MIN DING, RAJDEEP GREWAL, and JOHN LIECHTY*

Because most conjoint studies are conducted in hypothetical situations with no consumption consequences for the participants, the extent to which the studies are able to uncover "true" consumer preference structures is questionable. Experimental economics literature, with its emphasis on incentive alignment and hypothetical bias, suggests that more realistic incentive-aligned studies result in stronger out-of-sample predictive performance of actual purchase behaviors and provide better estimates of consumer preference structures than do hypothetical studies. To test this hypothesis, the authors design an experiment with conventional (hypothetical) conditions and parallel incentive-aligned counterparts. Using Chinese dinner specials as the context, the authors conduct a field experiment in a Chinese restaurant during dinnertime. The results provide strong evidence in favor of incentive-aligned choice conjoint analysis, in that incentive-aligned choice conjoint outperforms hypothetical choice conjoint in out-of-sample predictions. To determine the robustness of the results, the authors conduct a second study that uses snacks as the context and considers only the choice treatments. This study confirms the results by providing strong evidence in favor of incentive-aligned choice analysis in out-of-sample predictions. The results provide a strong motivation for conjoint practitioners to consider conducting studies in realistic settings using incentive structures that require participants to "live with" their decisions.

Incentive-Aligned Conjoint Analysis

Conjoint analysis, which has developed into a widely applied methodology for making inferences about consumer preferences and for uncovering empirical demand functions (Carrol and Green 1995), has many substantive applications in marketing, such as those for 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). Conjoint analysis also has been applied successfully in practice (Cattin and Wittink 1982; Wittink and Cattin 1989; Wittink, Vriens, and Burhenne 1994), and there is extensive literature on the subject (for reviews, see Green, Krieger, and Wind 2001; Green and Srinivasan 1978, 1990). As a result, there are many variants of conjoint analysis based on the way preference scores are elicited (e.g., ratings, rankings, self-explicated, constant sum, choice), the type of designs used (e.g., full factorial, fractional factorial, adaptive), the type of models estimated (e.g., regression, logit, probit, hierarchical Bayes), and the estimation procedures used to make inferences (e.g., maximum likelihood, Markov chain Monte Carlo). Despite these differences, most methods have certain common elements. Data collection requires consumers to rate, rank, or select alternative products, and the goal of the data analysis is to find the set of partworths that, given a compositional rule, is most consistent with the respondent's overall preferences (Green and Srinivasan 1978).

Although early research on conjoint analysis rarely used out-of-sample predictions to assess model validity, scholars have suggested that such predictions are the strongest means to assess the validity of conjoint studies (Green and Srinivasan 1990). As a result, three types of validation or prediction tasks—aggregate-level market share predictions (e.g., Srinivasan et al. 1981), individual-level predictions of purchase intentions (e.g., Leigh, MacKay, and Summers 1984), and individual-level predictions of actual behaviors (e.g., Srinivasan 1988; see also Green and Srinivasan 1990)—have dominated the conjoint landscape. However, each method has limitations.

*Min Ding is Assistant Professor of Marketing (e-mail: mqd9@psu.edu), Rajdeep Grewal is Assistant Professor of Marketing (e-mail: rug2@psu.edu), and John Liechty is Assistant Professor of Marketing and Statistics (e-mail: jc112@psu.edu), Smeal College of Business Administration, Pennsylvania State University. The authors contributed equally to this manuscript. The authors thank Gary Bolton, Eric Bradlow, Wayne DeSarbo, Jehoshua Eliashberg, Tony Kwasnica, Gary Lilien, and Arvind Ramgaswamy for their constructive comments and two anonymous *JMR* reviewers for their feedback. This research was supported by the Marketing Department at Pennsylvania State University. The authors greatly appreciate the cooperation of New Chinatown restaurant at State College, Pa., and the Laboratory for Economic Management and Auctions at Pennsylvania State University.

Journal of Marketing Research
Vol. XLII (February 2005), 67–82

First, scholars have attempted to predict real-world current (e.g., Davidson 1973; Page and Rosenbaum 1987) or future (e.g., Robinson 1980; Srinivasan et al. 1981) market share using conjoint tasks. Such aggregate-level predictions have confounding problems that are related to the effects of marketing-mix variables other than product design. For example, in his conjoint study of North Atlantic Air, Robinson (1980) uses airfares, discounts, and travel restrictions, in addition to information obtained from the conjoint exercise, to predict future market shares. In such exercises, it becomes difficult to separate the marketing-mix effects, such as advertising and promotions, from the conjoint task effects.

Second, predictions of purchase intentions are unreliable because stated preferences often differ from revealed preferences, which are derived from actual purchase behaviors (Green and Srinivasan 1990). Although attempts have been made to improve the reliability and validity of purchase intention predictions by using Pareto optimal choice sets in the prediction tasks (i.e., choice sets in which none of the alternatives is dominated by the remaining alternatives) (e.g., Elrod, Louviere, and Davey 1992; Johnson, Meyer, and Ghose 1989), these attempts have not always been successful, because non-Pareto, traditional holdout sets can sometimes be more difficult to predict than Pareto choice sets (Green, Helsen, and Shandler 1988). In addition, Pareto optimal choice sets offer no way to link purchase intention to actual purchase behavior.

Third, individual-level predictions of actual behavior usually are carried out through intervention studies in which the researchers perform a conjoint exercise while consumers are involved in actual decision making. For example, Srinivasan (1988), Srinivasan and Park (1997), and Wittink and Montgomery (1979) predict the jobs that MBA students will choose among multiple offers on the basis of self-explicated and rating scores. Wright and Kriewall (1980) predict whether high school seniors will apply to certain universities on the basis of student preferences revealed through a conjoint task. Because these intervention studies involve real decisions that are likely to affect the respondents in profound ways (e.g., the job preferences of MBA students; Wittink and Montgomery 1979), participants are likely to be motivated to reveal their "true" preferences.[1] Despite the merits of intervention studies, they often are not practical, because, in general, intervention tasks are not feasible.

In hypothetical data-collection exercises, participants may not experience strong incentives to expend the cognitive efforts needed to provide researchers with an accurate answer. A rich literature in experimental economics argues that such data can be inconsistent, erratic, and, in many cases, untrustworthy (e.g., see metastudies on the role of incentives; Camerer and Hogarth 1999; Smith and Walker 1993). The theoretical underpinning of this argument is based on the induced value theory (Smith 1976), which states that three conditions must be satisfied to solicit incentive-compatible behavior: monotonicity, salience, and dominance.[2]

The most relevant condition for conjoint analysis is salience, which requires that the reward be directly related to the decisions the participant makes during a study. Most practitioners of conjoint studies pay consumers some money for participation. However, paying a respondent a fixed amount is not salient, because there is no relationship between the respondent's performance/actions and the reward (money) he or she receives. As a result, there is no reason to expect that the respondent's behavior during a study will be consistent with his or her behavior during a similar, real-world, economic activity. (In other words, there are neither rewards nor penalties for respondents to correctly or incorrectly state their preferences.) On the basis of a metastudy of 74 research papers, Camerer and Hogarth (1999, p. 8) find that salient incentives tend to "shift behavior away from an overly socially desirable presentation of oneself to a more realistic one: when [salient] incentives are low participants say they would be more risk-preferring and generous than they actually are when [salient] incentives are increased."

A related stream of literature explicitly studies hypothetical bias in the context of the contingent valuation method (for a review, see Diamond and Hausman 1994). The contingent valuation method suggests that what participants say they would do in hypothetical situations does not necessarily correspond to what they actually do; that is, stated preferences do not always match revealed preferences. For example, in the context of deer-hunting permits, Bishop and Heberlein (1986) find that willingness-to-pay values were significantly overstated in the hypothetical condition than in the actual cash condition. List (2001) shows that sports-card dealers significantly overstated their bids for a sports card in a hypothetical condition compared with the real action ($107.89 versus $59.56). Finally, List and Shogren (1998) find that the selling price for a gift is significantly higher in real situations than in hypothetical situations.

On the basis of the literature on incentive alignment and hypothetical bias, we hypothesize that state-of-the-art conjoint data-collection techniques may fail to uncover preferences that align with actual purchase behavior because of the hypothetical settings in which the data are collected. In hypothetical research settings, respondents may discount their budget constraints or simply state preferences that are inconsistent with their actual behavior (e.g., because of a preference structure expected by peers). To induce realism in hypothetical tasks, we propose to use incentive structures that align with actual purchase behaviors. On the basis of the induced value theory (Smith 1976), we expect that an incentive-aligned conjoint analysis outperforms traditional hypothetical conjoint analysis in predicting actual behaviors. As a consequence, we also expect that the preference structure that incentive-aligned conjoint uncovers is different from the preference structure of hypothetical conjoint analysis. Specifically, as the contingent valuation method

[1]Another setting in which consumers have an incentive to act in a manner that reveals their "true" preference is Internet-based, mass-customization efforts that use consumers' preference ratings to personalize a decision support system. We thank an anonymous reviewer for highlighting this point.

[2]Monotonicity means that respondents must prefer more reward to less reward and not become satiated as the reward increases. This requirement is satisfied easily if money is used as the reward. Dominance requires that the respondents' utilities from the experiment come predominantly from the reward medium and that other influences are negligible. A salient reward must be great enough to satisfy the dominance requirement.

(Diamond and Hausman 1994; List 2001) suggests, budget constraints tend to be discounted in hypothetical situations, so we expect that price plays a more prominent role in incentive-aligned conjoint. In addition, socially desirable answers, such as lower preference for red meats or higher willingness to donate money to social causes, are less likely in incentive-aligned conditions (Camerer and Hogarth 1999), which may result in greater heterogeneity for socially desirables attributes. In other words, both participants who want and those who do not want to choose a socially desirable alternative will tend to choose that alternative during a hypothetical setting, but in an incentive-aligned setting, those who do not want to choose the socially desirable alternative will tend to reject that alternative, which will result in increased heterogeneity. In summary, we expect that incentive-aligned conjoint (1) outperforms traditional hypothetical conjoint in out-of-sample predictions of actual behaviors and (2) results in preference structures that give greater importance to price and may exhibit greater heterogeneity for socially desirable product attributes in the incentive-aligned condition.

To test these hypotheses, we conducted two field experiments. The first experiment, Study 1, had four conditions: the conventional (hypothetical) choice conjoint; the conventional (hypothetical) contingent valuation, or stated-price, method; and their corresponding incentive-aligned versions. The context we used was Chinese dinners. Study 2 had two conditions: the conventional (hypothetical) choice conjoint and its corresponding incentive-aligned version; Study 2 used snacks as its context. The results from the experiments demonstrate that conventional conjoint analysis exhibits hypothetical bias and that incentive-aligned choice conjoint significantly improves the out-of-sample predictions of actual purchase behaviors. The structure of partworths and the relative importance of various attributes also differ for incentive-aligned conditions compared with traditional hypothetical conditions.

STUDY 1: CHINESE DINNER SPECIAL

To examine the possibility of hypothetical bias, we required a research context that (1) represented a real economic decision for the participants (undergraduate and graduate students at a major U.S. university); (2) had a large set of attributes, each with several levels; (3) could generate new products easily through different combinations of the attributes; and (4) provided an easy means to induce realism in the product category as a result of ease of implementation. Chinese dinner specials meet these four criteria: (1) university students are interested in Chinese food, (2) Chinese dinner specials have a sufficient number of attributes, (3) these attributes can be used to generate product options, and (4) Chinese food can be prepared easily in real time and consumed by the participants right after the experiment. Therefore, Chinese dinner specials serve as the context for this study.

Qualitative Investigation

We first conducted qualitative investigations to understand the key attributes of Chinese dinner specials. Using an actual menu from the Chinese restaurant in which the experiment was conducted, we interviewed 10 undergraduate students to determine the attributes of a Chinese dinner special that were important to them and that they perceived as important to their peers. We then summarized the results and used them to develop a formal survey, which we gave to two groups of undergraduate marketing students (50 students in total) to obtain a better understanding of attribute importance and to assess the appropriate levels for the attributes. On the basis of the survey, we identified a total of eight important attributes associated with Chinese dinner specials: two attributes had two levels, five attributes had three levels, and one attributes had four levels (see Appendix A).

Experimental Design

We developed four different experimental treatments, namely, hypothetical choice conjoint, hypothetical contingent valuation method, incentive-aligned choice conjoint, and incentive-aligned contingent valuation method. Participants in the hypothetical treatments were not bound by their responses with regard to various tasks, but the participants in the incentive-aligned treatments were told that they had to live with their choices. (In this case, through certain random mechanisms, they were given one of the dinner specials they selected.)

In line with Lazari and Anderson (1994), to manage respondent fatigue, we used a fractional factorial design, which generated 108 profiles (Chinese meals). In the choice conditions, we therefore created three groups of 12 choice sets. Each choice set had three profiles (Chinese meals) and a "none of the above" option. We randomly assigned 9 participants to each of the three groups in the choice conditions (hypothetical and incentive aligned), which resulted in a total of 27 participants in each of the choice-based treatments. To ensure that the contingent valuation method was based on the profiles used for the choice method, we evenly divided the choice profiles (without the price attribute) into nine groups of 12 profiles. Then, for the hypothetical contingent and incentive-aligned contingent valuation methods, we randomly assigned 3 participants to each of the nine groups, which resulted in a total of 27 participants in each contingent valuation treatment. Because this was a between-subjects design, each participant appeared in only one of the four treatment groups.

The treatments constitute Part 1 of the experiment; the exact instructions given to the participants are included in Appendix A. Part 2 of the experiment, which was the same for all participants, was a holdout task. During the holdout task, each participant chose a meal from a menu of 20 different Chinese dinner specials (none of which appeared in Part 1; see Appendix B) or chose nothing at all (a total of 21 options). For all the participants, the choice made during Part 2 was real; that is, the restaurant served the meal they chose, and the cost of the meal was deducted from the $10 each participant received for the experiment. For participants in the incentive conditions, we used a random device to determine whether the meal they received came from Part 1 or Part 2 of the study. Finally, Part 3 comprised a brief exit survey that captured information about demographics, prior experience with Chinese food, and whether the participant understood the instructions in Parts 1 and 2.

Pilot Experiment

One of the reasons that conjoint analysis may not perform well is because the respondents are not serious about the purchase at the time of the study but answer the hypothetical questions as if they were. The incentive-aligned methods (incentive-aligned choice conjoint and incentive-aligned contingent valuation method), by definition and unlike the hypothetical methods, will not result in purchase

if the participant is not serious about purchasing, and the methods will automatically identify those participants. We conducted a pilot study to understand the existence and scope of such participants. Specifically, we recruited 41 participants for the hypothetical choice conjoint and used conditions that mirrored the settings common in conjoint studies. The differences between this pilot study and the main study are as follows: (1) We conducted the pilot study in a classroom, whereas we conducted the main study in a restaurant; (2) we conducted the pilot study during regular class time, whereas we conducted the main study at dinnertime; (3) the holdout task (Part 2) of the pilot study consisted of four choice tasks that were similar to those in Part 1, in that one of the four options was randomly chosen and the participant was given a coupon for his or her preferred meal plus the difference between $10 and the value of the coupon, whereas in the main study, participants chose from a menu of 20 specials; (4) participants received a coupon for a Chinese dinner special to be redeemed at a future date in the pilot study, whereas in the main study, participants consumed the meal at the end of the experiment; and (5) we did not screen participants in the pilot study, whereas in the main study, we instructed participants during recruiting that they should come only if they were interested in eating a Chinese meal.

The results confirm our assertion. For Part 1 of the pilot study, the hypothetical conjoint portion consisted of 12 choice tasks in which participants chose among four options (three different dinners or "none of the above"). Every participant selected at least one dinner from the 12 choice sets, and participants chose the "none of the above" option 25% of the time. For the four choice tasks in Part 2, participants chose among four options (three different dinners or "none of the above") and were told that the researchers would randomly select one of the choice tasks and that they would have to buy the chosen meal. In Part 2, slightly more than half (21 of 41) of the participants selected "none of the above" for each of the four choice tasks. Overall, the participants chose the "none of the above" option 67% of the time. Thus, in Part 1, the hypothetical condition, the respondents behaved as if they were interested in the Chinese dinner specials by preferring a meal to the "none of the above" option, but they behaved differently when asked to make a real purchase decision (Part 2).

Although the results provide evidence that the incentive-aligned approach induces different behavior, particularly among participants who are not serious about purchasing the product, the results do not answer a more important and insightful question: Does an incentive-aligned approach improve the quality of answers even when participants are serious about the purchase decision? In other words, even after participants who are not interested in the product are screened out (which constitutes a large percentage of participants in a typical commercial conjoint study), will the incentive-aligned approaches outperform the traditional approaches to predict actual purchase behavior, and will the approaches yield substantively different preference structures?

Recruiting Participants for the Main Experiment

A recruiting e-mail was sent to a mailing list maintained by the Experimental Economics lab at a large U.S. university. The e-mail stated that participants were needed for a market research experiment to be conducted during dinnertime (5:00 P.M.–6:00 P.M., Monday–Thursday) at a local Chinese restaurant. Participants would have a chance to purchase a Chinese dinner special of their choice during the experiment, which would be cooked by the restaurant and be ready for consumption by the end of the experiment. The e-mail explicitly stated that only people interested in eating at this restaurant that evening, provided that they could find the right meal at a good price, should participate. Each participant would be paid $10 for their participation, part of which they could use to purchase a Chinese dinner special. A total of 108 undergraduate and graduate students participated in the main experiment, with an average of 12 students per session. Only 11 of the 108 participants did not choose to buy a meal in the holdout task (3 from the incentive-aligned contingent valuation method group, 3 from the incentive-aligned choice conjoint group, and 5 from the hypothetical choice conjoint group).

Experimental Procedure

We scheduled the data-collection sessions from 5:00 P.M.–6:00 P.M., and we conducted the incentive-aligned and nonaligned versions (incentive-aligned choice conjoint or incentive-aligned contingent valuation method) on the same day or on successive days to minimize sample variations. Consistent with practices in experimental economics, we conducted the experiment by following a written procedure, which we describe subsequently.

For the hypothetical choice conjoint and hypothetical contingent valuation method, the participants completed the consent form and Parts 1, 2, and 3 in sequence, and experimenters collected each completed part before the next part was distributed. The restaurant served the meal that was selected in Part 2. Cash reimbursements ($10 less the cost of the dinner) were paid on completion of Part 3. Participants were dismissed after they were paid, and as did the participants in the two incentive-aligned conditions, most ate the dinner in the restaurant, though a few took the dinner home.

The procedure we used for the incentive-aligned contingent valuation method treatment is called "BDM" (Becker, DeGroot, and Marschak 1964). This procedure has been used widely in economics and was introduced recently into marketing to measure willingness to pay at the point of purchase (Wertenbroch and Skiera 2002). In addition to the consent form, participants were given written instructions that stated that they would have two chances to select a Chinese dinner special, once in Part 1 and once in Part 2. A random device would be used to decide which selection they would actually receive. Participants were then given Part 1. After completing Part 1, each participant went through a two-step process in which they chose a dinner special by randomly selecting a number between 1 and 12; then, they randomly drew a piece of paper from an envelope that gave the price for that special (the possible prices ranged from $.25 to $8 in $.25 increments). In line with the BDM procedure, if the randomly selected price was equal to or lower than the price stated by a participant, the participant received the dinner special and needed only to pay the randomly selected price. The participant did not receive a dinner special if the randomly selected price was higher than the stated price. This procedure ensured that it was in

the best interest of the participants to state their true valuation for a dinner special. After completing Part 1, participants were given Part 2. After completing Part 2, each participant randomly selected a ball from a container with two balls (labeled Part 1 and Part 2) to decide which dinner choice they would actually consume. Finally, participants were given Part 3. They were given the cash balance ($10 less the cost of the special) and were dismissed after handing in Part 3.

The procedure for incentive-aligned choice conjoint was the same as that for the incentive-aligned contingent valuation method, except that the dinner special that participants chose in Part 1 was their preferred option from a randomly chosen choice set (which resulted from choosing a random number between 1 and 12), and they did not need to select a price randomly, because they paid the price of the selected dinner. The randomizing mechanism we used to determine which of the participants' choices was fulfilled is called "random lottery procedure" and is used widely in experimental economics (Starmer and Sugden 1991). Investigators use this mechanism to minimize reference point and wealth effects while collecting a large amount of data. For this study, it also ensures realism, in that a participant is unlikely to eat more than one Chinese meal at a given time in his or her real consumption episode.

Estimation Procedure

To provide the best possible comparison between the incentive-aligned and hypothetical approaches, we used state-of-the-art models and estimation methods to assess participants' preferences, in-sample fit, and out-of-sample predictions. To analyze the rating data (which resulted from the stated price and BDM contingent valuation methods), we used a random-effects hierarchical Bayesian regression model that is similar to the model that Lenk and colleagues (1996) specify. The regression likelihood is as follows:

$$y_{it} =_d N\left(\beta_i^T x_{it}, \sigma^2\right), \tag{1}$$

where y_{it} is the tth contingent valuation given by the ith participant, $=_d$ is equal in distribution, N is the normal density, x_{it} describes the tth meal evaluated by the ith participant, and β_i is a vector of contingent valuation partworths for the ith participant. We assumed a hierarchical shrinkage specification for the individual partworths, where, a priori,

$$\beta_i =_d N\left(\bar{\beta}, \Lambda\right). \tag{2}$$

This specification allows for individual-level partworth estimates β_i but still permits an estimate of the aggregate or average partworths $\bar{\beta}$, as well as an estimate of the amount of heterogeneity for each partworth Λ. On the basis of in-sample and out-of-sample model performance, we used a simplified version of the model by assuming that Λ is a diagonal matrix.[3] Furthermore, we assumed vague conjugate priors for $\bar{\beta}$, Λ, and σ^2.

To analyze the choice data, we used a random-effects hierarchical Bayesian multinomial logit model that is similar to the model that Allenby, Arora, and Ginter (1998) specify.[4] The probability that the ith participant chooses the jth alternative from the tth choice set is given by

$$\Pr\left(y_{it} = j\right) = \frac{\exp\left\{\tilde{\beta}_i^T x_{itj}\right\}}{\sum_{\ell} \exp\left\{\tilde{\beta}_i^T x_{it\ell}\right\}}. \tag{3}$$

Again, we assumed a hierarchical shrinkage specification for the individual partworths, where, a priori,

$$\tilde{\beta}_i =_d N\left(\bar{\tilde{\beta}}, \tilde{\Lambda}\right). \tag{4}$$

As with the regression model, we were able to estimate individual-level partworth parameters, average partworth parameters, and the partworth heterogeneity. Again, on the basis of in-sample and out-of-sample model performance, we assumed that $\tilde{\Lambda}$ was a diagonal matrix.[5] Furthermore, we assumed vague conjugate priors for $\bar{\tilde{\beta}}$ and $\tilde{\Lambda}$. (Note that unlike the choice partworth parameters $\bar{\tilde{\beta}}$, the contingent valuation partworth parameters $\bar{\beta}$ do not have a price sensitivity element, because the contingent valuation is given in terms of the price that the participant is willing to pay for the proposed product.)

We tested a range of different prior values to ensure that the reported results were invariant to the prior specification. In addition, we assessed the convergence properties of the Markov chain Monte Carlo analysis to ensure that the algorithm had converged to the target density, as induced by the model specification, before making marginal summaries of the posterior density.

Results

We assessed in-sample goodness of fit for the logit models by calculating the percentage of times the model accurately identified the choice from the four alternatives (the hit rate), among which one alternative was "none of the above." In addition, we estimated the marginal probability of the data given a model (reported on a log scale) using the estimation method that Newton and Raftery (1994) provide, which can be used to form Bayes factors. With the logit model, the hypothetical choice conjoint, resulted in a better in-sample fit than did the incentive-aligned choice conjoint (i.e., a hit rate of 32% and 41% for the incentive-aligned choice conjoint and hypothetical choice conjoint, respectively). We obtained log-marginal probability values of –1109 and –788 for incentive-aligned choice conjoint and hypothetical choice conjoint, respectively.[6] The R-square values of .97 and .96 for incentive-aligned contingent valuation conjoint and hypothetical contingent valuation conjoint, respectively, show good overall model fit (the log-marginal probability values for incentive-aligned contingent

[3]We considered both a diagonal and a full matrix version of each model and found that the in-sample fit criteria (Bayes factors) moderately supported a full matrix version of the model but that the out-of-sample predictive performance favored a diagonal matrix version of the model.

[4]As pointed out by one of the reviewers, Sawtooth's hierarchical Bayesian software implements a model that is similar to the model that we implemented, which provides other researchers ready access to this model.

[5]As with the regression model, we considered both a diagonal and a full matrix version of each model. For the choice models, we found that both the in-sample fit criteria (Bayes factors) and the predictive performance criteria favored a diagonal matrix version of the model.

[6]The in-sample fit criteria are only reported for completeness and are based on the same model applied to different data sets. Note that this approach differs from existing literature in which such comparison is usually between different models applied to the same data set.

valuation conjoint and hypothetical contingent valuation conjoint are –401 and –368, respectively).

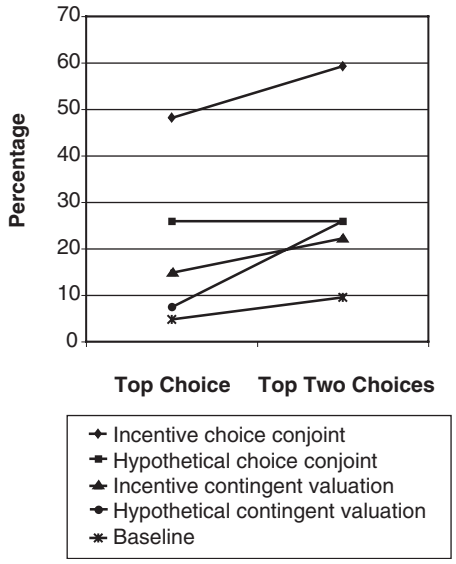
The incentive-aligned data result in significantly better out-of-sample predictions than the hypothetical results, which supports our primary hypothesis (see Figure 1). The incentive-aligned choice conjoint forecasts the correct purchase 48% of the time, which represents a more-than-tenfold improvement over the naive forecast rate of approximately 5% and is almost twice as good as the hypothetical choice conjoint forecast, which is correct 26% of the time. The results for the top two choices are equally impressive, with 59% and 26% correct predictions in the incentive-aligned choice and hypothetical choice conditions, respectively. The incentive-aligned contingent valuation method analysis results in better out-of-sample forecasts than the hypothetical contingent valuation method, with 15% compared with 7% correct.[7] We plot these out-of-sample numbers along with the naive baseline predictions in Figure 1. The superior out-of-sample predictive performance of incentive-aligned methods is evident from Figure 1. Figure 1 also provides support for the superior performance of choice methods when compared with contingent valuation methods. We discuss several possible explanations for this result in favor of the choice methods in the next section.

In addition to resulting in better out-of-sample forecasts, the aggregate parameter estimates based on the incentive-aligned tasks are markedly different from the estimates of the non-incentive-aligned tasks (see Table 1).[8] Perhaps the most striking finding is that, on average, the participants from the incentive-aligned task are more price sensitive (–1.59) than are participants from the non-incentive-aligned task (–.99), and their price sensitivity is spread over a larger range, as is indicated by the difference in the heterogeneity of the slopes (.44 and .20 for the incentive-aligned and non-incentive-aligned tasks, respectively). This finding is consistent with a notion in experimental economics literature, which suggests that participants discount budget constraints in hypothetical conditions (Diamond and Hausman 1994; List 2001). In addition, the average importance of the size of the meal, given by the Quart partworth, is almost three standard deviations above zero (2.8 = 1.29/.46) for participants in the hypothetical conjoint task and slightly greater than one standard deviation above zero (1.03 = .40/.39) for the incentive-aligned task, which indicates that when the task is incentive aligned, the effect of the size of the meal becomes negligible. An explanation is that though the additional quantity warrants higher valuation in theory, participants may be less likely to associate it with similarly high valuation in a real purchase experience because they know that they are unlikely to eat the additional amount (or, if they do, that it will not do them any good).

Another notable insight is that the levels of individual heterogeneity, as given by the diagonal elements of Λ and $\tilde{\Lambda}$, are markedly different for the incentive and hypothetical treatments. Consistent with assertions in experimental economics literature (e.g., Camerer and Hogarth 1999), the heterogeneity for socially desirable alternatives should increase when participants are presented with a real decision than when they are presented with a hypothetical decision. For example, consider the heterogeneity of the partworths for chicken and shrimp (which may be considered healthier alternatives to beef). The heterogeneity of the partworths of chicken and shrimp is smaller for the hypothetical choice conjoint than for the incentive-aligned choice conjoint.[9] This result suggests that in the hypothetical setting, some participants conformed to the social norm of selecting healthy alternatives, whereas many participants who followed this norm in the hypothetical setting likely abandoned it in the incentive-aligned setting because they had to live with (or, in this case, eat) their choice, thereby increasing the heterogeneity in the incentive-aligned condition. However, we observe a similar shift in heterogeneity of

[7]These results seem comparable to previous research that has used actual purchase decisions as validation. For example, for natural experiments involving MBA job choices, Wittink and Montgomery (1979), Srinivasan (1988), and Srinivasan and Park (1997) find that the first preference predictions range from 64% to 76%, compared with random choice results of 26% to 36%. Similarly, Dahan and colleagues (2002) report a 50% to 59% correct predictive performance for five new-to-market laptop computer bags, compared with a random choice outcome of 20%. With incentive-aligned choice conjoint, the predictions were correct 48% of the time compared with the random choice outcome of less than 5% (1 of 21).

Figure 1
PREDICTIVE PERFORMANCE FOR HOLDOUT TASK: STUDY 1



[8]Because the choice models result in much better out-of-sample predictive ability, we focus our discussion on the difference between the parameter estimates from these analyses. We include the contingent valuation results to be consistent with the practice in experimental economics of reporting all study results.

[9]With an independent sample of respondents from the same population as that in Study 1, we collected data on the social desirability of chicken, shrimp, and beef. For each attribute, the participants responded to the following three statements on a 1–7 "agree–disagree" scale: (1) I think it is socially desirable to eat beef, (2) my friends and family would agree that it is socially desirable to consume beef, and (3) there is a general perception that consuming beef is socially desirable. Because the Cronbach's alpha for the three scales exceeded .84, we averaged the three items to construct social desirability measures for chicken, shrimp, and beef. Paired sample t-tests showed that for the 37 respondents, eating chicken was more socially desirable than eating beef ($t = 2.91$, $p < .01$), eating shrimp was more socially desirable than eating beef ($t = 2.12$, $p < .05$), and there was no difference in the social desirability of chicken and shrimp ($t = –.63$, $p > .53$). The results show that chicken and shrimp are more socially desirable among the participant population than is beef.

Table 1

STUDY 1: SUMMARIES OF PARAMETER ESTIMATES FOR CHOICE CONJOINT METHOD (RANDOM-EFFECTS LOGIT ANALYSIS)

Parameter[b]	*Intercept*	*Hot and Sour Soup*	*Egg Drop Soup*	*Brown Rice*	*Noodles*	*Szechwan Sauce*	*Sweet and Sour Sauce*	*Standard Vegetables*	*Exotic Vegetables*	*Beef*	*Chicken*	*Shrimp*	*Pork Spring Roll*	*Quart*	*Price ($)*
Incentive Conjoint															
Slope (mean[a])	2.62	–.76	.43	.28	.42	–.17	.01	1.46	.23	2.66	3.42	2.48	–.60	.40	–1.59
Slope (standard deviation[a])	.65	.59	.47	.48	.54	.59	.52	.44	.52	.49	.64	.61	.47	.39	.19
Slope (heterogeneity[a])	.82	7.15	3.11	.99	1.18	4.91	3.13	1.02	2.99	2.07	6.14	5.44	3.74	1.26	.44
Hypothetical Conjoint															
Slope (mean[a])	.67	.56	.55	.58	–.09	.38	.14	1.85	1.00	3.65	4.11	3.28	–.80	1.29	–.99
Slope (standard deviation[a])	.57	.45	.50	.52	.52	.41	.36	.44	.55	.91	.93	.91	.51	.46	.26
Slope (heterogeneity[a])	1.55	1.58	3.46	3.05	3.44	1.78	.89	.49	3.84	1.76	2.68	4.53	2.90	2.93	.20

[a]Posterior mean and standard deviation of $\bar{\bar{\beta}}$; heterogeneity and posterior mean of diagonal of $\tilde{\Lambda}$. Note that a simple t-statistic calculation, (posterior mean)/(posterior standard deviation), gives guidance with respect to whether the marginal posterior density for each parameter is far from zero or whether the estimates have influence.

[b]Soups are compared with no soup, brown rice and noodles are compared with white rice, sauces are compared with brown sauce, vegetables are compared with no vegetables, meats are compared with no meat, pork spring roll is compared with vegetable spring roll, and quart size is compared with pint size.

beef, even though the magnitude of the change in heterogeneity is much smaller (the magnitude of change in the heterogeneity of beef is approximately .3 compared with 3.5 for chicken and .9 for shrimp). Nonetheless, these notions on social desirability need further exploration.

Furthermore, the heterogeneity of exotic vegetables (a riskier option than standard vegetables) reduces from 3.84 for hypothetical choice conjoint to 2.99 for incentive-aligned choice conjoint. This decrease in heterogeneity is accompanied by a decrease in the average partworth weight from 1.00 in the hypothetical choice conjoint to .23 in the incentive-aligned choice conjoint. Some participants in the hypothetical choice conjoint may have chosen exotic vegetables to try new items, but in the incentive-aligned choice conjoint, the participants seem to be more risk averse and do not prefer the novel exotic vegetable attribute. Similar insights hold for the regression-based parameter estimates (see Table 2).

According to the pilot study, it is almost certain that had we not screened out participants who were not serious about the purchase decision, the incentive-aligned contingent valuation method would have performed better than the hypothetical contingent valuation method and that the incentive-aligned choice conjoint would have performed better than the hypothetical choice conjoint. It is informative that the incentive-aligned methods outperform their hypothetical counterparts, even after participants who are not serious about purchase have been excluded from the study. This result suggests that it is important to find ways to use incentives that are aligned with purchase behavior in conjoint studies.

STUDY 2: SNACK COMBO

To test the robustness of the Study 1 findings, especially the strong increase in out-of-sample forecasting accuracy, we conducted a second study that focused on the choice conjoint. The task context for this second study was a snack combo. Specifically, participants identified their preferences for a snack combo that could have one (or none) of the following four attributes: a drink (water, Coca-Cola, Diet Coke, iced tea, or orange juice), a cookie (peanut butter, chocolate fudge, or oatmeal raisin), a Korean cereal bar (white, dark, or strawberry chocolate), and a piece of fruit (banana or apple). Each snack combo was priced at one of three levels ($1.00, $1.75, or $2.50).

We chose the snack combo context because, as were the Chinese dinners in Study 1, a snack combo is a familiar context for our participants and because we could choose multiple levels easily for each attribute. However, the snack combo context is also appealing because it differs from the Chinese dinner special context in several ways. First, because the attributes in the snack combo are different snack categories (e.g., cookies and fruit), they are less likely to interact with one another than the attributes in the Chinese dinner special (e.g., the value of a particular sauce is likely to interact with a specific meat or vegetable). Second, there is neither uncertainty nor inconsistency with regard to the product quality in Study 2. The quality of the Chinese dinner special depended on the restaurant and the cook, which added to the uncertainty of the conjoint study and the holdout task in Study 1. In contrast, the snack combo consisted of items that could be bought at a grocery store and, except for the fruit, were prepackaged, brand-name products.

To assess the robustness of the results, we changed the study design in several ways. First, we included an unfamiliar attribute: the Korean cereal bar. With this attribute, we can test risk preference and willingness to try new things explicitly. Second, we used an orthogonal design to generate a total of 27 conjoint tasks (each task had four choices, three snack combos, or "none of the above"), which enabled us to ask every participant to complete all 27 tasks, whereas in Study 1, we divided the total tasks into three groups, and each participant evaluated only one-third of the profiles. Therefore, in Study 1, we needed to pool information across participants to obtain the parameter estimates. Third, we divided the participants into two sessions. Although the experiments for each session were the same, the 30 snack combos in the holdout task appeared in different sequences to minimize the impact of any potential order effects.

Experiment

We visited stores and cafés frequented by the participants to identify popular brands of drinks and cookies and to obtain reasonable price levels. To ensure that our Study 1 risk preference results could be generalized to a context in which the attributes were completely new to the participants (in contrast to the exotic vegetables in Study 1), we visited a local Oriental-foods store and chose three varieties of a Korean cereal bar. According to the store owner, the bar had just been introduced to the local market and was not available in mainstream channels (e.g., chain grocery store, which we verified). We used the built-in routine in SPSS to generate 27 conjoint tasks (with three snack combo profiles in each task, for a total of 81 different snack combos) and another 30 unique snack combos for the holdout task.

We recruited 59 senior undergraduate students from the same U.S. university as in Study 1. We conducted the experiment over two sessions, and we randomly assigned participants in each session to either the incentive-aligned choice conjoint or the hypothetical choice conjoint. We brought snacks into the room before the start of each session. We packaged food items in each snack combo in a large freezer bag, and we stored the drinks in a cooler with ice. Participants knew precisely the brand and quality of snack they could buy.

The experimental procedure for both the incentive-aligned and the hypothetical choice conjoint conditions was similar to that in Study 1 (Appendix C). After the participants completed the conjoint task, we gave them instruction for the holdout task (Appendix D) in which they selected 1 of the 30 possible snack combos or no snack combo. The participants in the hypothetical choice conjoint condition received $3, out of which they could buy any of the 30 snack combos in the holdout task. We also gave the participants in the incentive-aligned choice conjoint condition $3, and they had two opportunities to buy a snack combo, one in the conjoint task (selected using a random lottery, as in Study 1) and one in the holdout task. We randomly selected one of the opportunities for each participant, and we then fulfilled his or her choice. At the completion of the experiment, we gave the participants the snack combo of their choice, and they received the balance of $3 less the price of the snack combo they chose.

Table 2

STUDY 1: SUMMARIES OF PARAMETER ESTIMATES FOR CONTINGENT VALUATION METHOD (RANDOM-EFFECTS REGRESSION ANALYSIS)

Parameter[b]	*Intercept*	*Hot and Sour Soup*	*Egg Drop Soup*	*Brown Rice*	*Noodles*	*Szechwan Sauce*	*Sweet and Sour Sauce*	*Standard Vegetables*	*Exotic Vegetables*	*Beef*	*Chicken*	*Shrimp*	*Pork Spring Roll*	*Quart*	*Price ($)*
Incentive Conjoint															
Slope (mean[a])	2.36	.27	.55	.05	.13	–.05	.16	.48	.37	1.14	1.06	.89	.19	.64	.66
Slope (standard deviation[a])	.38	.29	.30	.29	.29	.27	.27	.27	.27	.35	.30	.32	.28	.27	.10
Slope (heterogeneity[a])	2.06	1.23	1.04	1.01	1.05	1.02	1.02	1.06	1.11	1.69	1.23	1.61	1.64	1.18	—
Hypothetical Conjoint															
Slope (mean[a])	2.88	.34	.12	.02	–.02	.07	.09	.45	.21	1.23	1.28	1.67	.05	.86	.55
Slope (standard deviation[a])	.32	.27	.28	.28	.26	.26	.25	.25	.30	.31	.27	.28	.21	.25	.10
Slope (heterogeneity[a])	1.23	1.03	1.14	.98	.94	.97	.96	.96	1.81	1.22	.91	1.07	.77	1.00	—

[a]Posterior mean and standard deviation of $\bar{\beta}$; heterogeneity and posterior mean of diagonal of Λ. Note that a simple t-statistic calculation, (posterior mean)/(posterior standard deviation), gives guidance with respect to whether the marginal posterior density for each parameter is far from zero or whether the estimates have influence.

[b]Soups are compared with no soup, brown rice and noodles are compared with white rice, sauces are compared with brown sauce, vegetables are compared with no vegetables, meats are compared with no meat, pork spring roll is compared with vegetable spring roll, and quart size is compared with pint size.

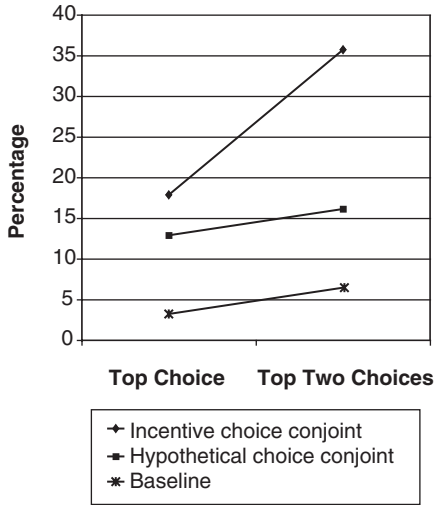
Results

Using the same estimation approach in terms of in-sample hit rate and log-marginal probability,[10] we find that the incentive-aligned choice conjoint condition (hit rate: 39%; log-marginal probability: –2619) outperforms the hypothetical choice conjoint condition (hit rate: 32%; log-marginal probability: –2795) for the snacks data set, in contrast to the in-sample results in Study 1. For out-of-sample predictions, the incentive-aligned choice conjoint condition (top choice: 18%; top two choices: 36%) also outperforms the hypothetical choice conjoint condition (top choice: 13%; top two choices: 16%), in support of our hypothesis that incentive-aligned conjoint predicts actual purchase behavior better than does hypothetical conjoint. In Figure 2, we plot the out-of-sample predictive performance of incentive-aligned choice conjoint and hypothetical choice conjoint relative to the naive baseline prediction (i.e., 1 of 31 for top choice and 2 of 31 for top two choices). This figure confirms the findings of Study 1 that incentive-aligned choice conjoint is superior to its hypothetical counterpart.

As we did in Study 1, we find markedly different aggregate parameter estimates for incentive-aligned and hypothetical tasks. Participants were more price sensitive in the incentive-aligned condition (–4.18) than in the hypothetical condition (–2.61). Price heterogeneity appears much higher in the incentive-aligned condition (5.60) than in the hypothetical condition (.90). Similarly, as we show in Table 3, the slope and heterogeneity parameters differ for other attribute-level combinations. For the novel attribute (i.e., Korean cereal bars), we find lower slope and heterogeneity in the incentive-aligned condition than in the hypothetical condition. The results suggest that in a hypothetical setting, participants tend to overstate their levels of risk preference and willingness to try new things than they do when they are in a real purchasing environment. Indeed, this conjecture about risk preferences and other conjectures about social desirability bias merit further scrutiny.

[10]Again, the in-sample fit criteria are only reported for completeness and are based on the same model applied to different data sets. Note that this is different from the existing literature, which usually compares different models applied to the same data set.

Figure 2
PREDICTIVE PERFORMANCE FOR HOLDOUT TASK: STUDY 2



GENERAL DISCUSSION

We rely on the literature of induced value theory (Smith 1976) and hypothetical bias (Diamond and Hausman 1994) to imply that contemporary conjoint-based methods may poorly identify consumer preference. Specifically, traditional conjoint techniques deal with hypothetical situations, and experimental economics literature suggests that a hypothetical setting does not motivate participants sufficiently to reveal their "true" preferences. We propose overcoming this weakness with incentive-aligned conjoint methods, specifically, the use of incentive-aligned versions of choice-based conjoint and the contingent valuation method.

Our results across the two studies provide strong evidence in favor of the incentive-aligned choice conjoint in terms of out-of-sample predictions of purchase decisions. We also find that participants in the incentive-aligned choice conjoint condition have systematically different preference structures than do participants in other conditions. Notably, the participants in the incentive-aligned choice conjoint condition, as compared with those in the hypothetical condition, have higher price sensitivity, exhibit lower risk seeking and willingness to try new things, and are less prone to socially desirable behaviors.

The benefits of incentive-aligned conditions for marketing researchers are evident and substantial. Marketing researchers should use incentive-aligned conjoint, and marketing academics should study incentive-aligned conjoint further to better understand the linkages between stated (hypothetical) and revealed (incentive-aligned) preferences.

The strong findings across both studies in favor of incentive alignment for individual decision making suggest that marketing academics should also investigate incentive alignment for both group decision making in consumer markets for which group norms can play an important role (e.g., family decision making) and organizational decision making in business markets.[11] Furthermore, academics should explore whether use of hypothetical and incentive-aligned conditions in combination improves out-of-sample predictions. For example, a conceivable research design could involve three stages: (1) hypothetical conjoint, (2) incentive-aligned conjoint, and (3) a holdout task.[12] In addition to providing potential prediction benefits, such a research design may suggest the manner in which people change their preferences. Comparative studies with process measures for both hypothetical and incentive conditions could illuminate the differences in the decision-making tasks in the two conditions.

From a managerial perspective, the most relevant issue is to identify and test various implementation strategies that align peoples' incentives for a wide range of products. We believe that the basic guidelines can be implemented for various product categories, especially when the attributes

[11]In both studies, we attempted to ensure that the experiments focused on individual decision making. The participants were not allowed to communicate and/or see one another's decision in either study. In the Chinese dinner study, we required participants who arrived in a group to sit at different tables.

[12]We thank an anonymous *JMR* reviewer for this suggestion.

Table 3

STUDY 2: SUMMARIES OF PARAMETER ESTIMATES FOR CHOICE CONJOINT METHOD (RANDOM-EFFECTS LOGIT ANALYSIS)

Parameter[b]	*Intercept (Coke)*	*Diet Coke*	*Orange Juice*	*Ice Tea*	*Water*	*Oatmeal Cookie*	*Chocolate Fudge Cookie*	*Peanut-butter Cookie*	*White Chocolate Korean Snack Bar*	*Strawberry Korean Snack Bar*	*Dark Chocolate Korean Snack Bar*	*Banana*	*Apple*	*Price ($)*
Incentive Conjoint														
Slope (mean[a])	2.11	–1.84	.16	1.10	1.43	1.87	2.27	2.15	.35	.38	.47	.88	.62	–4.18
Slope (standard deviation[a])	.92	1.24	.63	.41	.78	.58	.50	.51	.30	.36	.35	.57	.53	.57
Slope (heterogeneity[a])	12.96	40.37	8.56	2.26	11.14	6.49	3.73	3.59	.50	1.08	2.03	5.71	4.69	5.60
Hypothetical Conjoint														
Slope (mean[a])	1.48	–3.94	–.44	–.29	–.13	.57	2.22	2.80	.63	.73	.79	1.44	.85	–2.61
Slope (standard deviation[a])	.59	.90	.49	.59	.63	.51	.49	.54	.41	.34	.36	.30	.26	.25
Slope (heterogeneity[a])	5.79	18.79	5.52	8.57	9.35	4.49	4.82	5.95	3.16	1.97	2.56	1.32	.91	.90

[a]Posterior mean and standard deviation of $\bar{\bar{\beta}}$; heterogeneity and posterior mean of diagonal of $\tilde{\Lambda}$. Note that a simple t-statistic calculation, (posterior mean)/(posterior standard deviation), gives guidance with respect to whether the marginal posterior density for each parameter is far from zero or whether the estimates have influence.

[b]Cookies are compared with no cookie, Korean snacks are compared with no Korean snacks, and fruits are compared with no fruit.

are well understood and the product can be made available. However, a serious implementation challenge remains for expensive or complex products (e.g., automobiles) and for novel products for which a prototype may not exist. In the case of expensive products, it may not be cost effective to offer a real product to each study participant. Therefore, market researchers must ensure that the potential earning is greater than the opportunity cost. For example, an automobile company interested in an incentive-aligned conjoint study only needs to offer one or two automobiles to one or two randomly selected participants, as long as the potential earning (value of the car multiplied by the likelihood of winning) is greater than each person's opportunity cost. However, key to this lottery approach is that the winning participant must receive the car that matches his or her stated preference in the study. Another challenge is to obtain a sample that is interested in the product at the time of the exercise. By recruiting only participants who self-select (as we did in the Chinese dinner special context) as being interested in the product at the time of the exercise, this challenge may be met. In addition, to help eliminate participants who are only interested in monetary payments, part of the compensation should be the product rather than cash.

For novel products, incentive alignment can be truly difficult, as can other issues such as forecasting. Borrowing a page from Urban and colleagues (1997), virtual representations, such as information acceleration, may enable researchers to collect data (though hypothetical) on new product concepts. By tabulating the necessary adjustments to make hypothetical conjoint consistent with incentive-aligned conjoint for various existing product categories, researchers could then conduct hypothetical conjoint exercises and adjust the preference structure by using products similar to the new product. Thus, calibrating hypothetical conjoint with incentive-aligned conjoint becomes a critical research issue.

In conclusion, for managers, the guiding principle is simply to align respondents' interests to the actual decision outcome and to ensure that the incentive is not trivial compared with their opportunity cost. For example, the participant could receive a coupon redeemable only for the preferred choice he or she made during the study. As is apparent by the Nobel Prize given to Vernon Smith for his pioneering work in experimental economics, and as our results demonstrate, incentive alignment can significantly improve conjoint analysis. Our research suggests that marketing researchers should use incentive alignment to assess consumer preferences and should continue to conduct further research in the context of conjoint and other experiments that pertain to consumer behavior.

APPENDIX A: EXPERIMENT INSTRUCTION FOR STUDY 1 (PART 1)

In Appendix A, we provide the exact instructions for the experiment conducted in Study 1. Each participant first read the "General Instruction," followed by "Specific Instructions" for the experimental condition to which he or she was assigned. We also include one conjoint task example and one contingent valuation task example.

General Instruction (for all Experimental Conditions)

You are about to participate in an experiment designed to understand how people like you value a variety of different Chinese meals (dinner specials). We would ask that you pay close attention to the different meals being offered and determine an accurate value for each meal.

You will receive $10 for participating in this experiment.

Before proceeding with the remainder of the study, we would like to familiarize you with the type of meals that you will be considering. Each dinner special will be described by eight attributes: (During the course of the study, you may wish to refer to the following table.)

Attribute	*Levels*			
Soup	No soup	Hot and sour soup		Egg drop soup
Rice/noodle	White rice	Brown rice		Noodles
Sauce	Brown sauce	Szechwan sauce (hot and spicy)		Sweet and sour sauce
Vegetables	No vegetables	Standard vegetables[a]		Tofu and exotic vegetables[b]
Meat	No meat	Beef	Chicken	Shrimp
Spring roll	Vegetable spring roll			Pork spring roll
Quantity	Pint			Quart (two pints)
Price	$3.99	$4.99		$5.99

[a]Standard vegetables include common vegetables that you would find in a supermarket (e.g., broccoli, green peppers, green beans, mushrooms, snow peas, etc.).

[b]Exotic vegetables include vegetables usually found in Asia (e.g., bamboo shoots, Shanghai bokchoy, green mustard, Chinese egg, among others).

Specific Instructions

Non-incentive-aligned conjoint. You will be shown 12 sets of three meals. For each set of three meals, imagine that you were asked to choose between no meal and one of these three different meals at the stated price. Select the most attractive option (which could include not selecting any of the meals).

Incentive-aligned conjoint. You now have an opportunity to select a Chinese dinner special and have it cooked here in the restaurant before you leave. Here is how it works. You will be shown 12 sets of three meals. For each set of three meals, please choose between no meal and one of these three different meals at the stated price (which could include not selecting any of the meals). After you complete your selection, we will randomly choose a set from these 12 sets, and your choice for that set will be fulfilled. If you have selected no meal for that set, you will be given $10 cash to take home; if you have selected a meal for that set, the restaurant will cook that meal for you, and

you will be given $10 minus the price of that meal as stated.

An Example of the Conjoint Task

Choice Set 1			
Attributes	*Meal 1*	*Meal 2*	*Meal 3*
Soup	Hot and sour	Hot and sour	Hot and sour
Rice/noodle	Noodles	White rice	Brown rice
Sauce	Brown sauce	Szechwan sauce	Sweet and sour sauce
Vegetables	No vegetables	Standard vegetables	Tofu and exotic vegetables
Meat	No meat	Beef	Chicken
Spring roll	Vegetable	Pork	Pork
Quantity	Pint	Pint	Quart
Price	$3.99	$5.99	$5.99

Please indicate which meal you would choose (circle your choice).

- Meal 1
- Meal 2
- Meal 3
- None of the above

Non-incentive-aligned contingent valuation. You will be shown 12 different meals. For each meal, imagine that you were to state the price you would be willing to pay for the meal, and then write down the price for the meal.

Incentive-aligned contingent valuation (BDM procedure). You will be presented with 12 meals in this part of the experiment; please tell us the highest price you would be willing to pay for each meal. After you state your maximum prices for all 12 meals, we will determine which meal you will actually buy and how much you will pay for it based on the following procedure.

First, you will be asked to draw a ball from an envelope, which contains 12 balls labeled 1–12. You will be able to purchase the meal that has the same number as the one written on the ball.

Next, you will be asked to draw a ball/ticket from another envelope. The balls/tickets are labeled with different prices; the range of these prices is reasonable for a Chinese dinner special, neither too high nor too low. If you draw a price that is less than or equal to the price you choose for that meal, you will have to buy the special for the price you drew from the envelope. If the price you draw is greater than the price you choose, you will not be able to buy that particular meal. This procedure ensures that it is best for you to truthfully reveal the maximum price you are willing to pay for each meal. If you choose a price that is high, you may actually have to pay that high price. If you choose a price that is low, you may be disappointed if you can't buy the meal at the low price because you drew a price that is higher than the price you choose but lower than your "true" price. Note that you cannot influence the purchase price with the price you choose. Because you draw the purchase price from the envelope, it is completely random and independent of whatever you choose. For example, if you state that your maximum price for Meal 3 is $7.24, and you draw a price of $2.30, you will receive Meal 3 for $2.30 and receive the remaining $7.70 ($10 – $2.30) in cash.

An Example of the Contingent Valuation Task

Meal 1	
Attributes	*Description*
Soup	Egg drop
Rice/noodle	Noodles
Sauce	Brown sauce
Vegetables	Tofu and exotic vegetables
Meat	Beef
Spring roll	Vegetable
Quantity	Pint

Please indicate how much you would be willing to pay for the above meal: $_____.

APPENDIX B: HOLDOUT TASK FOR STUDY 1 (PART 2)[13]

Now you have to choose a single meal out of the 20 possible dinner specials presented in this part of the experiment. You may choose to select none of the 20 meals and thereby elect not to purchase. If you choose a meal, you will have to pay for it. For example, if you select to purchase the third dinner special meal at $4.99, we will give you $5.01 ($10 – $4.99) in cash, and the restaurant will cook that meal for you while you wait.

Please examine the meals on the next two pages [see Table B1][14] and indicate your choice below:

- Choose Meal No. _____.
- Do not wish to purchase any of the 20 dinner specials _____.

APPENDIX C: EXPERIMENTAL INSTRUCTIONS FOR STUDY 2

In Appendix C, we provide the exact instructions for the experiment conducted in Study 2. Each participant read the specific instructions for the experimental condition to which he or she was assigned. We also include one conjoint task example.

Instruction for Hypothetical (Traditional) Conjoint

You are about to participate in an experiment designed to understand how people like you value a variety of snacks. We would ask that you pay close attention to the different snacks being offered and determine an accurate value for each meal. You will be shown 27 sets of three snack combos. For each set of three combos, imagine that you were asked to choose between no snack and one of these three different snacks at the stated price. Select the most attractive option (which could include not selecting any of the snacks).

Instruction for Incentive-Aligned Conjoint

You are about to participate in an experiment designed to understand how people like you value a variety of snacks.

[13]The participants in the incentive-aligned conditions were informed that a random device would be used to determine the meal that they would consume, which would come from their choice in Part 1 or Part 2 of the experiment.

[14]In the Appendices B–D, the original instructions have been slightly modified to present the tables according to the requirements of *JMR*.

Table B1
DESCRIPTION OF POSSIBLE MEALS

Attributes	*Meal 1*	*Meal 2*	*Meal 3*	*Meal 4*	*Meal 5*	*Meal 6*	*Meal 7*	*Meal 8*	*Meal 9*	*Meal 10*
Soup	None	Hot and sour soup	Egg drop soup	None	Hot and sour soup	None	None	Egg drop soup	Hot and sour soup	Egg drop soup
Rice/noodles	Brown rice	Brown rice	Noodles	Brown rice	Brown rice	White rice	Brown rice	White rice	Noodles	White rice
Type of sauce	Brown sauce	Sweet and sour sauce	Brown sauce	Sweet and sour sauce	Brown sauce	Brown sauce	Szechwan sauce	Sweet and sour sauce	Sweet and sour sauce	Szechwan sauce
Type of vegetables	Tofu and exotic vegetables	No vegetables	Standard vegetables	Tofu and exotic vegetables	Standard vegetables	Tofu and exotic vegetables	Standard vegetables	Standard vegetables	Tofu and exotic vegetables	Standard vegetables
Type of meat	No meat	Shrimp	Chicken	Shrimp	Shrimp	Beef	Beef	Shrimp	Shrimp	No meat
Spring roll	Pork spring roll	Vegetable spring roll	Vegetable spring roll	Vegetable spring roll	Pork spring roll	Vegetable spring roll	Vegetable spring roll	Vegetable spring roll	Pork spring roll	Vegetable spring roll
Quantity	Pint	Quart	Quart	Pint	Pint	Quart	Quart	Quart	Pint	Quart
Price	$4.99	$5.99	$4.99	$5.99	$4.99	$3.99	$4.99	$5.99	$3.99	$4.99

Attributes	*Meal 11*	*Meal 12*	*Meal 13*	*Meal 14*	*Meal 15*	*Meal 16*	*Meal 17*	*Meal 18*	*Meal 19*	*Meal 20*
Soup	Hot and sour soup	Hot and sour soup	None	Egg drop soup	Hot and sour soup	None	Egg drop soup	Hot and sour soup	Hot and sour soup	None
Rice/noodles	White rice	White rice	Noodles	Noodles	Brown rice	White rice	Noodles	Noodles	Noodles	Noodles
Type of sauce	Brown sauce	Szechwan sauce	Szechwan sauce	Brown sauce	Sweet and sour sauce	Brown sauce	Sweet and sour sauce	Brown sauce	Szechwan sauce	Sweet and sour sauce
Type of vegetables	Tofu and exotic vegetables	Tofu and exotic vegetables	No vegetables	Standard vegetables	Tofu and exotic vegetables	Standard vegetables	Standard vegetables	No vegetables	Tofu and exotic vegetables	Tofu and exotic vegetables
Type of meat	Beef	Shrimp	Beef	Beef	Shrimp	No meat	Chicken	Beef	Chicken	No meat
Spring roll	Vegetable spring roll	Pork spring roll	Pork spring roll	Pork spring roll	Vegetable spring roll	Vegetable spring roll	Vegetable spring roll	Pork spring roll	Pork spring roll	Vegetable spring roll
Quantity	Pint	Pint	Pint	Quart	Quart	Pint	Pint	Quart	Quart	Pint
Price	$3.99	$4.99	$3.99	$5.99	$5.99	$3.99	$3.99	$4.99	$5.99	$4.99

We would ask that you pay close attention to the different snacks being offered and determine an accurate value for each combo. You now have an opportunity to purchase a snack combo. Here is how it works. You will be shown 27 sets of three combos. For each set of three combos, please choose between no snack and one of these three different combos at the stated price. After you complete your selection, we will randomly choose a set from these 27 sets, and your choice for that set will be fulfilled. If you have selected no snack for that set, you will be given $3 cash; if you have selected a snack combo for that set, you will be given $3 minus the price of that combo as stated, in addition to the actual snack combo.

Remember, the choice you make here in the experiment will be fulfilled (you will receive the actual snack combo selected by you) [see Table C1 for an example of a conjoint task].

APPENDIX D: HOLDOUT TASK FOR STUDY 2

You will receive $3, and you can use it to purchase the snack. Please select the combo [see Table D1] that you will be interested to buy (just one) or, in the case you are not interested in any of them, indicate as such.

I want to buy Combo # _____.

I do not want to buy any combo _____.

Table C1

AN EXAMPLE OF THE CONJOINT TASK

1	$2.50	Water	Peanut butter	Korean strawberry cereal bar	Banana
2	$1.75	Orange juice	Peanut butter	No Korean cereal bar	Apple
3	$2.50	Diet Coke	Chocolate fudge	Korean white chocolate cereal bar	Banana

Please indicate your most preferred choice:

_____ Combo 1

_____ Combo 2

_____ Combo 3

_____ Don't want to purchase any combo from this page.

Table D1

30 AVAILABLE SNACK COMBOS

Item #	*Price ($)*	*Drink Included*	*Cookie Included*	*Korean Cereal Bar Included*	*Fruit Included*
1	1.00	Coke	No cookie	No Korean cereal bar	Banana
2	1.00	Ice tea	Oatmeal raisin	Korean white chocolate cereal bar	Banana
3	1.75	Coke	Chocolate fudge	No Korean cereal bar	Apple
4	2.50	Coke	Oatmeal raisin	No Korean cereal bar	No fruit
5	1.75	Ice tea	Chocolate fudge	Korean white chocolate cereal bar	Banana
6	1.75	Coke	Chocolate fudge	Korean dark chocolate cereal bar	Apple
7	1.00	Water	Chocolate fudge	Korean dark chocolate cereal bar	Apple
8	1.00	Coke	Peanut butter	Korean dark chocolate cereal bar	No fruit
9	2.50	Orange juice	Chocolate fudge	Korean dark chocolate cereal bar	Apple
10	2.50	Water	No cookie	Korean strawberry cereal bar	Banana
11	2.50	Diet Coke	No cookie	Korean white chocolate cereal bar	Apple
12	1.00	Diet Coke	Peanut butter	Korean white chocolate cereal bar	No fruit
13	2.50	Orange juice	Peanut butter	No Korean cereal bar	Apple
14	1.75	Water	Oatmeal raisin	No Korean cereal bar	No fruit
15	2.50	Orange juice	Peanut butter	Korean dark chocolate cereal bar	Apple
16	2.50	Diet Coke	No cookie	Korean dark chocolate cereal bar	No fruit
17	2.50	Orange juice	Chocolate fudge	No Korean cereal bar	Banana
18	1.00	Coke	Chocolate fudge	Korean strawberry cereal bar	No fruit
19	2.50	Orange juice	Oatmeal raisin	Korean white chocolate cereal bar	No fruit
20	1.75	Ice tea	No cookie	Korean dark chocolate cereal bar	No fruit
21	1.75	Water	Peanut butter	Korean white chocolate cereal bar	Banana
22	2.50	Diet Coke	Peanut butter	Korean dark chocolate cereal bar	No fruit
23	1.00	Ice tea	Chocolate fudge	Korean strawberry cereal bar	No fruit
24	2.50	Water	Chocolate fudge	No Korean cereal bar	No fruit
25	1.75	Ice tea	Peanut butter	Korean dark chocolate cereal bar	Banana
26	2.50	Ice tea	Chocolate fudge	Korean white chocolate cereal bar	Banana
27	2.50	Coke	No cookie	Korean strawberry cereal bar	Banana
28	2.50	Coke	Oatmeal raisin	Korean strawberry cereal bar	No fruit
29	1.75	Ice tea	Chocolate fudge	Korean dark chocolate cereal bar	Banana
30	1.75	Orange juice	No cookie	Korean strawberry cereal bar	No fruit

REFERENCES

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

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak (1964), "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 9 (July), 226–32.

Bishop, R. and T.A. Heberlein (1986), "Does Contingent Valuation Work?" in *Valuing Environmental Goods: A State of the Arts Assessment of Contingent Valuation Method*, R. Cummings, D. Brookshire, and W. Schulze, eds. Totowa, NJ: Rowman & Allenheld.

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), 385–91.

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

Dahan, E., J.R. Hauser, D. Simester, and O. Toubia (2002), "Application and Test of Web-Based Adaptive Polyhedral Conjoint Analysis," Working Paper No. 146, Center for eBusiness@MIT.

Davidson, J.D. (1973), "Forecasting Traffic on STOL," *Operations Research Quarterly*, 24 (4), 561–69.

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.

Elrod, Terry, Jordan J. Louviere, and Krishnakumar S. Davey (1992), "An Empirical Comparison of Rating-Based and Choice-Based Conjoint Models," *Journal of Marketing Research*, 29 (August), 368–77.

Green, Paul E., Kristiaan Helsen, and Bruce Shandler (1988), "Conjoint Validity Under Alternative Profile Presentation," *Journal of Consumer Research*, 15 (December), 392–97.

——— 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), 117–32.

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

——— 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.

Johnson, Eric, Robert J. Meyer, and Sanjay Ghose (1989), "When Choice Models Fail: Compensatory Models in Negatively Correlated Environments," *Journal of Marketing Research*, 26 (August), 255–70.

Kohli, Rajiv and Vijay Mahajan (1991), "A Reservation-Price Model for Optimal Pricing of Multiattribute Products in Conjoint Analysis," *Journal of Marketing Research*, 28 (August), 347–54.

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

Leigh, Thomas W., David B. MacKay, and John O. Summers (1984), "Reliability and Validity in Conjoint Analysis and Self-Explicated Weights: A Comparison," *Journal of Marketing Research*, 21 (November), 456–62.

Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green, and Martin R. Young (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, 15 (2), 173–91.

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.

——— and J.F. Shogren (1998), "The Deadweight Loss of Christmas: Comment," *American Economic Review*, 88 (5), 1350–55.

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), 334–42.

Newton, Michael A. and Adrian E. Raftery (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 56, 3–48.

Page, Albert L. and Harold F. Rosenbaum (1987), "Redesigning Product Lines with Conjoint Analysis: How Sunbeam Does It," *Journal of Product Innovation Management*, 4 (2), 120–37.

Robinson, Patrick J. (1980), "Application of Conjoint Analysis to Pricing Problems," in *Proceedings of the First ORSA/TIMS Special Interest Conference on Marketing Measurement and Analysis*, D.B. Montgomery and D.R. Wittink, eds. Cambridge, MA: Marketing Science Institute, 183–205.

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.

Srinivasan, V. (1988), "A Conjunctive-Compensatory Approach to the Self-Explication of Multiattributed Preferences," *Decision Sciences*, 19 (2), 295–305.

———, Peter G. Flaschsbart, Jarir S. Dajani, and Rolfe G. Hartley (1981), "Forecasting the Effectiveness of Work-Trip Gasoline Conservation Policies Through Conjoint Analysis," *Journal of Marketing*, 45 (Summer), 157–72.

——— and C.S. Park (1997), "Surprising Robustness of Self-Explicated Approach to Customer Preference Structure Measurement," *Journal of Marketing Research*, 34 (May), 286–91.

Starmer, C. and R. Sugden (1991), "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation," *American Economic Review*, 81 (4), 971–78.

Urban, Glen L., John R. Hauser, William J. Qualls, Bruce D. Weinberg, Jonathan D. Bohlmann, and Roberta A. Chicos (1997), "Validation and Lessons from the Field: Applications of Information Acceleration," *Journal of Marketing Research*, 34 (February), 143–53.

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

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

——— and David B. Montgomery (1979), "Predictive Validity of Trade-Off Analysis for Alternative Segmentation Schemes," in *Educators' Conference Proceedings*, Series 44, Neil Beckwith, Michael Houston, Robert Mittelstaedt, Kent B. Monroe, and Scott Ward, eds. Chicago: American Marketing Association, 69–73.

———, 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.

Wright, Peter and Mary Ann Kriewall (1980), "State-of-Mind Effects on the Accuracy with Which Utility Function Predicts Marketplace Choice," *Journal of Marketing Research*, 17 (August), 277–93.