MARKETING SCIENCE INSTITUTE

Measuring Consumer Willingness to Pay at the Point of Purchase

Klaus Wertenbroch and Bernd Skiera

WORKING PAPER • REPORT NO. 01-105 • 2001

WORKING PAPER SERIES

M A R K E T I N G   S C I E N C E   I N S T I T U T E

Measuring Consumer Willingness to Pay at the Point of Purchase

Klaus Wertenbroch and Bernd Skiera

WORKING PAPER • REPORT NO. 01-105 • 2001

W O R K I N G   P A P E R   S E R I E S

We are grateful to Sönke Albers, Karen Gedenk, Wagner Kamakura, Jordan Louviere, Subrata Sen, Russ Winer, Dick Wittink, and three anonymous reviewers for helpful comments on earlier versions and presentations of this paper, and we thank Birga Schweiger and Hristo Giochev for their help with the data collection.

Measuring Consumer Willingness to Pay at the Point of Purchase

Klaus Wertenbroch and Bernd Skiera

Market researchers rely on measures of consumers' willingness to pay (WTP) in estimating product demand and in designing optimal price schedules. Existing market research techniques for measuring WTP differ in whether they provide an incentive for consumers to reveal their true WTP and in whether they simulate actual point-of-purchase contexts.

In this study, authors Wertenbroch and Skiera present an empirical comparison of several procedures for eliciting WTP that are applicable directly at the point of purchase. Specifically, they test the applicability of a lottery-based procedure to measuring consumer WTP.

Studies

The researchers sampled target consumers and provided them with an opportunity to buy the product in question. Respondents were asked to announce their true WTP for the product at the particular point of purchase. Next, they randomly drew a price from a prespecified distribution in an urn or envelope. If the drawn price was less than or equal to the WTP they indicated, they were required to buy the product at the price they drew. If the drawn price exceeded their offer, they were not allowed to buy the product. This mechanism ensured that their best response strategy was to announce their true WTP. Rather than receiving a participation fee, respondents actually paid out-of-pocket—but never more than the product was really worth to them, thus ensuring the ethical soundness of the procedure. That is, they never walked away with less value than they had when they began the task.

Findings and Implications

Three studies demonstrate that the lottery procedure provides a feasible, reliable, and valid market research procedure to elicit consumer WTP in specific point-of-purchase settings in fast-moving consumer goods markets. It entails relatively little cost, time, and effort to administer.

A key result across all three studies is that consumers report substantially lower WTP under the lottery procedure than under hypothetical response formats. This suggests that hypothetical methods may lead managers to overprice relative to consumers' true WTP. They also show that the lottery procedure provides a better measure of consumers' true point-of-purchase WTP because consumers tend to round prices when responses are only hypothetical, whereas responses under the lottery procedure are more differentiated.

Hypothetical responses depend on the prices consumers *normally* pay in the category, whereas the incentive constraint under the lottery procedure helps respondents determine their WTP based on the *specific point-of-purchase context*, under the controlled influence of marketing mix variables (e.g., WTP for candy bars on display at supermarket checkout counters, where retailers may try to induce impulse buying). Thus, responses reflect how ready to buy consumers are under real transaction conditions.

A key benefit of the lottery-based method is that it allows market researchers to create opportunities for transactions at real points of purchase under the actual marketing mix conditions that the marketer desires. This is especially useful for new products, for which no actual purchase data yet exist. The method can serve as a stand-alone (off- or online) procedure, or it can be combined with existing preference elicitation techniques and pretest market research to develop better insights into the factors that influence consumer choice and product valuations. It can also provide an alternative to auctions in order to price-discriminate in regular online transactions.

Klaus Wertenbroch is Associate Professor of Marketing, INSEAD, France. Bernd Skiera is Professor of Electronic Commerce, Johann Wolfgang von Goethe-Universität, Germany.

Contents

Figures

Introduction

> *I am inclined to offer Mr. Vieweg from Berlin an epic poem,* Herrmann and Dorothea*, which will have approximately 2,000 hexameters. . . . Concerning the royalty we will proceed as follows: I will hand over to Mr. Counsel Böttiger a sealed note which contains my demand, and I wait for what Mr. Vieweg will suggest to offer for my work. If his offer is lower than my demand, then I take my note back, unopened, and the negotiation is broken. If, however, his offer is higher, then I will not ask for more than what is written in the note to be opened by Mr. Böttiger.*
>
> – Johann Wolfgang von Goethe in a letter on January 16, 1797, cited in Moldovanu and Tietzel (1998)

Goethe's scheme to elicit a price from Vieweg for his manuscript is perhaps the earliest documented example of enticing a buyer in an incentive-compatible format to truthfully reveal his willingness to pay, or reservation price, presaging Nobel Laureate William Vickrey's (1961) analysis of the problem by a century and a half. To Vieweg, Goethe's sealed price represents a random variable that is distributed independently of his own WTP (Moldovanu and Tietzel 1998). Hence, Vieweg's dominant response strategy is to bid exactly his WTP. WTP denotes the maximum price a buyer is willing to pay for a given quantity of a good. It is a ratio-scaled measure of the subjective value the buyer assigns to that quantity. He or she buys the item from a set of alternatives, for which his or her WTP exceeds purchase price the most. As in the above example, knowledge of consumer WTP is crucial in estimating demand and designing optimal pricing schedules. Existing market research elicitation techniques differ according to whether they provide an incentive to consumers to reveal their true WTP and according to whether they simulate actual point-of-purchase contexts.

We present an empirical comparison of several such procedures for eliciting WTP at the point of purchase. In particular, we test the applicability of Becker, DeGroot, and Marschak's (1964) well-known BDM procedure for measuring the utility of lotteries to measuring consumer WTP in market research. The BDM procedure allows us to combine incentive-compatibility with WTP elicitation in relevant point-of-purchase contexts. We begin with a brief discussion of existing approaches to WTP measurement in market research. Next, we develop the theoretical and methodological properties of the BDM procedure for use in market research. We then describe two point-of-purchase field studies that apply the procedure to inexpensive grocery items. The studies demonstrate its feasibility, reliability, and validity and show its superior performance as compared to a conventional method based on survey responses. A follow-up experiment compares BDM with a typical hypothetical choice-based approach. The experiment shows the feasibility of the BDM procedure across domains, applying it to an inexpensive durable. It suggests that BDM yields lower WTP estimates than hypothetical methods because of the incentive constraint rather than the cognitive effort it requires of respondents. The experiment also controls for strategic response behavior. Our closing discussion focuses on the limitations of applying the BDM procedure in market research.

Determining Willingness to Pay in Market Research

Transactions Data

Market researchers estimate WTP either from actual market transactions (revealed preferences, e.g., from scanner data) or from survey data (stated preferences). Transactions data such as scanner and simulated test market data (e.g., Silk and Urban 1978) are incentive-compatible and have high external validity because actual purchases are observed under realistic marketing mix conditions. For example, test market simulations such as ACNielsen's BASES system (www.bases.com) provide consumers with opportunities to buy real products from competitive choice sets at experimentally manipulated price points. Participants receive a "participation fee" that they can either keep or spend on the available products at the posted prices. Thus, demand estimates cannot be biased downward simply due to possible liquidity constraints. However, posted prices in real or simulated markets typically vary only within limited ranges (Ben-Akiva et al. 1994). Hence, transactions data reveal only that a buyer's WTP is at least as high as the posted price and that a nonbuyer's WTP is lower than that price. An individual's true WTP remains unknown, preventing marketers from extracting maximum consumer surplus. Moreover, except for test market simulations, transactions data are unavailable for new products that have not yet been sold under market conditions.

Survey Data

In contrast, the key advantage of survey data is that they can be elicited in concept testing and new product development, or the evaluation of nonmarket public goods (e.g., Cameron and James 1987). The most successful methodology in market research is conjoint analysis (Green and Srinivasan 1990), but survey data can also come from contingent valuation (Mitchell and Carson 1989). Conjoint analysis is designed to determine trade-offs between product features or attributes (including price), and differences in utilities (WTP) are inferred from subjects' rankings or ratings of alternatives, or WTP is elicited as a dependent variable as the sum of money that would make subjects indifferent between a bundle of attributes and the money (e.g., Kalish and Nelson 1991; Rao and Soni 1994). Contingent valuation and related approaches (e.g., Jones 1975; Kalish and Nelson 1991) require respondents to state their WTP for entire goods or for attribute-level changes directly (open-ended contingent valuation) or to make single or repeated choices about whether they would buy a good at a given price (closed-ended contingent valuation). On the downside, the external validity of these approaches may be limited as they provide little incentive to consumers to truthfully reveal their WTP because responses are hypothetical (Hoffman et al. 1993).[1] Response incentive effects can occur in contingent valuation when the survey itself prompts respondents to make inferences about the value of the good (Alberini, Kanninen, and Carson 1997; Carson, Groves, and Machina 1999). Finally, Hoffman et al. (1993) suggested that consumers' revealed preferences (i.e., choice reactions to posted prices) may arise from different reference frames than their stated preferences (i.e., stated reservation prices). Because the demand revealed under these two approaches may differ systematically, a variety of WTP measurement procedures may be needed for cross-validation purposes.

Vickrey Auctions

Market researchers need methods that are both applicable at the point of purchase and that provide incentive-compatible estimates of WTP derived from real transactions. To address the latter goal, Hoffman et al. (1993) advocated the use of experimental Vickrey auctions. Vickrey (1961) suggested that incentive compatibility is ensured if a given bid determines only whether the bidder has the right to buy the good that is auctioned off. In a sealed-bid auction, the actual purchase price is determined solely by the other participants' bids. The *n* highest bidders in a Vickrey auction (also called $(n+1)^{th}$-price, sealed-bid auction) win the good at the price of the $(n+1)^{th}$-highest bid. The dominant strategy is to bid exactly one's WTP (Kagel 1995; Vickrey 1961). Unlike methods based on stated preference data, Vickrey auctions provide bidders with an incentive to truthfully reveal their WTP because they must buy the good in a real transaction if their bid wins the auction.

Despite these theoretical advantages, Vickrey auctions exhibit practical and empirical limitations. First, auctions may pose operational problems, as consumers/bidders have to meet in a research facility, typically at a substantial setup cost. Second, auction bidding processes do not naturally mimic consumer point-of-purchase decisionmaking processes in normal retail settings (Hoffman et al. 1993). In contrast to the practically unrestricted supply of goods in actual retail settings, bidders in an auction compete with each other for a limited stock. In practice, the decision of how much to bid is thus not only a question of the bidder's true valuation of the good but also of ensuring that he or she places the winning bid. Thus, participants in Vickrey auctions of objects with private values tend to violate the incentive constraint by bidding more than the objects are worth (e.g., Kagel 1995; Kagel, Harstad, and Levin 1987). These empirical violations of incentive compatibility may limit the external validity and usefulness of Vickrey auctions in marketing research. Hoffman et al. (1993) acknowledge the challenge of designing "learning trials, task frames and instructions that explain incentive-compatible auctions so as to minimize both the incidence and impact of such strategic behavior" (p. 334).

BDM

To address some of the theoretical, empirical, and practical limitations of conventional transaction- and survey-based methods simultaneously, we apply BDM to the elicitation of WTP at the point of purchase. This application is designed to be theoretically incentive-compatible, realistic, transparent to respondents, and operationally efficient. Unlike Vickrey auctions, it allows researchers to determine individual consumers' WTP in relevant and typical purchase settings in the field. Becker, DeGroot, and Marschak's (1964) original procedure measured the utility of lotteries by eliciting minimum selling prices (willingness to accept, or WTA) for gambles in an incentive-compatible format, by determining actual transaction prices randomly (i.e., by drawing a ball marked with a price from an urn). So the distribution of BDM transaction prices is exogenous to respondents' WTPs, just like in Vickrey auctions. Experimental researchers in behavioral decision theory have widely used BDM-type random preference elicitation procedures with consumer goods as stimuli (e.g., Kahneman, Knetsch, and Thaler 1990; Wertenbroch 1998), but market research practitioners have not relied on this approach.

Incentive-Compatible Elicitation at the Point of Purchase

A key benefit of applying BDM is that it allows market researchers to create opportunities for transactions at real points of purchase under the actual marketing mix conditions that the marketer desires. Target consumers are sampled and told that they have a chance to buy the product without having to spend more money on the purchase than they truly want to. They learn that the buying price p for the product is not yet set and will be determined randomly. Next, we explain, and then apply, the following procedure. We ask consumers to offer a price s for the product, which should equal the highest price they are willing to pay for the product. Next, we randomly determine p from a prespecified distribution (which is unknown to respondents). Letting consumers themselves draw a ticket marked with a price from an urn should increase their confidence in the randomness of the price setting mechanism and underscore the futility of misrepresentation. If the drawn price p is less than or equal to their offer s, they are required to buy the product at price p. If p exceeds their offer, they are not allowed to buy the product. Figure 1 illustrates the procedure; sample instructions can be found in the appendix. The dominant strategy is to offer one's true WTP because, for any distribution of buying prices,

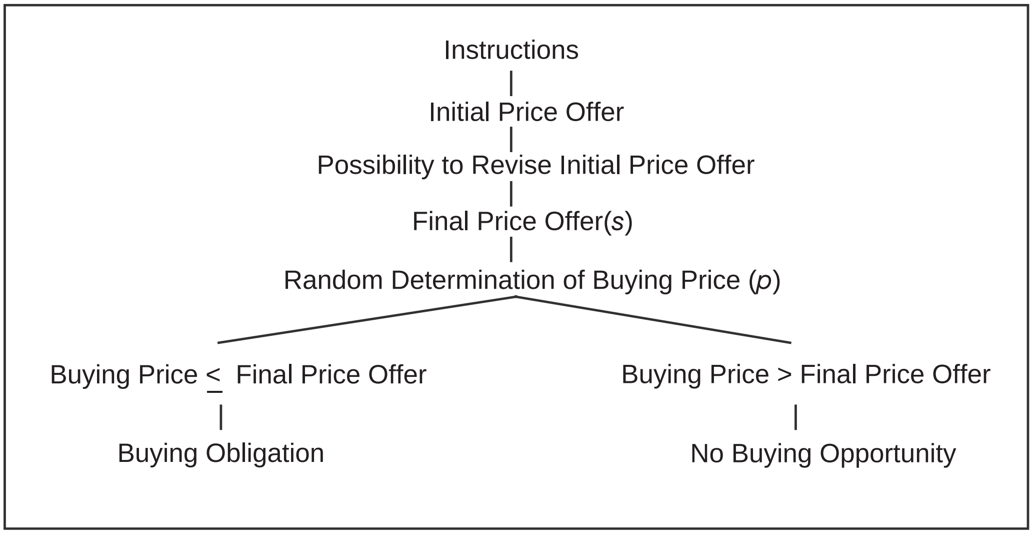
1. *understating* one's true WTP ($s < \text{WTP}$) reduces the chance of buying at a gain (where the forgone gain is $\text{WTP} - p \geq 0$ for all $s < p < \text{WTP}$), without increasing the actual gain if the consumer has to buy (if $p \leq s$) since understating cannot affect the buying price p;

2. *overstating* one's true WTP ($\text{WTP} < s$) increases the chance of buying at a loss (where the incurred loss is $\text{WTP} - p < 0$ for all $\text{WTP} < p \leq s$).

Realistic Purchase Settings

BDM has both strengths and weaknesses compared to existing methods. WTP can be elicited right at the point of purchase so it can vary as a function of the actual purchasing context and competitive set, which should enhance external validity. This also makes BDM easier to administer and should mitigate the overbidding found in Vickrey auctions (e.g., Kagel 1995), where consumers convene in an artificial format that is unrepresentative of the actual purchase context and that may trigger unrepresentative competitive bidding behavior. The realism in the elicitation context that BDM affords is crucial. Thaler (1985) found that (hypothetically) incentive-compatible WTPs differ dramatically for identical items (cold beer) depending on the point of purchase (a fancy resort hotel or a rundown grocery store) as the transaction context itself induces different levels of utility and WTP. Experimental choice researchers have long argued that choice tasks need to be designed realistically, approximating as closely as possible the actual purchase context (Carson, Groves, and Machina 1999), as consumers often construct their preferences in response to the choice context rather than retrieving a previously formed value (Bettman, Luce, and Payne 1998).

Figure 1. Flow Chart of BDM Procedure



"Out-of-Pocket" Transactions

As an additional incentive for respondents to consider their price offers carefully, we suggest that they not be given any compensation for their participation. Thus, they must have enough cash to pay out of their own pockets. This "out-of-pocket" obligation forces respondents to consider their real readiness to buy and minimizes possible distortions caused by the windfall character of any extra compensation. Thaler and Johnson (1990) have shown that propensity to spend varies with whether the funds stem from such windfall gains. Hence, prices that are elicited with existing procedures may overestimate WTP in everyday market transactions because subjects receive a participation fee so that they never leave the lab with less money than when they entered it. Under BDM, however, respondents need not be compensated for coming to a research facility (see Hoffman et al. 1993) because they are intercepted at real points of purchase. Making respondents pay out of pocket renders BDM particularly suitable for determining WTP for unplanned, low cost purchases of consumer packaged goods. For these, any liquidity effects should be small or negligible. The downside of the out-of-pocket feature is that measuring WTP for more expensive items such as durables may bias WTP downward due to uncontrolled liquidity constraints. Thus, respondents would have to be allowed to pay by check or credit card. Moreover, consumers typically buy big-ticket items after much more deliberation so that they may not be as willing to participate in BDM for these items.

The Distribution of Prices

Valid WTP estimation requires that respondents trust the interviewer and expect to participate in a fair transaction. Thus, the distribution, from which buying prices are drawn, has to have a range, within which all prices appear fair. If the distribution is skewed toward high prices, consumers may infer that they are being "cheated," similar to bidders' fears of experimenter misrepresentation in Vickrey auctions (Rothkopf and Harstad 1995). Subjects should not be told about the range and moments of the distribution to avoid anchoring, which affects selling price bids in BDM (Bohm, Lindén, and Sonnegård 1997). The choice of distribution type and moments is flexible. It depends on the researcher's budget and objectives, as the distribution affects the researcher's expected revenue. The distribution itself cannot influence WTP responses. However, it is possible that the lottery nature of the task may bias responses. For example, for existing products, respondents might cap their bids at the remembered market price if they see the task as an opportunity to "win" the item at a price below market. The possibility of such underbidding marks all incentive-compatible procedures; we return to the issue of strategic response biases below.

In sum, BDM permits the elicitation of incentive-compatible WTP without convening consumers in groups in a laboratory. Respondents do not compete with others for the same product. They pay out of pocket in real purchase locations such as malls or stores, subject to the intended purchase conditions and realistic purchasing motives. This keeps the costs of BDM low compared to laboratory-based market research unless BDM buying prices happen to be much lower than the cost of goods sold. We now turn to an empirical assessment of BDM vis-à-vis other elicitation procedures.

Studies 1 and 2: Testing BDM in the Field

Two independent studies tested the feasibility, reliability, and validity of applying BDM at the point of purchase. Study 1 elicited WTP for a can of Coca-Cola on a public beach in Kiel, Germany. Study 2 elicited WTP for a piece of pound cake on a commuter ferry in Kiel. No substitutes were available in either location. Our WTP distributions are specific to these monopolistic contexts. This point-of-purchase specificity is a key feature of WTP, as WTP distributions cannot be generalized across choices from different competitive sets. But for any given point of purchase (including monopoly settings), we can compare the performance of BDM with that of other methods for measuring WTP.

As a benchmark for methods based on stated preferences, we asked consumers to state their WTP hypothetically (e.g., Gabor and Granger 1966; Jones 1975; Kalish and Nelson 1991). According to Carson, Groves, and Machina (1999), this continuous response format is a special type of open-ended contingent valuation that corresponds to price matching in the decisionmaking literature (e.g., Tversky, Slovic, and Kahneman 1990).[2] Except for the lack of incentive compatibility, we elicited these matching prices under the same conditions as WTP under BDM, that is, in the actual purchase contexts.

Method

Subjects, Design, and Procedure. Four-hundred randomly selected consumers participated in the studies, 200 beach visitors in the first and 200 ferry passengers in the second. Within each study, 100 consumers' WTP was elicited via price matching (control group) and the other 100 consumers' WTP via BDM (test group) in a two-level between-subjects design, in which subjects were randomly assigned to conditions. To control for environmental conditions that might affect demand for the products, we ran each study under equal weather conditions across four consecutive days. The same interviewer conducted all interviews, with an equal number of consumers in both groups being interviewed every day during the same time interval to further assure comparability of conditions.

The interviewer approached subjects individually and introduced herself as an academic marketing researcher from the local university. Subjects in the control groups (price matching) were shown a can of classic Coca-Cola or a piece of pound cake and were asked for the maximum price they would be willing to pay for it if it were for sale. In the test groups, we applied the BDM procedure outlined above (Figure 1). Thus, each test subject was actually offered a can of Coke or a piece of cake for purchase. The specific instructions read to subjects by the interviewer are listed in the appendix, including instructions on revising their offers. When respondents had determined their final price offer, they drew a ball from the urn to determine the purchase price.

In both studies, the distribution of potential purchase prices in the urn was uniform. Coke prices ranged from DM 0.50 to DM 2.50, in increments of DM 0.10 (DM 1.00 approximately equaled U.S. $0.55). Cake prices ranged from DM 0.80 to DM 2.00, in increments of DM 0.10. These ranges were wide enough to include the soft drink and pound cake prices that we had found in a survey of different local retail outlets. None of the characteristics of the price distributions in the urn were reported to subjects, even if they asked, to avoid anchoring their responses (cf. Bohm, Lindén, and Sonnegård1997).

Measures. We recorded the final price offer (s), the randomly determined purchase price (p), and whether subjects complied with their purchase obligations. We also examined various transparency and acceptability measures of BDM (Table 1), and subjects rated how thirsty (hungry) they were, how much they liked the items, how much they were craving them, and what price they normally paid for them (Table 2). Finally, BDM subjects who had to buy rated on 5-point Likert scales how much value and satisfaction they derived from their purchase, while BDM subjects whose price offers were less than the purchase price stated whether they regretted not having offered a higher price and whether they would have bought at a higher price.

Results

Table 1 shows that respondents perceived BDM as highly transparent and acceptable. The procedure was not confusing, nor was it difficult to understand why it was optimal to state exactly one's WTP. Hardly anyone approached refused to participate in the studies. Those who did decline to participate did not care at all for Coke or pound cake. All who agreed to participate did so without hesitation and visibly enjoyed drawing the purchase prices themselves to see whether they were eligible to buy a Coke or a piece of cake. Some respondents asked about the range of the possible purchase prices but were content when told that the distribution was "reasonable, with prices neither too high nor too low."

Table 1. Mean WTP and Transparency and Acceptability Ratings per Condition in Studies 1 and 2, *N* = 100 per condition (standard deviations in parentheses, minimum WTP = 0 in all conditions)

	Coke		Cake	
	BDM	**Price matching**	**BDM**	**Price matching**
Mean WTP (across four subsamples in each of the four conditions)§	DM 1.06 [a] (.66)	DM 1.35 [a] (.81)	DM 1.12 [b] (.56)	DM 1.68 [b] (.82)
Has this procedure been confusing for you? (reverse scored: 1 = very much so, 5 = not at all)§§	4.81 (.51)	n.a.	4.95 (.22)	n.a.
Is it clear why it is in your best interest to state exactly the price you are willing to pay?§§	4.21 (.78)	n.a.	4.14 (.69)	n.a.
Would you participate in a survey like this again? (% yes responses)	99.0%	n.a.	95.0%	n.a.

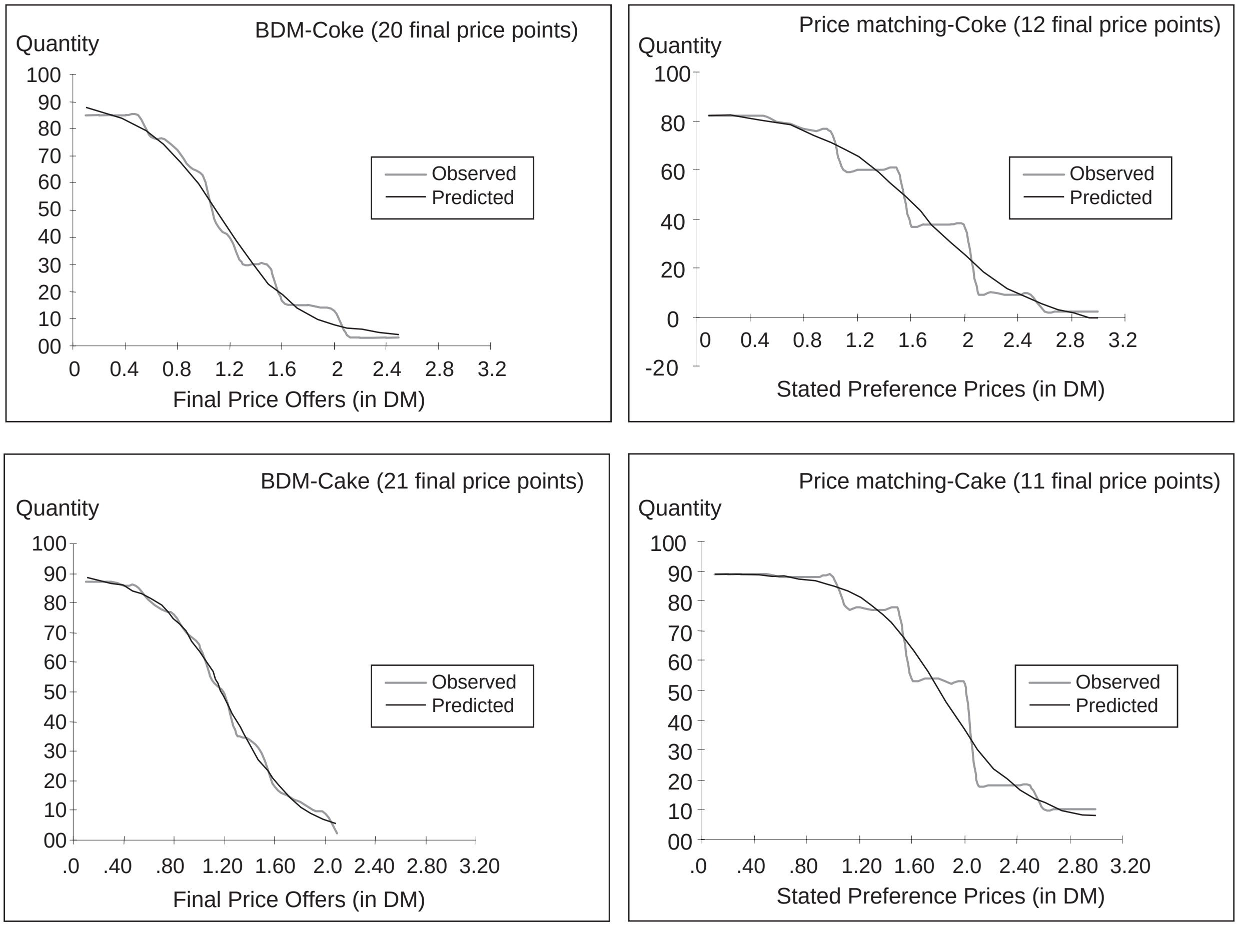
§ WTP is either final price offer (BDM) or stated price (price matching)
§§ responses on 5-point-scale (1 = not at all, 5 = very much so)
a values with same superscripts differ (at $p < .01$ in t-test with unequal variances)
b values with same superscripts differ (at $p < .0001$ in t-test with unequal variances)
n.a. not applicable

Table 2. Correlation of WTP with Measures of Face Validity in Studies 1 and 2

	Coke		Cake	
	BDM	Price matching	BDM	Price matching
How thirsty/hungry are you right now?	.2708** (N = 100)	.0809 (N = 100)	.2749** (N = 100)	.0079 (N = 100)
How much do you like Coca-Cola/ cake?	.3641**** (N = 100)	.3005** (N = 100)	.5915**** (N = 100)	.2270* (N = 100)
How much do you normally pay for a can of Coca-Cola/a piece of cake? (in DM)	.2566* (N = 87)	.3111** (N = 75)	.0337 (N = 92)	.4543**** (N = 94)
How much did you crave the Coca-Cola/the piece of cake?	.4679**** (N = 100)	n.a.	.6478**** (N = 100)	n.a.

Responses on 5-point-scale for thirst/hunger, liking, and craving (1 = not much, 5 = a lot)
n.a.: not applicable
*p < .05, ** p < .01, **** p < .0001 that r = 0.

Figure 2. Observed and Predicted Demand in Studies 1 and 2



Distribution of WTP. The cumulative distributions of WTPs in Figure 2 show that consumers stick to major price points when asked only hypothetically under price matching (observed demand is measured as the number of respondents who say they would buy at price $p \leq$ WTP). In contrast, WTPs elicited under BDM are distributed more smoothly and are more differentiated, suggesting greater accuracy. Table 1 shows that mean WTPs under BDM are lower than mean matching prices (Δ = DM -0.29, t = -2.65, p < .01 for Coke; Δ = DM -0.56, t = -5.70, p < .0001 for cake). For Coke, the mean final price offer is more than 20 percent below the mean matching price; for cake it is 33 percent less.

Reliability. We determined the reliability of the WTP measures by comparing mean WTPs across the four daily respondent subsamples within each condition (cf. Green, Tull, and Albaum 1988, p. 253). ANOVAs of nonzero WTPs failed to reveal any differences between the subsamples under BDM for Coke (F(3,82) = 0.80, p < 1) and for cake (F(3,96) = 1.14, p < 1). This suggests that BDM is reliable when point-of-purchase characteristics remain stable. In contrast, our data suggest that WTPs elicited under price matching may be less reliable. They appeared comparable across subsamples for Coke ($F(3,96) = 0.13$, $p < 1$) but were not equal for cakes ($F(3,96) = 2.63$, $p < .10$).

Face Validity of Both Methods. We determined the face validity of WTPs by correlating the final BDM-price offers and the matching prices with the respondent characteristics shown in Table 2. All correlation coefficients have the expected (positive) signs. We transformed the individual coefficients into Fisher-*Z* scores (see, e.g., Glass and Stanley 1970) in order to test whether the correlations differed across test and control groups. Subjects' WTPs for Coca-Cola correlated significantly with how thirsty they were under BDM ($Z_{BDM} = .27$) but not under price matching ($Z_{match} = .08$; $z = 1.37$, $p < .10$ for the difference in the two *Z*-coefficients). Subjects' liking for Coca-Cola showed a directionally similar pattern, but the difference in correlation coefficients failed to reach significance. Subjects' WTPs for cake correlated with how hungry they felt under BDM ($Z_{BDM} = .28$) but not under price matching ($Z_{match} = .01$; $z = 1.91$, $p < .05$). Similarly, the correlation between WTPs and how much they liked cake was stronger under BDM ($Z_{BDM} = .68$) than under price matching ($Z_{match} = .23$; $z = 3.13$, $p < .01$), and BDM responses in both studies correlated with subjects' craving for the items. As predicted, these results suggest that WTPs from BDM provide a more valid measure of subjects' preferences than those from price matching. How much subjects normally paid for Coca-Cola and cake tended to be less strongly correlated with WTP under BDM than under price matching (Coke: $Z_{BDM} = .26$ versus $Z_{match} = .32$, $z = -.37$, $p < 1$; cake: $Z_{BDM} = .03$ versus $Z_{match} = .49$, $z = -3.06$, $p < .01$). This may imply that price-matching subjects anchor their responses on a reference price instead of carefully determining their situation- and context-specific true WTP. Recall that matching prices were distributed less smoothly than BDM WTPs, also suggesting a simple anchoring or rounding process to determine responses (Figure 2). In contrast, responses under BDM seem to reflect more closely subjects' situation-specific individual demand.

Internal Validity of Both Methods. Logit analyses of purchase probabilities $\Pr(buy|p) = e^{a+b \cdot p}/(1+ e^{a+b \cdot p})$ indicated downward sloping demand with $a = 3.94$ ($p < .0001$) and $b = -3.44$ ($p < .0001$) in BDM-Coke, $a = 4.32$ ($p < .0001$) and $b = -2.52$ ($p < .0001$) in price matching-Coke, $a = 4.58$ ($p < .0001$) and $b = -3.72$ ($p < .0001$) in BDM-cake, and $a = 6.42$ ($p < .0001$) and $b = -3.48$ ($p < .0001$) in price matching-cake. Figure 2 shows that predicted demand (measured as the expected number of respondents who buy at price $p \leq$ WTP) tracks observed demand much better under BDM than under price matching. This was confirmed by an analysis of the Fisher-*Z* transformed correlations between observed and expected demand, which showed better fit under BDM than under price matching (Coke: $r_{BDM} = .9952$, $r_{match} = .9881$, $Z_{BDM} = 3.02 > Z_{match} = 2.56$, $z = 1.60$, $p < .10$; cake: $r_{BDM} = .9981$, $r_{match} = .9876$, $Z_{BDM} = 3.48 > Z_{match} = 2.54$; $z = 3.09$, $p < .01$). These results suggest that BDM provides a highly internally valid measure of WTP and outperforms price matching in modeling demand.

Criterion Validity of BDM. BDM also has high criterion validity, measured as the percentage of consumers who followed through with their purchase obligation (if $p \leq$ WTP). Only 2.5 percent (1 out of 41) refused to comply with their purchase obligation in the Coke study and 7.5 percent (3 out of 40) refused to buy in the cake study. Buyers' ratings of how satisfied they were with their purchases were high (M = 4.17, s = .83 for Coke, M = 4.03, s = .73 for cake on a 5-point scale), suggesting that most of them were happy with their purchase. The large majority of nonbuyers (if $p >$ WTP) did not regret that they had not made a higher price offer and reaffirmed their preference for not buying. This argues against the possibility of significant strategic underbidding. Only 3.4 percent (2 out of 59) in the Coke study and 6.7 percent (4 out of 60) in the cake study said they should have offered a higher price. Of these six, five subjects said that they would have actually liked to buy at the purchase price they had randomly drawn. Finally, we asked for ratings of how much value the purchase provided to those who bought (if $p \leq$ WTP) or would have provided at the drawn price to those who did not buy (if $p >$ WTP). The overall means were M = 2.91 (s = 1.22) for Coke and M = 3.23 (s = 1.07) for cake (on a 5-point scale). These value ratings are significantly correlated with a measure of consumer surplus (WTP – p), r = .66 (p < .0001) for Coke and r = .59 (p < .0001) for cake. Overall, the high purchase obligation compliance rates and satisfaction ratings, the low incidence of regret by nonbuyers, and the strong correlation between consumer surplus and transaction evaluations attest to the high criterion validity of BDM.

Discussion of Studies 1 and 2

Studies 1 and 2 suggest that BDM is a reliable and valid method to determine consumer WTP. It outperforms price matching, a conventional open-ended contingent valuation approach, on measures of reliability and face, internal, and criterion validity. However, several questions remain.[3]

Incentive Compatibility and Strategic Behavior. First, is BDM truly incentive-compatible (i.e., the dominant response strategy is to bid one's true WTP)? BDM is based on Vickrey's (1961) principle of making the distribution of purchase prices exogenous to respondents' valuations, and it mirrors BDM's (1964) procedure for eliciting WTA. Yet, as noted above, respondents in Vickrey auctions often bid above their WTP. Is it similarly possible that BDM respondents misrepresent their WTP? There are at least three possible reasons for this conjecture.[4]

First, BDM may not be truth-revealing if subjects see their responses as consequential beyond the immediate survey context (Carson, Groves, and Machina 1999). Specifically, if they believe that their responses will be used to set long-run market prices, they will have an incentive to *understate* their WTP. If they believe that their responses will determine the introduction of a desirable new product, they may see reasons to *overstate* their WTP. But such belief-based strategic misrepresentation may occur under all response formats (Carson et al. 1994), whether or not they are theoretically incentive-compatible (including Vickrey auctions and BDM). So these arguments hold for all marketing research methods and do not predict a difference in WTP between BDM and price matching, in contrast to our findings in Studies 1 and 2.

Second, if respondents are uncertain about their valuations or want to maintain control over the decision whether to buy, rather than having the transaction imposed on them, they may *overstate* their WTP to maximize their chances of becoming eligible to buy at the next stage. After all, if they then feel that the price drawn from the urn is too high, they may simply walk away from their purchase obligation. Note that this argument, too, applies to Vickrey auctions (Hoffman et al. 1993). But only a small fraction of our respondents in the two studies (4 out of 91) reneged on their purchase obligations. More importantly, Casey and Delquié (1995) found such strategic behavior in choice-based elicitation of WTP only when costs were explicitly framed as losses of money but not when they were framed as payments in a transaction (as under BDM), arguing against this notion of strategic rule violations.

A third possible reason for strategically *overstating* true WTP is that subjects may want to ensure that they do not walk away empty-handed once they are engaged in the elicitation process, a form of escalation of commitment. Similarly, Kagel, Harstad, and Levin (1987) suggested that respondents in Vickrey auctions strategically overbid in order to increase their chances of winning. The expected cost of such overstating is negligible (only pennies more) as respondents may well pay less than their stated prices (Kagel 1995). Study 3 is a laboratory experiment designed to test this argument.

Differential Attention. A second open question is whether BDM outperforms price matching and produces lower WTPs because it imposes an incentive constraint on respondents or simply because it induces more careful consideration of the value they place on an object. The preference reversal literature (e.g., Carmon and Simonson 1998) suggests that price matching induces greater price sensitivity than do choice tasks because directly asking respondents for their WTP makes price as an attribute more salient than in a usual market setting. Might BDM exacerbate the price sensitivity found in price-matching tasks? To determine the role of the incentive constraint versus the salience of price, Study 3 compares BDM with a choice-based elicitation procedure that is also not incentive-compatible but that forces subjects to carefully consider their responses and thus makes price equally salient.

Domain of Applicability. A third open question concerns the domain of applicability of BDM. The method is designed for products that can be made available at the site of the survey, at the point of purchase. Thus, studies 1 and 2 show that BDM performs well for inexpensive food products characterized by instantaneous demand. As a measure of face validity, note that "craving" is a better predictor of WTP for the two food and beverage items than "liking," consistent with the idea that demand for such items is subject to "visceral" factors whose intensity fluctuates with subjects' hunger and thirst (Loewenstein 1996). This result confirms that BDM is well suited for point-of-purchase contexts, in which consumers are interested in making an immediate purchase of a single item. However, applying it to big ticket durables would require greater liquidity from respondents (e.g., carrying checks or credit cards) as well as a more complex survey context, in which the researcher can enforce purchase obligations as well as deliver products immediately. Short of these more complex conditions, can we apply BDM to nonfood items such as inexpensive durables, for which demand is affected by current inventory? Study 3 examines this question as well.

Study 3: Strategic Behavior, Incentive Constraint, Applicable Domain

A laboratory experiment addresses these three issues. First, although there is no gold standard for measuring WTP because it is impossible to know subjects' unobservable true WTP, we can test for strategic overbidding due to an escalation of commitment by varying the conditions for such an escalation of commitment. Thus, we manipulate whether or not subjects receive compensation for their participation in BDM independently of their bids, enabling those who receive compensation not to walk away empty-handed even if they do not transact. We also examine if BDM responses reflect a desire to maintain control over the purchase decision as another possible reason for strategic behavior.

Second, we compare BDM with a repeated choice-based procedure that imposes comparable demands on respondents' efforts and attention. This allows us to isolate the effect of the incentive constraint (i.e., of the immediate behavioral consequences) on WTP from possible effects (1) of insufficient cognitive resources being devoted to the task and (2) of lower salience of price in the price-matching tasks in studies 1 and 2. Specifically, we use a variant of Gabor and Granger's (1966) price list procedure, in which subjects state whether or not they would want to buy an item at each of several price points. Many variants of this technique have appeared in the literature. In closed-ended or binary discrete choice contingent valuation, each subject states for only one price whether they would buy at that price (e.g., Cameron and James 1987). A more efficient variant, double-bounded discrete choice, is also often used in contingent valuation surveys. Subjects state two purchase decisions. Conditional on their response to an initial price, they are given a follow-up price that is either higher or lower than the initial one. The underlying WTP distribution as a function of a set of independent variables (product attributes, respondent characteristics, etc.) is estimated from the choice probabilities with maximum-likelihood techniques (Alberini, Kanninen, and Carson 1997; Hanemann, Loomis, and Kanninen 1991). We use this double-bounded approach and narrow down the range even further, within which a subject's WTP lies, following a choice bracketing technique reported by Casey and Delquié (1995).

The third purpose of Study 3 is to examine if BDM can also be applied to inexpensive durable goods. We elicit WTP for a ballpoint pen in a laboratory context where subjects cannot be assumed to be in the market for a pen. This departure from a typical point-of-purchase use of BDM allows us to test how sensitive the method is to variations in subjects' readiness to buy.

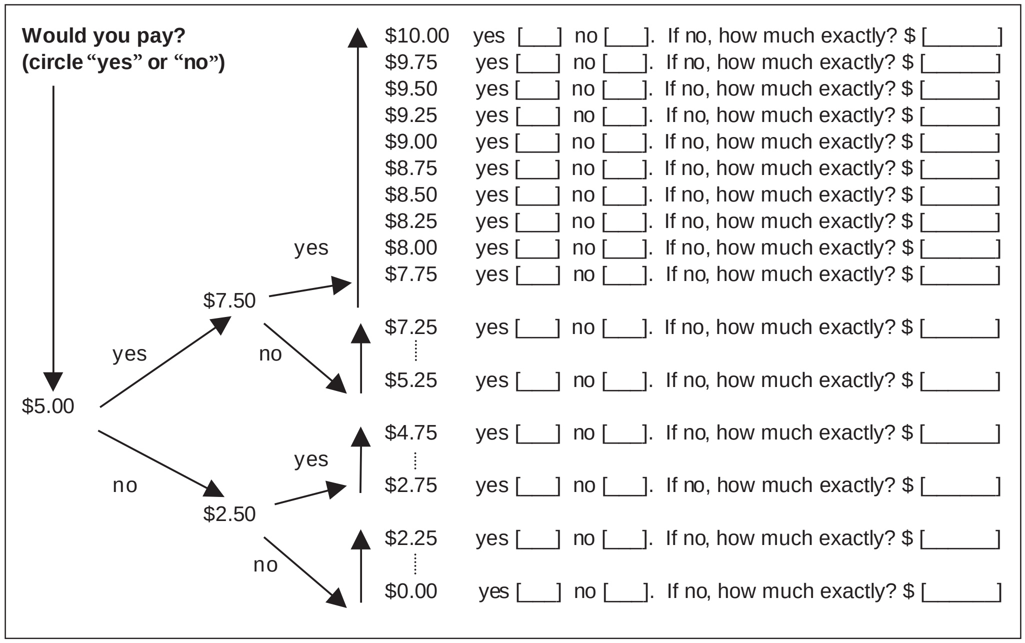
Method

Subjects and Stimuli. Two hundred and fifty-five undergraduate students at a private northeastern university were recruited to participate in this experiment. When they arrived at the experimental room, they were told that the experimenter was interested in determining their WTP at that moment (in the lab) for a newly designed ballpoint pen with an ergonomic cushion grip. The experimenter then showed them this focal pen, along with two other pens for comparison purposes, and let them examine and try out each of the pens. The focal pen cost $5.49 at retail, while the other two pens cost $9.99 and $0.20. We told subjects only the latter two prices.

Design and Procedure. After this introduction, subjects were randomly allocated to one of three conditions. Subjects in the first two conditions were offered an opportunity to purchase the focal pen from the experimenter in the laboratory under BDM. The distribution of prices in the urn was uniform, ranging from $0.00 to $10.00 in increments of $0.25, in line with retail prices of the reference pens. We did not reveal this distribution to the subjects. The specific instructions and procedure followed those for studies 1 and 2 (see appendix). In one of these conditions (BDM-M&Ms), we gave subjects a bag of chocolate candy before they started the procedure. In the other condition (BDM-no M&Ms), they did not get any compensation. This manipulation was meant to test whether the possibility of walking away from the procedure empty-handed affected WTP.

Subjects in the third condition (BRACKETS) were given a bag of candy as compensation for their participation but were not offered an opportunity to purchase the pen. Instead, they were asked to make a series of hypothetical buy/don't buy choices at different price points, imagining that the experimenter was selling the pen in the laboratory. Figure 3 describes the procedure. Specifically, subjects were asked whether they would buy the pen for $5.00. If the response was "no" ("yes"), a follow-up price of $2.50 ($7.50) would be presented. Contingent on a subject's response to that price, one of four lists of nine or ten additional prices was then presented in steps of $0.25. This narrowed down the price range to a small enough interval so that the experimenter asked subjects directly how much exactly they were willing to pay. The presentation order of these incremental prices (increasing or decreasing) was counterbalanced across subjects and had no effect on subjects' responses so that we pooled the data. BDM and BRACKETS both took 8 to 10 minutes to administer.

Figure 3. Choice Bracketing Procedure (BRACKETS condition) in Study 3



Manipulation Checks. Immediately after describing the procedure, we presented subjects with checks on whether giving them M&Ms reduced the conditions for an escalation of commitment that might cause strategic responses. We asked them to rate on 9-point scales whether they felt they were getting something tangible from participating, and whether they felt they would walk away from the experiment empty-handed if they did not transact under BDM (see Table 3 for the actual questions).

Dependent Measures. Dependent measures were subjects' WTP and their subjective estimates of how much they normally paid for ballpoint pens. Subjects also rated how important it was for them to maintain control over whether or not they could buy, and they stated which of the three pens they would most prefer to buy at the given prices, that is, their stated WTP for the focal pen and the listed prices for the two reference pens. Finally, we also asked them to rate on 9-point scales how much of a need they felt for the focal pen, how attractive and ergonomic they found it, and how important it was for them to write with an attractive and ergonomic writing instrument (Table 3).[5]

Results and Discussion

Manipulation Checks. The manipulation of the conditions for an escalation of commitment was successful (Table 3). While there was no difference between subjects who received a bag of M&Ms in BDM-M&Ms and in BRACKETS (t = -.14; p < .89), both felt more strongly than subjects in BDM-no M&Ms that they were getting something tangible from participating (t = 9.82; p < .0001 and t = 9.78; p < .0001, respectively).[6] Moreover, under BDM, subjects who received M&Ms felt more than those who did not that they would get a fair deal rather than walk away empty-handed if they did not transact ($t = 4.30$; $p < .0001$).

Table 3. Manipulation Checks, Distribution of WTP in Study 3 and Means of Respondent Characteristics across Conditions (standard deviations in parentheses; minimum prices in all conditions were zero; one missing value each in BDM-no M&Ms and in BRACKETS conditions)

QUESTION	BDM-M&Ms ($N = 80$)	BDM-no M&Ms ($N = 85$)	BRACKETS ($N = 90$)
Are you getting something tangible out of your participation in this survey? §§	6.64 [a] (2.21)	3.22 [a,b] (2.26)	6.59 [b] (2.29)
Would not being able to buy feel more like "walking away empty-handed" (= 1) or more like "a fair deal" (= 9)? §§	6.19 [a] (2.52)	4.34 [a] (2.99)	n.a.
Mean WTP §	\$1.24 [a] (.93)	\$1.33 [b] (1.21)	\$3.04 [a,b] (1.87)
How much do you normally pay for a ballpoint pen? (in \$)	\$ 1.12 (.86)	\$ 1.34 (1.49)	\$ 1.35 (1.18)
WTP minus normally paid price for pens	\$.12 [a] (1.11)	\$.15 [b] (2.45)	\$1.85 [a,b] **** (1.94)
How important was it to you to maintain control over whether or not you could buy? §§§	-.96 *** (2.34)	-.96 *** (2.48)	n.a.
Which of the three pens at the given prices would you most prefer to buy? [share of focal alternative, $\chi^2 = 22.8$***]	85% [a]	81% [b]	56% [a,b]
How much do you feel you need this particular ballpoint pen? §§	2.13 [c] (1.46)	2.44 (1.86)	2.69 [c] (1.93)
Weighted attractiveness of • How attractive and ergonomic did you find the ballpoint pen I offered you? • How important is it to you to write with an attractive & ergonomic writing instrument? §§, †	3.93 (1.68)	4.13 (1.99)	4.27 (1.70)

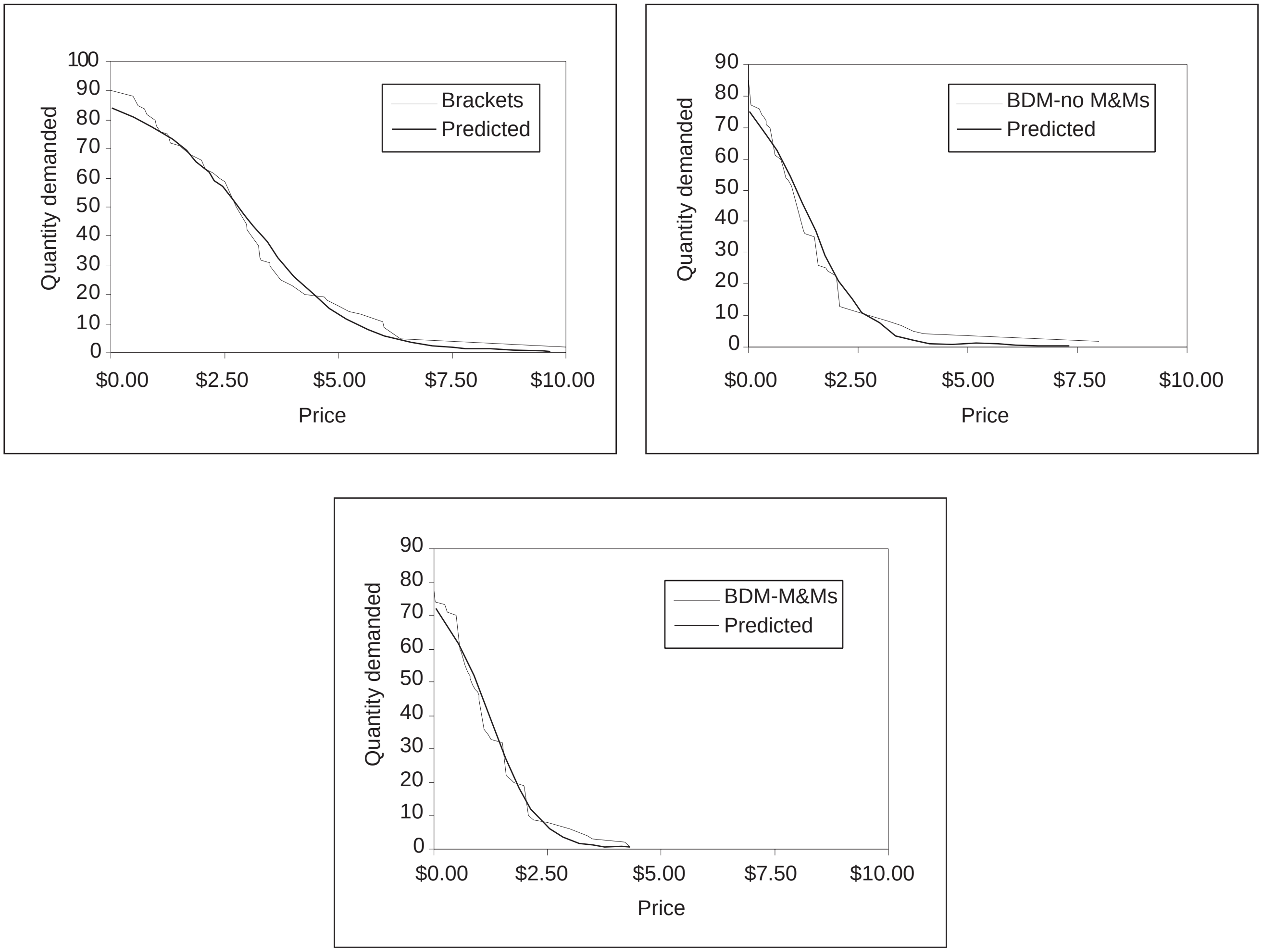
§ price is either final price offer (BDM) or stated price (BRACKETS)
§§ responses on 9-point-scale (1 = not at all, 9 = very much)
§§§ responses on 9-point-scale (1 = not at all, 9 = very much) re-scaled from –4 to +4
† square root of product of attractiveness and importance ratings
[a,b] values with same superscripts differ at $p < .0001$ in t-test with unequal variances
[c] values with same superscripts differ at $p < .05$ in t-test with unequal variances; different from 0 at *** $p < .001$, **** $p < .0001$
n.a. not applicable.

Table 4. Multiple Regression Predictors of WTP in Study 3 (standard errors in parentheses; two missing values each in BDM-no M&Ms and in BRACKETS conditions)

Question	BDM-M&Ms (*N = 80*)	BDM-no M&Ms (*N = 83*)	BRACKETS (*N = 88*)
Intercept	.19 (.25)	.48 (.32)	1.01 (.50)
How much do you feel you need this particular ballpoint pen?	.13 (.08)	.00 (.08)	.22 (.12)
Weighted attractiveness of • How attractive and ergonomic did you find the ballpoint pen I offered you? • How important is it to you to write with an attractive and ergonomic writing instrument? †	.16 * (.07)	.17 * (.07)	.18 (.14)
How much do you normally pay for a ballpoint pen? ($)	.15 (.11)	.10 (.09)	.53 ** (.17)
R^2	.22 ***	.10 *	.25 ****

Responses on 9-point-scale (1 = not at all, 9 = very much)
†square root of product of attractiveness and importance ratings
* $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$.

Figure 4. Observed and Predicted Demand in Study 3



Strategic Behavior. This experiment had three purposes. First, we wanted to test for strategic overbidding due to escalation of commitment or to a desire to maintain control over the purchase decision. As shown in Table 3, a *t*-test failed to reveal a difference in mean WTP between the two BDM conditions ($t = -.50$; $p < .62$), which we would have expected if BDM induced an escalation of commitment and led subjects to strategic overbidding in order not to walk away from the procedure empty-handed. Thus, reducing the conditions for such an escalation of commitment by paying subjects for their participation does not significantly dampen WTP under BDM. Also, if BDM encouraged overbidding because subjects with uncertain preferences wanted to make it to the second stage of the procedure to ensure that they would have the final choice of whether to buy or simply walk away, then they should display a desire to maintain control over whether or not to buy. But the negative mean ratings in Table 3 show that subjects in neither BDM condition felt it was important to maintain control over the purchase situation ($t = -3.68$; $p < .001$ with M&Ms and $t = -3.59$; $p < .001$ without M&Ms).

The Effect of the Incentive Constraint. The second purpose of this experiment was to examine if differences in WTP estimates under BDM and price matching could be due to the incentive constraint imposed under BDM or whether subjects simply devote insufficient cognitive resources to their price-matching responses. The BRACKETS condition had been designed to elicit comparable cognitive effort from subjects and make price equally salient so that any difference in WTP between BDM and BRACKETS would derive from the incentive constraint under BDM. Logit analyses of purchase probabilities $\Pr(\text{buy}|p) = e^{a+b \cdot p}/(1 + e^{a+b \cdot p})$ indicated downward sloping demand, with $a = 2.72$ ($p < .0001$) and $b = -.89$ ($p < .0001$) for BRACKETS, $a = 2.20$ ($p < .0001$) and $b = -1.84$ ($p < .0001$) for BDM-M&Ms, and $a = 1.96$ ($p < .0001$) and $b = -1.49$ ($p < .0001$) for BDM-no M&Ms. Analyses of Fisher-Z transformed correlations between observed and predicted demand (Figure 4) showed no differences in model fit between BRACKETS ($r = .992$, $Z = 2.74$) and BDM [$r = .989$, $Z = 2.61$ with M&Ms ($z = -.56$, $p < 1$) and $r = .988$, $Z = 2.56$ without M&Ms ($z = -.75$, $p < 1$)]. This suggests that BDM has similar reliability and internal validity as an elicitation method that requires subjects to consider more carefully the value they place on an object. However, Table 3 shows that, as predicted, mean WTP was higher in BRACKETS than under either BDM condition ($t = 7.45$; $p < .0001$ with M&Ms and $t = 5.21$; $p < .0001$ without M&Ms). As BRACKETS was designed to make price equally salient, this suggests that the incentive constraint alone can cause differences in WTP estimates between BDM and contingent-valuation approaches. This result replicates our findings in studies 1 and 2, which showed that WTPs under nonincentive-compatible price matching were higher than under BDM.

Domain of Applicability. The third purpose of the experiment was to examine the performance of BDM when applied not to food but to an inexpensive durable, for which demand depends on current inventory. Table 3 shows that a greater percentage of subjects preferred the focal pen (at their stated WTP) over the comparison pens (at the given prices) in either BDM condition than in BRACKETS ($\chi^2 = 22.8$; $p < .001$). This suggests that even for nonfood items, for which demand is not driven by visceral consumption impulses, BDM allows subjects to derive their WTP in a given purchase situation more accurately and with greater certainty about their preferences.

To determine the causes of this superior performance, we compared subjects' WTPs with their estimates of the prices they normally paid for a ball-point pen (Table 3). In BRACKETS, mean WTP exceeded the mean estimate of normally paid prices ($t = 8.96$, $p < .0001$), which was not the case in either BDM condition ($t = .95$; $p < 1$ with M&Ms and $t = .56$; $p < 1$ without M&Ms). We also regressed WTP on subjects' ratings of their need for the focal pen, the square root of the product of their ratings of the pen's attractiveness and their importance weights, and their estimates of normally paid prices (Table 4). The only predictor of WTP in BRACKETS was subjects' estimates of normally paid prices ($b = .89$, $t = 5.23$, $p < .0001$). Consistent with the findings in studies 1 and 2, this suggests that WTP is driven by subjects' price memories rather than an assessment of the value of the good in the current purchase context when it is elicited under a hypothetical response format. It is as if subjects start with what they normally pay and then adjust that estimate upward in order to derive their WTP. In contrast, WTP under BDM was predicted by the weighted attractiveness ratings ($b = .16$, $t = 2.30$; $p < .05$ with M&Ms and $b = .30$, $t = 2.51$; $p < .05$ without M&Ms), suggesting as predicted that BDM leads subjects to derive their WTP as a function of the perceived value of the good in the specific purchase situation.

Overall Discussion and Conclusions

Summary

Our studies demonstrate that BDM provides a feasible, reliable, and valid market research procedure to elicit consumer WTP in specific point-of-purchase settings in fast-moving consumer goods markets. BDM entails relatively little cost, time, and effort to administer. Studies 1 and 2 showed that face, internal, and criterion validities were high and compared favorably to a conventional open-ended contingent valuation approach of hypothetically asking consumers to state their WTP (price matching) for grocery items. Study 3 showed that BDM is equally well suited for inexpensive durables, that the incentive constraint is key to its performance, and that there was no evidence of overbidding, in contrast to empirical findings in tests of Vickrey auctions (Kagel 1995).

A key result across all three studies is that consumers reported substantially lower WTP under BDM than under hypothetical response formats. This difference adds to similar findings of hypothetical bias vis-à-vis other incentive-compatible formats across 39 studies reviewed by Harrison and Rutström (1999), including experimental results by Neill et al. (1994) for open-ended and Cummings, Harrison, and Rutström (1995) for dichotomous choice contingent valuation. For example, Neill et al. (1994) compared WTP in continent valuation with WTP in hypothetical and (otherwise identical) real Vickrey auctions. They found that WTP under the two hypothetical response formats exceeded WTP in real Vickrey auctions by far, implying that the lack of economic commitment rather than the absence of a structured institution in contingent valuation causes the bias in hypothetical WTP responses. This is also borne out in Study 3, with higher WTP in the fairly structured BRACKETS condition than in BDM. Our findings suggest that stated-preference methods may lead managers to overprice relative to consumers' true WTP, unless hypothetical survey responses can be recalibrated if the bias is known and stable (Harrison and Rutström 2001).

In addition, our results suggest that BDM provides a better measure of consumers' true point-of-purchase WTP because consumers in studies 1 and 2 tended to round prices when responses were only hypothetical, whereas BDM WTPs were more differentiated. Study 3 showed that nonincentive-compatible responses depend on the prices consumers normally pay in the category, whereas the incentive constraint under BDM helps respondents determine their WTP based on the point-of-purchase context.

Limitations and Future Research

Our research has a number of clear limitations. Although BDM is theoretically incentive-compatible (Kagel 1995), incentive-compatibility in practice requires unambiguous agreement about sellers' and buyers' goals. Thus, all field-based methods for measuring WTP, including BDM, are subject to possible uncontrollable belief-based strategic misrepresentation. Second, because BDM relies on actual transactions, it can only be applied to existing products (i.e., old products at new points of purchase or new products and prototypes). Unlike conjoint analysis, it cannot be used in concept design and new product development. Third, liquidity constraints may bias demand downward for higher priced products and big-ticket items. Yet, these constraints may be overcome by allowing participants to pay by check or credit card or to take a loan from the researcher (see Cummings, Harrison, and Rutström 1995). Importantly, if BDM is applied at the targeted point of purchase, the results will reflect how ready to buy consumers are under real transaction conditions. Conceptually, BDM is thus suited for products and services across all price ranges, given sufficient experimenter credibility and trustworthiness. Nonetheless, future research should examine the relative performance and feasibility of BDM for big ticket items. Finally, we tested BDM in monopolist purchase settings, as the specific point of purchase did not matter for our validation purposes. What happens when substitute products are available at the point of purchase? We predict that WTP varies inversely with the surplus consumers can derive from these substitutes.

Conclusion

BDM allows researchers to elicit WTP in an incentive-compatible manner in specific point-of-purchase contexts and under the controlled influence of marketing mix variables (e.g., WTP for candy bars on display at supermarket checkout counters where a retailer may try to induce impulse buying). BDM can serve as a standalone (off- or online) procedure as shown here, or it can be combined with existing preference elicitation techniques and pretest market research to develop better insights into the factors that influence consumer choice and product valuations (Huber et al. 1993).

BDM also provides an alternative to auctions in order to price-discriminate in regular online transactions (using a transparent online random device). The surface features of BDM might resemble those of reverse auctions. For example, at *Priceline.com* retail customers make binding bids, backed by a required credit card authorization. A participating vendor can then reject or accept the bid, depending on the desired margin. This enables vendors to generate incremental revenue and capacity utilization without disrupting their existing distribution channels or retail pricing structures. However, reverse auctions are not incentive-compatible. Customers pay what they bid (*s*) so they must bid less than their true WTP if they want to obtain surplus from the transaction. BDM solves that problem. The random mechanism provides an incentive for customers to bid their true WTP. Thus, a vendor can better price-discriminate against successful bidders.[7] The vendor can also make follow-up offers to those who are not allowed to buy, based on their WTP. For example, online auctioneer *Onsale* successfully targeted nonwinning bidders with tailor-made e-mail offers that were based on the bids it had elicited from them earlier (Moon 1999). So BDM could serve to sell Goethe's *Herrmann and Dorothea* under incentive-compatible conditions again one day.

Appendix: Instructions Given by the Interviewer

Below are the BDM instructions that the interviewer read to the subjects in studies 1 and 2.

Initial instructions

"Hello! I am a researcher from the University of Kiel and am conducting a marketing survey here on the beach (ferry). The survey takes only a few minutes. I am wondering if you would like to participate. You will need a small amount of money, because I will offer you an opportunity to buy a can of Coca-Cola (a piece of pound cake). You will not have to spend any more for the Coke (cake) than you really want to. I'd like to know how much money you are willing to spend for this can of Coke (piece of cake) here on the beach (ferry). The purchase price is not yet determined. Please tell me the highest price you would be willing to pay. You may then draw a ball from this urn. The balls are labeled with different prices. If you draw a price that is less than or equal to the price you tell me, you will have to buy the Coke (cake) for the price you drew from the urn. If the price you draw is greater than the price you tell me, you will not be able to buy the Coke (cake). This procedure ensures that it is best for you to truthfully reveal the maximum price you are willing to pay. If you tell me a price that is higher, you may actually have to pay that higher price. If you tell me a price that is lower, you may be disappointed if you can't buy if you draw a price that is higher than the price you tell me but lower than your 'true' price. Note that you cannot influence the purchase price with the price you tell me. Because you draw the purchase price from the urn, it is completely random and independent of whatever you tell me. Do you have any questions?"

Initial price offer

"Now, what is the maximum amount you are willing to pay for this can of Coke (piece of cake)? (*The subject states a price*.) If you now draw a price that is less than or equal to the price you just stated, I will sell you the Coke (cake) at the price you drew from the urn. However, if you now draw a price that exceeds the one you just stated, I will not sell you the Coke (cake)."

Option to revise

"If you now drew a price that is DM 0.10 higher than the price you just stated, would you consider buying the Coke (cake) after all? If so, please tell me the true maximum price, at which you would be willing to buy." (*The subject continues to state higher prices, until he or she would not consider a purchase anymore at a price that is DM 0.10 higher than stated.*)

Notes

1. Conjoint analysis is best described as incentive-neutral because respondents have no incentive to be inaccurate in indicating their preferences for price-quality bundles. We thank a reviewer for this suggestion.

2. Closed-ended contingent valuation does not provide a measure of WTP at the level of individual respondents (see Cameron and James 1987), which we require for our reliability and validity analysis.

3. We thank the editor and the reviewers for emphasizing these issues.

4. Continuous response formats (e.g., Vickrey auctions, BDM) assume that preferences adhere to expected utility theory. They lose their truth-revealing properties under violations of the independence axiom (e.g., Holt 1986). None of the reasons for misrepresentation imply a theoretical flaw in BDM but merely possible empirical challenges, as pointed out by one reviewer.

5. We also applied the same transparency, acceptability, compliance, satisfaction, and regret measures as in studies 1 and 2. The results were similar, so we do not report them here.

6. All *t*-tests are based on unequal variances. Between-subjects ANOVAs with planned contrasts of all dependent measures revealed the same effects.

7. For every purchase, the vendor forgoes, and the bidder keeps, WTP – s with Priceline and WTP – p with BDM. So vendors should prefer BDM if E(pBDM) > E(sPriceline), which depends on the distribution of buying prices (pBDM) that the vendor chooses. Note that the lower bound of the distribution would normally be given by the vendor's marginal cost. Vendors may give customers information on the underlying price distribution to create anchoring effects (Bohm, Lindén, and Sonnegård 1997).

References

Alberini, Anna, Barbara Kanninen, and Richard T. Carson (1997), "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data." *Land Economics* 73 (August), 309-24.

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

Ben-Akiva, Moshe, M. Bradley, T. Morikawa, J. Benjamin, Thomas P. Novak, P. Thomas, Hamen Oppewal, and Vithala Rao (1994), "Combining Revealed and Stated Preferences Data." *Marketing Letters* 5 (4), 335-50.

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

Bohm, Peter, Johan Lindén, and Joakim Sonnegård (1997), "Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanisms vs Markets." *Economic Journal* 107 (July), 1079-89.

Cameron, Trudy A., and Michelle D. James (1987), "Estimating Willingness to Pay from Survey Data: An Alternative Pre-Test-Market Evaluation Procedure." *Journal of Marketing Research* 24 (November), 389-95.

Carmon, Ziv, and Itamar Simonson (1998), "Price-Quality Tradeoffs in Choice vs. Matching: New Insights into the Prominence Effect." *Journal of Consumer Psychology* 7, 323-43.

Carson, Richard T., Theodore Groves, and Mark J. Machina (1999), "Incentive and Informational Properties of Preference Questions." San Diego, CA: University of California at San Diego, Department of Economics, Working paper.

_________, Jordan J. Louviere, Donald A. Anderson, Phipps Arabie, David S. Bunch, David A. Hensher, Richard M. Johnson, Warren F. Kuhfeld, Dan Steinberg, Joffre Swait, Harry Timmermans, and James B. Wiley (1994), "Experimental Analysis of Choice." *Marketing Letters* 5 (4), 351-68.

Casey, Jeff T., and Philippe Delquié (1995), "Stated vs Implicit Willingness to Pay under Risk." *Organizational Behavior and Human Decision Processes* 61 (February), 123-37.

Cummings, Ronald G., Glenn W. Harrison, and E. Elisabet Rutström (1995), "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive-Compatible." *American Economic Review* 85, 260-67.

Gabor, André, and C. W. J. Granger (1966), "Price as an Indicator of Quality: Report on an Enquiry." *Economica* 33 (February), 43-70.

Glass, G. V., and J. C. Stanley (1970), *Statistical Methods in Education and Psychology*. Englewood Cliffs, NJ: Prentice Hall.

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

________, Donald S. Tull, and Gerald Albaum (1988), *Research for Marketing Decisions*. Englewood Cliffs, NJ: Prentice Hall.

Hanemann, W. Michael, J. Loomis, and B. Kanninen (1991), "Statistical Efficiency of Double Bounded Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics* 73, 1255-63.

Harrison, Glenn W., and E. Elisabet Rutström (forthcoming 2001), "Experimental Evidence of Hypothetical Bias in Value Elicitation Methods." In *Handbook of Results in Experimental Economics*, eds. Charles R. Plott and Vernon L. Smith. New York: Elsevier.

Hoffman, Elizabeth, Dale J. Menkhaus, Dipankar Chakravarti, Ray A. Field, and Glen D. Whipple (1993), "Using Laboratory Experimental Auctions in Marketing Research: A Case Study of New Packaging for Fresh Beef." *Marketing Science* 12 (3), 318-38.

Holt, Charles A. (1986), "Preference Reversals and the Independence Axiom." *American Economic Review* 76 (June), 508-15.

Huber, Joel, Dick R. Wittink, John A. Fiedler, and Richard Miller (1993), "The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice." *Journal of Marketing Research* 30 (February), 105-14.

Jones, D. Frank (1975), "A Survey Technique to Measure Demand under Various Pricing Strategies." *Journal of Marketing* 39 (July), 75-7.

Kagel, John H. (1995), "Auctions: A Survey of Experimental Research." In *The Handbook of Experimental Economics*, eds. John H. Kagel and Alvin E. Roth, 501-85. Princeton, NJ: Princeton University Press.

________, Ronald M. Harstad, and Dan Levin (1987), "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study." *Econometrica* 55, 1275-1304.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy* 98, 1325-48.

Kalish, Shlomo, and Paul Nelson (1991), "A Comparison of Ranking, Rating and Reservation Price Measurement in Conjoint Analysis." *Marketing Letters* 2 (4), 327-35.

Loewenstein, George F. (1996), "Out of Control: Visceral Influences on Behavior." *Organizational Behavior and Human Decision Processes* 3 (March), 272-92.

Mitchell, Robert C., and Richard T. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Baltimore, MD: Johns Hopkins University Press.

Moldovanu, Benny, and Manfred Tietzel (1998), "Goethe's Second-Price Auction." *Journal of Political Economy* 106, 854-9.

Moon, Youngme (1999), "Onsale, Inc." *Harvard Business School Case Study #* 9-599-091.

Neill, Helen R., Ronald G. Cummings, Philip T. Ganderton, Glenn W. Harrison, and Thomas McGuckin (1994), "Hypothetical Surveys, Provision Rules, and Real Economic Commitments." *Land Economics* 70 (May), 145-54.

Rao, Vithala R., and Harsh Soni (1994), "Bundle Pricing Using the Balance Model." Ithaca, NY: Cornell University, Johnson Graduate School of Management, Working paper.

Rothkopf, Michael H., and Ronald M. Harstad (1995), "Two Models of Bid-Taker Cheating in Vickrey Auctions." *Journal of Business* 68 (April), 257-67.

Silk, Alvin J., and Glen L. Urban (1978), "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology." *Journal of Marketing Research* 15 (May), 171-91.

Thaler, Richard (1985), "Mental Accounting and Consumer Choice." *Marketing Science* 4, 199-214.

________, and Eric J. Johnson (1990), "Gambling with the House Money and Trying to Break Even: Effects of Prior Outcomes on Risky Choice." *Management Science* 36, 643-60.

Tversky, Amos, Paul Slovic, and Daniel Kahneman (1990), "The Causes of Preference Reversals." *American Economic Review* 80, 204-17.

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

Wertenbroch, Klaus (1998), "Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice." *Marketing Science* 17, 317-37.

For Further Reading

Other MSI working papers on pricing include:

00-117	**Justifying Profitable Pricing** by Joel E. Urbany
00-110	**What Makes Consumers Pay More for National Brands than for Store Brands: Image or Quality?** By Raj Sethuraman
99-128	**Price-matching Refund Policies as Signals of Store Price Image** by Joydeep Srivastava
98-131	**Transaction Decoupling: The Effects of Price Bundling on the Decision to Consume** by Dilip Soman and John T. Gourville
98-123	**Behavioral Explanations for Asymmetric Price Competition** by Makoto Abe
98-116	**Divide and Prosper: Effects of Partitioned Prices on Consumers' Price Recall and Demand** by Vicki G. Morwitz, Eric A. Greenleaf, and Eric J. Johnson
97-126	**Why Do Consumers Pay More for National Brands than for Store Brands?** By Raj Sethuraman and Catherine Cole
96-103	**The Effects of Price-Comparison Advertising on Buyers' Perceptions of Acquisition Value and Transaction Value** by Dhruv Grewal, Kent B. Monroe, and R. Krishnan
95-114	**Behavioral Perspectives on Pricing**, conference summary by Hooman Estelami and Peter Palij
95-105	**National Brand and Store Brand Competition: Who Hurts Whom?** By Raj Sethuraman
90-112	**Consumer Knowledge of Normal Prices: An Exploratory Study and Framework** by Joel E. Urbany and Peter R. Dickson

Other MSI working papers on measurement and analysis include:

99-121	**HEC Symposium on Advances in Choice Theory** edited by Gilles Laurent

98-115 **Field Experiments with Uni- and Multi-inventory Adaptive Survey Designs for Likert-type Data**
by Jagdip Singh

97-112 **When Do Purchase Intentions Predict Sales?**
by Vicki Morwitz, Joel H. Steckel, and Alok Gupta

93-114 **Seeing the Voice of the Customer: The Zaltman Metaphor Elicitation Technique**
by Gerald Zaltman and Robin A. Higie

91-120 **Measuring Overall Judgements and Attribute Evaluations: Overall First vs. Attribute First**
by K. Sridhar Moorthy

91-102 **Determining Inter-Brand Substitutability Through Survey Measurement of Consumer Preference Structures**
by Randolph E. Bucklin and V. Srinivasan

90-117 **Adaptive Designs for Likert-Type Data: An Approach for Implementing Marketing Surveys**
by Jagdip Singh, Roy D. Howell, and Gary K. Rhoads

90-101 **Direct Perceptual Measurement: Methodology and Validity**
by Michael D. Johnson and David A. Horne

MARKETING SCIENCE INSTITUTE
1000 Massachusetts Avenue
Cambridge, MA 02138 USA
Phone: 617.491.2060
Web: www.msi.org