblog.*

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. Unfortunately, this task is not trivial for most applications. Standard guidelines in the experimental economics literature (Smith 1976) require that participants be paid for their performance on all tasks, which implies that all product variations a participant is asked to evaluate should be available for the participant to purchase or consume if he or she so chooses. Any missing variations will likely make the conjoint study hypothetical. In most product categories, conjoint practitioners and researchers have access to only a few product variations, making it unclear how to incentive align conjoint participants using the established guidelines. Existing literature in both marketing and economics is silent on how to address this important issue.

An ideal solution to this problem should satisfy several criteria. First, it should be theory driven and provide general guidelines for conjoint practice. Second, it should not require significant changes to existing conjoint practice. Third, the solution should be applicable to varied product categories and development stages by demanding as few variations of real products as possible. Fourth, the additional burden on practitioners and participants should be minimal.

This article develops a truth-telling mechanism that satisfies these four criteria using the theory of incomplete information games (more specifically, mechanism design theory; for an accessible introduction, see Mas-Colell, Whinston, and Green 1995). I 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. It also confirms that hypothetical bias (defined in the economics literature as the bias induced by the hypo-

© 2007, American Marketing Association
ISSN: 0022-2437 (print), 1547-7193 (electronic)

Journal of Marketing Research
Vol. XLIV (May 2007), 214–223

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

thetical nature of a task) exists for expensive durable products and for frequently purchased inexpensive products (as reported in the literature). The rest of the article is organized as follows: I first specify the mechanism and present theoretical insights. I then describe the empirical study and discuss the results. I conclude with a general discussion 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 they are given the proper incentives by the principal. The objective of a design is to identify a mechanism for which it is in the best interest of the agents to report their true types. 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. 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.

In general, the existing mechanism design literature studies two topics: social efficiency (e.g., Clarke 1971; Groves 1973; Vickrey 1961) and revenue maximization (e.g., Myerson 1981; Wilson 1993). However, none of the existing mechanisms can be readily applied to the conjoint context because they pertain to products/goods that either are currently available (e.g., auction design) or will be available in the near future if participants so choose (e.g., public goods). In addition, the existing mechanism design literature has almost exclusively restricted its attention to cases in which the agent's private information (type) is one dimensional, whereas conjoint participants' private information is multidimensional.

Thus, I specify a new, theoretically sound mechanism to motivate truth telling in conjoint analysis. 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. In turn, the participant's responses to the conjoint questions can be considered his or her strategy. The information in this game is asymmetric: A participant's type is known to him- or herself but not to the experimenter. In standard (unincentivized) conjoint, 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.

When a conjoint study is recast as an incomplete information game, the task of incentivizing conjoint participants becomes equivalent to modifying this game such that the truth-telling strategy becomes the BNE. I accomplish this by adding a task in which a participant can purchase a real product, depending on his or her *inferred* willingness to pay (WTP). The intuition behind this modification is that an experimenter can infer a participant's preference structure on the basis of his or her responses, which 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, I 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 equal his or her true WTP. The BDM procedure, which guarantees that 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 price is drawn randomly from a (typically, uniform) distribution, and (3) the outcome is determined as follows: If the price drawn is higher than the stated WTP, the participant will not be able to purchase the item, but if the price drawn is lower than or equal to the stated WTP, the participant will be able to purchase the item but will pay only the randomly drawn 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 the mechanism proposed herein, I replace the stated WTP with inferred WTP.

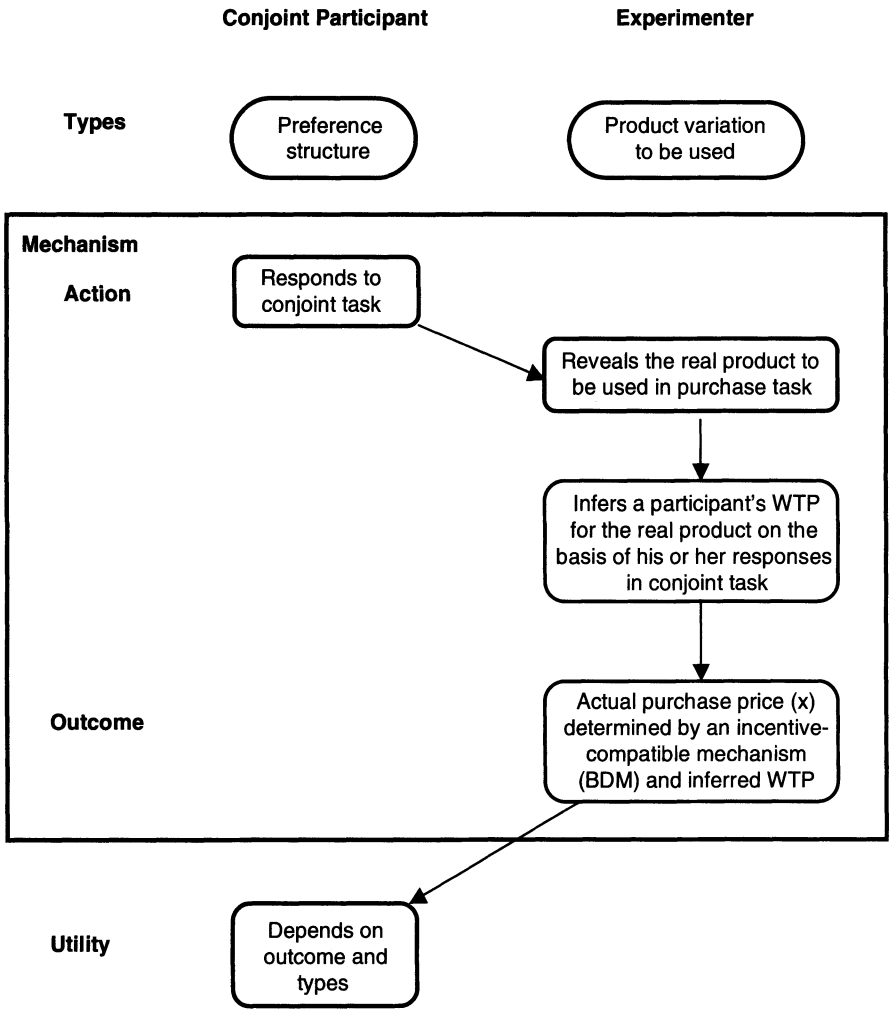
Figure 1 presents the complete mechanism graphically. Note that participants are informed about the entire process (game) before the start of the conjoint task. The mechanism proceeds as follows: First, each participant completes the standard conjoint task as usual, which can use any type of conjoint 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 the data, the experimenter estimates the participants' preferences (partworths) using their conjoint responses and infers each participant's WTP for the real product. Fourth, using the inferred WTP, the BDM procedure determines whether a participant will be able to purchase the real product and, if so, at what price.[1]

The BDM procedure ensures that it is in the best interest of a participant to have his or her inferred WTP equal to his or her true WTP, but because of errors in conjoint (e.g., design, estimation, quantity and quality of participants' responses), the BDM procedure itself does not necessarily guarantee that it will be in the best interest of a participant to respond truthfully in conjoint. The Appendix (a formal

[1]An alternative incentive-compatible procedure commonly used in the economics literature is the Vickrey auction (Vickrey 1961), which is inferior to BDM in the current context for three reasons: (1) It requires multiple players, (2) people participating in auctions against other people tend to deviate from their true WTP, and (3) it substantially reduces truth-telling incentives for participants who believe that their WTPs are smaller than those of most other participants.

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

Figure 1
A GRAPHIC REPRESENTATION OF THE TRUTH-TELLING MECHANISM



treatment of the mechanism) shows that truth telling in conjoint indeed represents the unique BNE in this mechanism under general conditions. 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 the non-truth-telling strategies have an expected inferred WTP that is substantially different from the true WTP (for the precise condition, see the Appendix).

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 is 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. (It is assumed that conjoint methodologies [estimation] are not biased when participants truthfully state their preferences.) Although there is little reason to believe that an estimation procedure will produce a smaller variance for a non-truth-telling strategy than for the truth-telling strategy, practitioners can take comfort (Theorem 1b) in the notion 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 leads to improved predictive performance. Furthermore, it aims to test whether lottery incentives (only a certain percentage of randomly selected participants

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

receive rewards based on their decision) work effectively for the mechanism.[2]

The second objective is to further explore patterns of differences in partworths between the hypothetical condition and an incentive-aligned condition (e.g., under the truth-telling mechanism). In their study of Chinese dinner specials and snack combos, Ding, Grewal, and Liechty (2005) establish two replicable findings in inexpensive food products: (1) Hypothetical bias exists, and (2) price sensitivity is both smaller and less heterogeneous in the standard (hypothetical) condition. I investigate these findings in products that are different from inexpensive food. In addition, I examine whether there are replicable hypothetical bias patterns associated with physical product features.[3] Taken together, these two objectives call for two identically designed experiments (thus, any changes in the hypothetical bias pattern between the two experiments for the same feature can be attributed only to the feature itself rather than to [artificial] design parameters) to test predictive performance; each has two conditions (hypothetical and truth telling), and both use new durable products that have a price range of a few hundred dollars and share several physical features.

The rest of this section discusses the products, validation task, design, participants, estimation, and fit and predictive performance. Finally, I discuss patterns of difference in estimated partworths between the hypothetical and truth-telling mechanism conditions for price and five common physical features present in both experiments. Unless otherwise noted, all descriptions apply to both experiments.

Product Category

After informal discussions with students at a major U.S. university where the study was conducted, I selected the Apple iPod product category for its overall match with the selection criteria. Most consumers purchase several accessories when they buy this product; at the time of the Nano and Shuffle launches, Apple's Web site even suggested seven iPod gift sets (Starter, Teens, College Students, Athletes, Commuters, Travelers, and Gadget Lovers), each of which consisted 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, I 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, I developed an initial list based on the common accessories that the Apple Web site recommended. Then, on the basis of a focus group session with ten students from the university, I selected a subset of this list for the actual experiments. The subset consists of an iPod (storage-size variation), a case/holder, headphones, speakers, car audio, power, and a warranty. 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 $2^2 3^5 4^1$ and includes four price levels.

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 makes it possible to examine replicable patterns of difference in partworths associated with these five physical features.

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 can 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 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 ranks as the most preferred. That is, small validation tasks are not appropriate for establishing the usefulness of a new method, such as the truth-telling mechanism. This guideline is followed in the construction of the validation task.

Design

In each experiment, two conditions were employed: 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 (in which the identity of the real product that could be purchased is revealed), external validity task, and a brief survey. The control condition does not include the purchasing task.

[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, 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]If a physical feature to be examined is present in at least two different products, each is used in a separate experiment. It is possible to test for possible repetition of feature-specific bias patterns in different contexts and thus contribute to empirical generalization in the future (defined by Bass [1995, p. G7] as "a pattern or regularity that repeats over different circumstances").

[4]Half of the accessories (excluding those that come with an iPod; e.g., basic earphones) were 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 are not good candidates for examining replicable hypothetical bias patterns.

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

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

The introduction contains both experimental instructions and a detailed description of the iPod and the 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. 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.

Participants in the truth-telling condition 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 (the conjoint task, the purchasing task, and the external validity task) are straightforward and follow 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 ($2^2 3^5 4^1$). Then, SAS was used to generate 72 profiles, with 20 additional nonduplicate profiles for the purchasing and external validity tasks. The 72 profiles were divided into 24 groups by means of the random sequence generated by SAS (with a few rearrangements to ensure that there was no dominant profile in any given group), and the groups were used as the 24 choice tasks (after an option of not purchasing was added to each group) in the conjoint task. In each choice task, a participant's job was to pick his or her most preferred option out of four (three different packages plus the not-purchasing option). Of the 20 additional profiles, 4 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) were eliminated. 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, one of the four eliminated profiles was used as the real product (without price) for the purchasing task in the truth-telling mechanism.

Participants

Participants in both experiments were recruited from the same undergraduate and graduate student population at a major U.S. university. Experiment 1 was conducted one month after the iPod Shuffle launch, and Experiment 2 took place one month after the iPod Nano launch. To ensure that 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. A total of 49 students participated in Experiment 1 and were randomly assigned to the control (24) and the 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). No one participated in both experiments.

Estimation

A random-effects hierarchical Bayesian multinomial logit model is used for estimation,[6] 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

$$\Pr(z_{it} = j) = \frac{\exp\{\beta^T_i d_{itj}\}}{\Sigma_l \exp\{\beta^T_i d_{itl}\}}, \tag{1}$$

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. A priori, it is assumed that

$$\beta_i \sim \text{Normal}\,(\bar{\beta}, \Lambda) \tag{2}$$

and that there are vague conjugate priors for $\bar{\beta}$ and Λ. The hierarchical Bayesian approach makes it possible to estimate individual-level partworth parameters (β_i), average partworth parameters ($\bar{\beta}$), and the partworth heterogeneity (Λ). Inferences were made after it was ensured that the convergence properties of the Markov chain Monte Carlo analysis were met. In addition, a range of different prior values was tested to ensure that the results were invariant to the prior specification. A participant's WTP was inferred for the iPod package in the purchasing task for each draw of the sampler after convergence, and the average was taken as the final estimate.

Fit and Predictive Performance

The estimated partworths for Experiment 1 appear in Table 1. To assess the in-sample goodness of fit for Experiment 1, the percentage of times the model correctly identified the choice in each of the 24 tasks in the conjoint experiment was calculated for each participant. The averages are

[6]This particular estimation approach precludes an experimenter from estimating WTP in real time. In the empirical study, participants in the truth-telling condition were told that the winner would be invited back the next day for the BDM procedure (if BDM is chosen on the basis of the coin flip).

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

Table 1
PARAMETER ESTIMATES FOR EXPERIMENT 1

		Control (Hypothetical)		*Truth-Telling Mechanism*	
Attribute	*Level*	*M*[e]	*Heterogeneity*[f]	*M*	*Heterogeneity*
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[a]				
	Apple + Nike Vapor[b]	.72 (.30)	.83 (.48)	.65 (.27)	.61 (.39)
	Apple + Nike Duro[c]	.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[d]		–5.87 (.66)	6.12 (2.83)	–5.43 (.47)	1.96 (1.25)

[a]Basic Apple earphone that comes with any iPod Shuffle purchase.
[b]Nike Vapor sport bud headphones.
[c]Nike Duro behind-the-head headphones.
[d]Price levels are \$129, \$159, \$189, and \$219. They are divided by 100 before estimation for ease of presentation.
[e]Posterior mean and standard deviation of $\bar{\beta}$.
[f]Posterior mean and standard deviation of diagonal of Λ.
Notes: The five common physical features present in both experiments are in bold.

Table 2
PREDICTIVE PERFORMANCE FOR EXPERIMENT 1

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

identical between the two conditions (78%). However, the predictive performance for the external validity task improves considerably for the truth-telling mechanism in Experiment 1 (Table 2). That is, the choices of 36% of the participants in the truth-telling mechanism can be correctly predicted, compared with only 17% in the control condition (p = .085)[7] and 6% with a naive prediction (1 of 17). The percentage of choices in the external validity task that agrees with one of the top two predicted options was also calculated to measure the sensitivity of the predictive performance. With this criterion, 64% of choices in the truth-telling mechanism can be correctly predicted, compared with 38% in the control condition (p = .043).

The estimated partworths for Experiment 2 appear in Table 3. 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 4). The choices of 34% of participants in the truth-telling mechanism can be correctly predicted, compared with 21% in the control condition (p = .067). With the top two predicted options, 56% of choices in the truth-telling mechanism can be correctly predicted, compared with 40% in the control condition (p = .047).[8]

[7]Bootstrap is used to obtain the probability of observing a difference in the number of correct predictions between the control and the truth-telling mechanism that would be at least this extreme if they were drawn from the same population (based on the standard procedure of resampling with replacement). The population in the bootstrap is created by aggregating the prediction outcomes (numbers of correct and incorrect predictions) from the control and truth-telling mechanism under each situation.

[8]To demonstrate empirically the importance of using a large and realistic validation task in benchmarking different methodologies (see the discussion under the "Validation Task" subsection) and to provide a baseline to existing literature, participants were asked in Experiment 2 to complete

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

Table 3
PARAMETER ESTIMATES FOR EXPERIMENT 2

		Control (Hypothetical)		Truth-Telling Mechanism	
Attribute	*Level*	*M*[f]	*Heterogeneity*[g]	*M*	*Heterogeneity*
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[a]				
	Apple + lanyard[b]	.44 (.15)	.25 (.10)	.51 (.16)	.46 (.19)
	Apple + Sony[c]	.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[d]	.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[e]		–5.21 (.42)	6.85 (1.78)	–5.63 (.25)	1.37 (.52)

[a]Basic Apple earphone that comes with any iPod Nano purchase.
[b]Apple Nano lanyard headphones.
[c]Sony Fontopia earphones.
[d]Tekkeon myPower for iPod Nano.
[e]Price levels are \$209, \$239, \$269, and \$299. They are divided by 100 before estimation for ease of presentation.
[f]Posterior mean and standard deviation of $\bar{\beta}$.
[g]Posterior mean and standard deviation of diagonal of Λ.
Notes: The five common physical features present in both experiments are in bold.

Table 4
PREDICTIVE PERFORMANCE FOR EXPERIMENT 2

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

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 by using a cost-effective lottery incentive structure, thus removing a practical hurdle to the implementation of the truth-telling mechanism for expensive products.

an additional and more traditional validation task (four choice questions, each of which contained four different iPod Nano packages and the option of not purchasing). According to both the top option and the top two options, the truth-telling mechanism led to better predictions, but the difference is not significant for the top option ($p = .151$) and is only marginally significant for the top two options ($p = .088$). This confirms the importance of using large and realistic validation tasks to compare methods.

Patterns of Differences in Partworths for Price and Five Common Physical Features[9]

Judging from the superior predictive performance under the truth-telling mechanism in both experiments, it is concluded that 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) because iPods and food represent the poles on two key product dimensions (price and frequency of purchase).

The patterns of differences associated with price sensitivity for iPod packages are characterized by a similar mean

[9]A more rigorous analysis of the patterns of difference is available on request.

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

and higher variance in both experiments. (The terms "similar," "lower," or "higher" are used to refer to the patterns in which the value [mean or variance] under the hypothetical condition is similar to, lower than, or higher than that under the truth-telling mechanism, respectively.) These results differ from those previously reported. Research in both economics (Diamond and Hausman 1994; List 2001) and marketing (Ding, Grewal, and Liechty 2005) has shown that, on average, people are less price sensitive in hypothetical conditions because they appear to discount their budget constraints. It is hypothesized 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 will likely 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 (the maximum difference is $90, and the minimum difference is $30). Ding, Grewal, and Liechty (2005) also report that participants in hypothetical conditions have less heterogeneous price sensitivities for the Chinese dinner special and the snack combo, contrary to what is observed for iPod packages. Further research is needed to understand how price sensitivity changes under hypothetical conditions.

Of the five common physical features (which can be examined for replicable patterns of differences in partworths), three (Monster iSpeaker, Creative speaker, and the power adapter) are characterized by the same type of patterns—namely, lower means and lower variances.[10] With the exception of the mean pattern for the power adapter (similar mean in Experiment 2), all the patterns observed in Experiment 1 were replicated in Experiment 2. The pattern of differences associated with the Sony cassette adapter is characterized by a higher mean and lower variance, except that the difference in mean disappeared in Experiment 2. The pattern of differences associated with the Belkin FM transmitter is characterized by a similar mean and lower variance in Experiment 1. Although the mean pattern was replicated in Experiment 2, its variance pattern in Experiment 2 contrasts with that in Experiment 1. The empirical evidence presented here suggests that the patterns of differences associated with physical features are mostly replicable and depend on the physical features. It appears that these patterns are related to how likely it is that the owner will use a physical feature, given purchase. It is conjectured that under hypothetical conditions, on average, participants tend to understate their valuation for physical features they are likely to use (e.g., speakers, the power adapter) and to overstate their valuation for physical features they are unlikely to use (cassette adapter). Further study is needed to test this conjecture.

On the basis of these empirical results, it is clear that the WTP for a complex product will not necessarily be lower under an incentive-aligned condition. In the case of the iPod, a participant may have higher utility for a speaker under the truth-telling mechanism and, as a result, may have a higher WTP for an iPod package that includes this speaker.

[10]If formal between-experiment comparisons were to be made, ratios of coefficients (e.g., against the price sensitivity) should be used. However, the results would not likely be much different in this study, because the price sensitivities are sufficiently similar between the experiments. I thank an anonymous reviewer for pointing this out.

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 the 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 the 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 $2,000 refrigerator may require only a 1 in 100 chance of winning). Finally, this mechanism does not impose any additional burdens on participants, other than needing to read extended (but easy-to-understand) instructions. (Note that a participant needs to learn only once if he or she participates in many different conjoint studies that contain this mechanism.)

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 demonstrates the following: (1) Hypothetical bias exists not only in inexpensive and frequently purchased product categories but also in expensive durable product categories, (2) the patterns of differences associated with price sensitivity for expensive durable products are different from those for frequently purchased inexpensive products, and (3) the patterns of differences in partworths associated with physical attributes are feature specific.

Given its sound theoretical foundation and empirical support, this mechanism should lead incentive-aligned conjoint studies to become field tested and then perhaps to 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 inexpensive products (e.g., a box of cereal that retails for $2.99). It is unlikely to generate large enough incentives for some participants to tell the truth in these cases, because the potential penalty is inherently lim-

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

ited by the maximum value of the product. Second, in line with 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 worthwhile to investigate how participants behave if they must pay for the product with their own money. Third, some conjoint participants may not completely understand the mechanism but may believe that they will be better off if they try harder to fill it out as conscientiously (and as truthfully) as possible. It would be worthwhile to investigate the mental process involved. Fourth, with regard to the effectiveness of the lottery incentive structure, it would be worthwhile 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 financially feasible for expensive products, such as an iPod. Finally, the current mechanism requires that one version of the product be physically available. As a result, it is not applicable to really new product concepts for which no physical product (not even a prototype) is available. This represents an important and fruitful direction for further research.

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

APPENDIX: FORMAL SPECIFICATION OF THE TRUTH-TELLING MECHANISM

This Appendix formally specifies the proposed incomplete information game and describes its key theoretical properties. In line with the structure that Mas-Colell, Whinston, and Green (1995) use, the incomplete information game is defined by (D1) types, (D2) probability, (D3) mechanism, and (D4) payoff (utility) functions. Specifically,

D1: Each participant has an N_1-dimensional type $t_1 = (t_1[1], t_1[2], \ldots, t_1[N_1]) \in \mathbf{R}^{N1}$, 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 $t_2 = (t_2[1], t_2[2], \ldots, t_2[N_2]) \in \mathbf{R}^{N2}$, 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. Alternatively, the total dimensions of a participant's type could be interpreted as the sum of the attribute levels of more than one product category, and the experimenter's type is a combination of products from these categories.

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 $o: A_1 \times A_2 \rightarrow Z$.

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 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 $o: A_1 \times A_2 \rightarrow Z$, 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, and 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: If it is assumed that the purchase price (x) is randomly drawn from a uniform distribution, $x \in [c_1, c_2]$ and $w^i \in [c_1, c_2]$, the expected utility for a type t_1 participant who chooses strategy a_1 in the conjoint study is

$$\text{(A1)} \quad eu_1 = \frac{(E\{[W(t_2) - c_1]^2\} - E\{[W(t_2) - m(t_2)]^2\} - E[v(t_2)])}{[2(c_2 - c_1)]},$$

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 (assuming a conjoint method is unbiased if truth-telling strategy is adopted):

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

$$\text{(A2)} \quad E[V(t_2)] \leq E[v(t_2)], \text{ for any participant strategy } a_1, \text{ or}$$

$$\text{(A3)} \quad E\{[m(t_2) - W(t_2)]^2\} \geq E[V(t_2)] - E[v(t_2)],$$

$$\text{for any strategies whose } E[V(t_2)] > E[v(t_2)],$$

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

REFERENCES

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1998), "On the Heterogeneity of Demand," *Journal of Marketing Research*, 35 (August), 384–89.

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

Bass, Frank 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), 226–32.

Bradlow, Eric T., Ye Hu, and Teck H. Ho (2004), "A Learning-Based Model for Imputing Missing Levels in Partial Conjoint

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions

Profiles," *Journal of Marketing Research*, 41 (November), 369–81.

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.

Clarke, Edward H. (1971), "Multipart Pricing of Public Goods," *Public Choice*, 11 (1), 19–33.

Diamond, Peter A. and Jerry A. Hausman (1994), "Contingent Valuation: Is Some Number Better Than No Number?" *Journal of Economic Perspectives*, 8 (4), 45–64.

Ding, Min, Rajdeep Grewal, and John Liechty (2005), "Incentive-Aligned Conjoint Analysis," *Journal of Marketing Research*, 42 (February), 67–82.

Green, Paul E. and Vithala R. Rao (1971), "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research*, 8 (August), 355–63.

Groves, Theodore (1973), "Incentives in Teams," *Econometrica*, 41 (4), 617–31.

List, John 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.

Mas-Colell, Andreu, Michael Whinston, and Jerry R. Green (1995), *Microeconomic Theory*. Oxford: Oxford University Press.

Myerson, Roger B. (1981), "Optimal Auction Design," *Mathematics of Operations Research*, 6 (1), 58–73.

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

Toubia, Olivier, John R. Hauser, and Duncan I. Simester (2004), "Polyhedral Methods for Adaptive Choice-Based Conjoint Analysis," *Journal of Marketing Research*, 41 (February), 116–31.

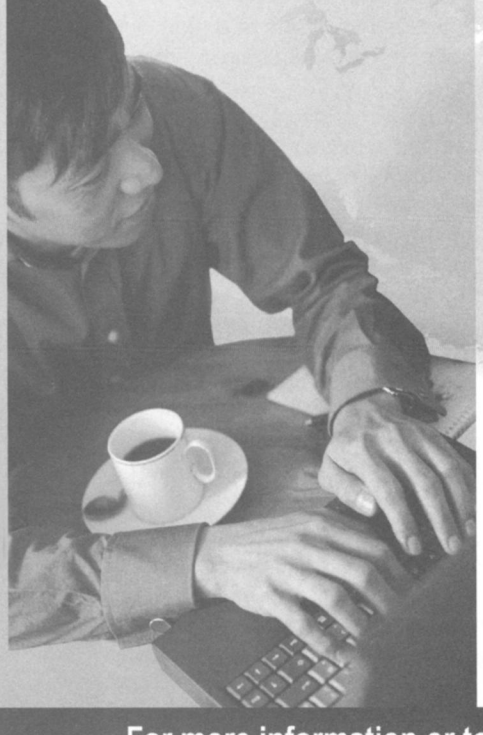
———, Duncan I. Simester, John R. Hauser, and Ely Dahan (2003), "Fast Polyhedral Adaptive Conjoint Estimation," *Marketing Science*, 22 (3), 273–303.

Vickrey, William (1961), "Counterspeculation, Auctions and Competitive Sealed Tenders," *Journal of Finance*, 16 (1), 8–37.

Wertenbroch, Klaus and Bernd Skiera (2002), "Measuring Consumers' Willingness to Pay at the Point of Purchase," *Journal of Marketing Research*, 39 (May), 228–41.

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

The Premier Professional Development Program for Researchers Worldwide



PRINCIPLES OF MARKETING RESEARCH

An Online Certificate Course

NEW! Pharmaceutical Supplements gives students the cutting edge skills to excel in the growing world of healthcare research.

PRINCIPLES OF MARKETING RESEARCH is a vigorous self-study training program based upon the Marketing Research Core Body of Knowledge (MRCBOK©). Delivered via the Internet, the program allows you to progress at your own pace, augmented by consultations with your assigned coach-mentor, upon request.

GRADUATE TESTIMONIAL

The **Principles of Marketing Research** *course is very beneficial for participants from any part of the world. The certificate program is highly recommended for anyone who is pursuing a market research career and needs scientific knowledge in the field.*

— Mookda Somsongkul, MANAGING DIRECTOR, RESEARCH MATTERS CO., LTD., BANGKOK, THAILAND

The *PRINCIPLES OF MARKETING RESEARCH* course has become the research profession's highest quality, most valued, and most widely supported educational program on the fundamentals of marketing research.

The investment in this program is $749 per employee among AMA members and their colleagues. Corporate group discounts are available for 10 or more employees enrolling during a 12-month period.

MRII MARKETING RESEARCH INSTITUTE INTERNATIONAL

AMERICAN MARKETING ASSOCIATION

The University of Georgia
Center for Continuing Education
Conference Center & Hotel

For more information or to register for the course, go to *www.principlesofmarketingresearch.org*, e-mail *Bob.Wells@georgiacenter.uga.edu*, or call +1-706-542-6692

This content downloaded from 132.174.254.159 on Fri, 17 Jul 2015 18:13:30 UTC
All use subject to JSTOR Terms and Conditions