DE GRUYTER OPEN — doi 10.2478 / gfkmir-2014-0040

{ *New Method* }

Measuring Consumers' Willingness to Pay
WHICH METHOD FITS BEST?

Klaus M. Miller, Reto Hofstetter, Harley Krohmer, and Z. John Zhang

Gauging the maximum willingness to pay (WTP) of a product accurately is a critical success factor that determines not only market performance but also financial results. A number of approaches have therefore been developed to accurately estimate consumers' willingness to pay. Here, four commonly used measurement approaches are compared using real purchase data as a benchmark. The relative strengths of each method are analyzed on the basis of statistical criteria and, more importantly, on their potential to predict managerially relevant criteria such as optimal price, quantity and profit. The results show a slight advantage of incentive-aligned approaches though the market settings need to be considered to choose the best-fitting procedure.

Superior Pricing – Higher Profitability

Prices have a strong and immediate impact on company profits. Therefore pricing decisions are gaining more and more importance. One major challenge for successful pricing is estimating the perceived value of a product to its customers correctly, which involves estimating customers' willingness to pay (WTP). Usually, if consumers are interested in buying a product or service, there is a maximum tolerable price, depending on the value the product can generate. If the product is more expensive consumers won't buy, but if the product is at or below that threshold they will purchase. Gauging this maximum price or maximum willingness to pay (WTP) accurately is necessary to position a product or service among competing offers, to decide optimal price related segmentation of a market and for decisions about changing or modifying prices. An underestimation or poor differentiation of that price may lead to wasted profit potential whereas an overestimation may mean losing potential customers.

Measuring Consumer's Willingness to Pay (WTP)

However, determining consumer's willingness to pay is not an easy task. First, it is challenging for consumers to actually estimate a product's value or to know what they might be prepared to spend, especially if a product is fairly new. Second, the consumer might know but be unwilling to say. Consumers might answer strategically, hoping that a lower stated willingness to pay will result in lower prices. Or there might be social influence at work. Respondents might overstate the amount they would spend because of the self-image they would like to create. Others again might refuse to talk about this issue at all.

Not surprisingly, there are many ways for measuring WTP as accurately as possible. One major distinction between the approaches is whether WTP is measured directly or indirectly. In practice, some marketing researchers favor the direct approach, asking consumers directly to state their WTP for a specific product through, say, an

THE AUTHORS

Klaus M. Miller is Assistant Professor of Marketing at the Institute of Marketing and Management, University of Bern, Switzerland, miller@imu.unibe.ch.

Reto Hofstetter is Assistant Professor of Marketing at the Center for Customer Insight, University of St. Gallen, Switzerland, reto.hofstetter@unisg.ch.

Harley Krohmer is Professor of Marketing, Chair of the Marketing Department, and Executive Director of the Institute of Marketing and Management, University of Bern, Switzerland, krohmer@imu.unibe.ch.

Z. John Zhang is the Murrel J. Ades Professor of Marketing at the Wharton School, University of Pennsylvania, USA, zjzhang@wharton. upenn.edu.

This article is an updated version of *Miller, K.M./Hofstetter, R./Krohmer, H./Zhang, Z.J.* (2011), "How Should Consumers' Willingness to Pay Be Measured? An Empirical Comparison of State-of-the-Art Approaches", Journal of Marketing Research, 48(1), 172 – 184 and is published with the permission of the American Marketing Association.

{ Box 1 }

ONLINE SURVEY – INNOVATIVE CLEANING PRODUCTS FOR HIGH-TECH EQUIPMENT

In the survey, 1,124 Swiss consumers were randomly assigned to one of five different experimental groups. In the ***open-ended question format (OE) group***, each participant had to directly state his or her individual hypothetical WTP for the cleaning product.

In the ***BDM group***, we determined actual WTP by using a BDM mechanism that had been applied successfully before. Participants were told that they were obligated to buy the cleaning product at the randomly determined price if the price was less than or equal to their stated WTP. However, if the randomly determined price was higher, a respondent would not have to buy the product. This mechanism ensures that participants have no incentive to indicate a price that is higher or lower than their true WTP.

In the ***CBC group***, we used a computer-generated, choice-based, conjoint design. We gave each respondent seven choice tasks and told them to imagine that he or she had to choose in an online shop among the product alternatives "right here" and "right now." Each choice task contained four cleaning products (i.e., conjoint stimuli) and a none-purchase option. Each conjoint stimulus was described by five attributes which we obtained in a pre-test. Attribute levels varied systematically (see Table 2).

In the ***ICBC group***, the conjoint procedure was exactly the same as for the CBC group. In addition, participants were informed that their responses in the conjoint task would be used to infer their WTP for a product. They were further instructed that after the completion of the survey, the product with the attributes preferred by the most people would be produced. The BDM mechanism embedded in CBC procedure ensured that participants had an incentive to reveal their true preferences.

In the ***REAL group***, we collected real transaction data by asking each participant whether he or she would be willing to buy the cleaning product at a certain price displayed in an online shop. The test site used for the experiment was similar to the real online shop of our cleaning product manufacturer. For the sake of comparability, price levels in the online shop corresponded to the price levels in our conjoint treatments (CBC and ICBC group). The price levels were randomly assigned to the participants, and each price level had an equal chance of appearing in the online shop.

Then, a series of statistical analyses were applied to compare the data sets and obtain WTP estimates.

open-ended question format. Others prefer an indirect approach, where WTP is calculated based on consumers' choices among several product alternatives and a none-choice option. Another classification discerns if consumers' hypothetical or actual WTP is determined. Table 1 gives an overview of the nature of the different approaches and lists the state-of-the-art methods for each field.

The different approaches all have advantages and drawbacks concerning the difficulties of measuring WTP, as described before. Hypothetical methods tend to overestimate WTP when compared with actual WTP from BDM and ICBC. However, there are situations when actual WTP cannot be measured in a study, e. g., when the prices at stake are very high or when products are highly individualized. Further, applying incentive-aligned approaches may not always be feasible due to the availability of product-prototypes or survey subjects and due to legal restrictions on the types of marketing research one can carry out. On the other hand, giving the respondents choice alternatives rather than direct questioning should make it easier for them to gauge their real preferences and actual value of alternatives.

Testing What Works Best Compared to Real Purchase Data

Making the right decision of how to measure WTP involves evaluating the drawbacks and advantages of each approach for each research setting individually. The type of products and the research objectives need to be clear to make a good decision. Many prior studies have tested differences among these approaches for different product types, but have not compared their results to what is ultimately of most interest: consumers' real WTP.

This contribution focuses on this "gap" and assesses whether the four approaches presented in Table 1 are statistically different from real purchase data, and which of these methods may lead marketing researchers to better pricing decisions. In a large-scale experimental design and field test (Box 1), WTP was collected for a new and inexpensive cleaning product for high-tech equipment (e.g., computer keyboards) using the four approaches and compared to real purchase data obtained from an online shop. The "five-in-one study" allowed a comprehensive assessment of each approach's ability to capture mean WTP and WTP distributions as well as managerially relevant criteria, such as the ability to predict the optimal price, quantity and profit to be expected. It also helps to better understand the strengths and weaknesses of each approach.

RESULTS OF THE COMPARISON BETWEEN METHODS

Comparing the Average WTP

In the dataset, all methods produce valid outcomes in measuring consumers' mean WTP. However, relative to the real purchase data, CBC shows by far the largest hypothetical bias (as can be seen from the ratio of the measured WTP to the benchmark) followed by OE, ICBC, and BDM. Hence, for our case study of an inexpensive cleaning product we can say that BDM performs best (see Table 3).

We further assessed the differences between the various methods to measure consumers' WTP. Here, directly stated WTPs in BDM and OE differed significantly (Δ = CHF 2.06) and indirectly stated WTPs in CBC and ICBC differed even more (Δ = CHF 5.52). One reason for the much higher WTP estimates in CBC may be that there were many more none choices under ICBC. In the ICBC group, 19 % of the participants chose the none-choice option, whereas under hypothetical CBC only 5 % chose the none-choice option and this difference results in a much larger intercept between the prices under CBC, but the difference in price sensitivity is small.

To sum up, mean WTP analysis showed statistically unbiased results for all methods. Further, we found that both hypothetical methods (OE, CBC), however, are significantly different from their incentive-aligned counterparts. The comparison of the means of the hypothetical methods with the real purchase benchmark showed that for our case study of an inexpensive cleaning product, CBC was more biased in absolute terms than OE.

As a consequence, hypothetical CBC may be more appropriate if a manager is primarily interested in the relative utilities of product attributes and price and less in predicting the actual best price.

Comparing WTP Distribution across Methods

Mean WTP is very important for both value-auditing and the valuation of a public good (e.g., clean air or water). However, for product pricing decisions, even an accurate estimate of mean WTP may not be very helpful to the marketing researcher for identifying the optimal price(s). For instance, if the data covers different segments with different value perceptions between segments, but similar evaluations within the segments, the mean might be misleading. Therefore, it is necessary to consider the entire WTP distribution (see Figure 1) in assessing the performance of an approach, not just the mean.

TABLE 1:
Overview of the Tested State-of-the-Art-Methods

Measuring Consumer's Willingness to Pay (WTP)		
	Direct Measurement (consumers state WTP directly)	**Indirect Measurement** (WTP is derived from several choice alternatives)
Hypothetical WTP (have no financial consequences for consumers)	**OE** Open-ended questions	**CBC** Choice-based conjoint analysis
	E.g. "What is the maximum you would be willing to pay to obtain X?"	WTP is calculated based on subjects' choices among several product alternatives and a none-choice option
Actual WTP (require real financial commitment, real purchases take place)	**BDM*** Incentive-compatible mechanism	**ICBC** Incentive-aligned choice based conjoint analysis
	A subject is obligated to purchase a product if the price drawn from a lottery is less than or equal to the subject's stated WTP	Subjects' WTP is inferred from their revealed preferences, they are also obligated to make a purchase using the BDM mechanism

* Becker, DeGroot, and Marschak (1964) mechanism

TABLE 2:
Attributes and Levels of the Cleaning Product Included in the Choice-Based Conjoint Analysis

Attribute	Levels	Number of Attribute Levels
Brand	CLEAN-A, CLEAN-B, CLEAN-C, CLEAN-D	4
Color	red, blue, green, yellow	4
Durability (Period of usage)	2 months, 4 months, 6 months, 8 months	4
Cleaning Power	Absorbs 90 % of dust & dirt, Absorbs 75 % of dust & dirt, Absorbs 60 % of dust & dirt	3
Price	CHF 1.59, CHF 4.79, CHF 7.95, CHF 11.10, CHF 14.30	5

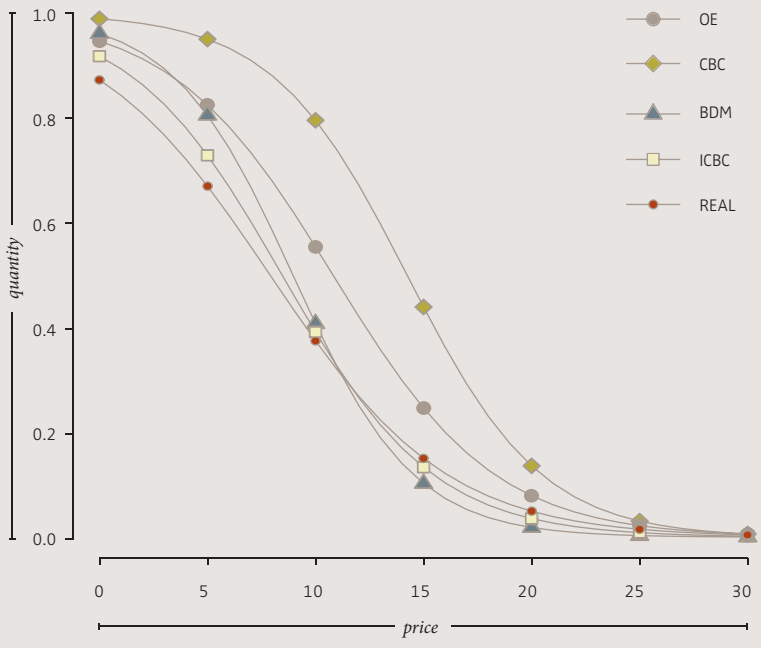
» The type of products and the research objectives need to be clear to make a good decision. «

The tests show that the demand curves of OE, BDM and ICBC are quite similar to the true demand curve from the real purchase data. However, significant differences in WTP distributions between CBC data and real purchase data could be observed. These results are consistent with the analysis of mean WTP values as discussed before. BDM tracks real demand best, followed by ICBC, OE, and CBC. It shows that even hypothetical methods can capture real demand well.

TABLE 3:
Comparison of Mean WTP and Confidence Intervals Across Methods

Method	n	Mean (Swiss Francs)	Ratio HWTP or AWTP / Benchmark
OE	279	11.03	1.30
CBC	310	14.92	1.76
BDM	183	8.96	1.06
ICBC	151	9.39	1.11
REAL (Benchmark)	201	8.46	n.a.

FIGURE 1:
Demand Curves for Cleaning Products Dependent on Measurement Approach



THE DIFFERENT APPROACHES AND BUSINESS DECISIONS

Do these differences matter for price-setting or sales forecasting, the ultimate test of a successful approach? An examination of how well each of these tools supports the business decision of choosing the profit-maximizing price can answer this question for the cleaning products in this study. First, the performance of the approaches in determining the demand curve within a range around the optimal price is compared. We then examine the ability of the different approaches to explicitly forecast the optimal price, quantity and profits.

Comparing Willingness to Pay Distributions Around the Optimal Price

The optimal price (CHF 8.50) and optimal price range based on market information (demand characteristics and costs) served as a starting point. Next, a confidence range for the optimal price of the real purchase data was constructed. Within this confidence range, we will find the optimal price with a probability of 95 %. Then comparing the WTP distributions from the various methods to the actual WTP within the confidence range produced the following results (see Figure 2). The diagrams show the optimal price and WTP along the demand curve of the real purchase data (filled circles) with WTP generated with the different approaches and the confidence ranges of each approach and the real data (each with dotted lines). The straight vertical line indicates the optimal price based on the real purchase data.

The WTPs from BDM overlap at any given price point in the range of the profit-maximizing price. In other words, the BDM data is very similar to the real purchase data. Further, partial overlaps can be observed for OE and ICBC distributions. However, CBC does not overlap at all in the relevant range for a pricing decision in our application. BDM shows the least deviation from the benchmark (Δ = .170), followed by ICBC (Δ = .661), OE (Δ = 1.840), and CBC (Δ = 4.376).

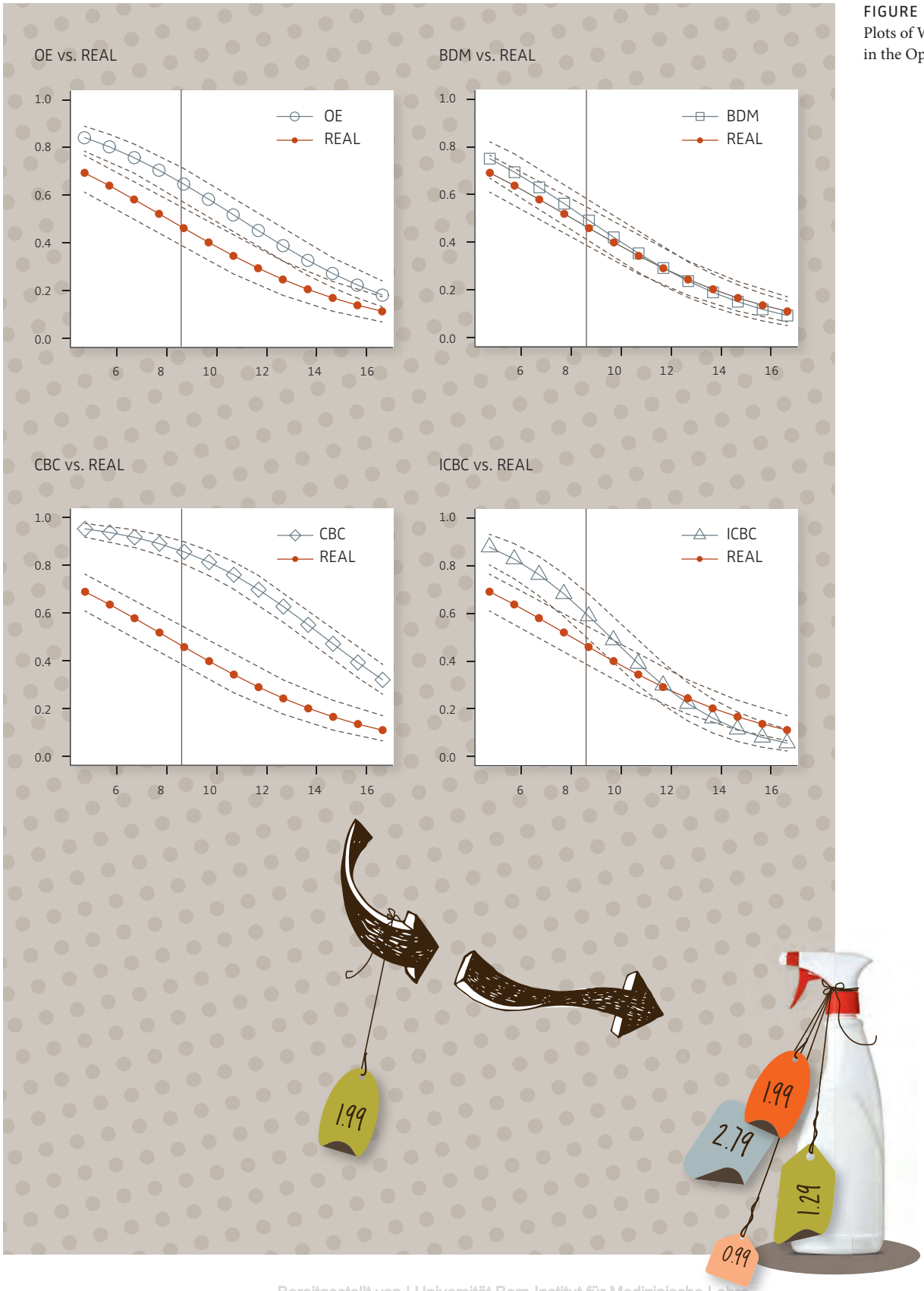


FIGURE 2:
Plots of WTP Distributions in the Optimal Price Range

Method	Optimal Price	Confidence Interval	Absolute difference to benchmark	Optimal Quantity	Confidence Interval	Absolute difference to benchmark	Optimal Profits	Confidence Interval	Absolute difference to benchmark
OE	9.681	[8.408, 11.353]	1.181	.576	[.466, .659]	.112	152,483.7	[134,948.4, 163,644]	45,978.6
CBC	11.494	[10.352, 12.837	2.994	.707	[.606, .777]	.243	225,799.3	[206,972.8, 237,568.7]	119,294.2
BDM	8.164	[6.782, 9.938	.336	.522	[.393, .617]	.058	114,459.9	[95,014.62, 125,842.5]	7,954.8
ICBC	7.925	[6.896, 9.342	.575	.652	[.523, .748]	.188	138,561.9	[120,520.8, 150133.3]	32,056.8
REAL	8.500	[6.872, 12.299	n.a.	.464	[.318, .602]	n.a.	106,505.1	[85,045.95, 123,140.7]	n.a.

Notes: Quantity scaled from [0,1], n.a. = not applicable

* The gray-shaded cells indicate that the confidence interval of the specific measure overlaps with the confidence interval of the corresponding benchmark measure obtained from the real purchase data. Hence, shaded areas imply no statistical difference between the estimated measure and the benchmark.

TABLE 4:
Overview of Estimates for Optimal Price, Quantity and Profit Across Methods

Comparing the Ability to Forecast Optimal Price, Quantity and Profits

Here the optimal price, quantity, and profit based on the real purchase data served as a benchmark for the performance of the individual approaches to measure consumers' WTP. The results of this analysis are summarized as follows:

> All methods seem to be equally able to forecast the ***optimal price***. All measures overlap with the confidence interval of the real data.

> For the ***optimal quantity***, only CBC performs significantly worse. Interestingly, the point estimates for the optimal price and quantity from OE and CBC do not fall in the confidence intervals generated with other methods for the respective measures, although the confidence intervals do overlap slightly.

> The findings for ***optimal profits*** paint a different picture. While forecasts from hypothetical approaches (OE and CBC) produce different results from the benchmark, those from incentive-aligned approaches do not. However, the absolute deviations from the benchmark are large for all approaches, especially for the hypothetical approaches. Therefore, managers should treat optimal profit estimates based on hypothetical data with care as these market research results may lead to significant economic differences. Finally, a rank order shows the following results: BDM yields the least deviation followed by ICBC, OE, and CBC for all three-point estimates (see table 4).

These findings suggest that in our application, the incentive-aligned methods are better able to forecast not only optimal price and quantity, but also profits. However, surprisingly, this analysis shows that hypothetical methods are also effective for forecasting optimal price and quantity, despite generating hypothetical bias.

» The BDM data is very similar to the real purchase data. «

Key Findings

> ***Incentive-aligned approaches performed best***
The results suggest that an incentive-aligned approach may be a more preferable choice for researchers and practitioners. However, this may not be true for all types of products (see next point). Further, other factors may limit the application of incentive-aligned approaches. For example, reasons such as cost, the unavailability of product-prototypes or survey subjects and legal restrictions as previously discussed.

> ***Type of product and purchasing context matter***
OE can outperform CBC in estimating mean WTP and WTP distribution, as well as making pricing decisions for an inexpensive, frequently purchased, non-durable product category like our cleaning product. According to previous findings, however, CBC may perform better when a product is less unique and faces more competing products, unlike the cleaning product in this study. Thus, indirect approaches such as conjoint analysis may be better suited for the product category where a more extensive decision process is involved (e.g., a digital camera) while direct approaches are less suitable for infrequently purchased products, and more suitable for offerings absent of any explicit competitive offering (e.g., products without any or only few direct competitors).

> ***Hypothetical bias might be less relevant***
Focussing on hypothetical bias in evaluating conjoint approaches is perhaps irrelevant for most marketing applications. Our analysis shows that even if a particular approach generates biased mean WTPs, and even if the estimated demand curve is different from the actual demand curve, the approach may still be useful in guiding marketing researchers to good pricing decisions. In particular, hypothetical CBC can be appropriate if managers are primarily interested in the relative utilities of product attributes and price and less in predicting the actual best price. If the research objective is to estimate WTP in relation to other product attributes, then OE and CBC can deliver valuable insights, despite some obvious concerns about the hypothetical nature of these approaches. •

FURTHER READING

Anders Gustaffsson, Andreas Herrmann, and Frank Huber, eds (2003): Conjoint Measurement. Methods and Applications, Berlin: Springer

Orme, Bryan K. (2003), "Which Conjoint Method Should I Use?" Sawtooth Software, ResearchPaper Series.

Voelckner, Franziska (2006), "An Empirical Comparison of Methods for Measuring Consumers 'Willingness to Pay", Marketing Letters, 17 (2), 137 – 149.

Wertenbroch, Klaus and Bernd Skiera (2002), "Measuring Consumers' Willingness to Pay at the Point of Purchase", Journal of Marketing Research, 39 (2), 228 – 241.

KEYWORDS:

Market Research, Pricing, Demand Estimation, Willingness to Pay, Hypothetical Bias