European Journal of Operational Research 219 (2012) 368–378

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Willingness-to-pay estimation with choice-based conjoint analysis: Addressing extreme response behavior with individually adapted designs

Sonja Gensler [a,*], Oliver Hinz [b], Bernd Skiera [c], Sven Theysohn [d]

[a] University of Groningen, Department of Marketing, Faculty of Economics and Business, P.O. Box 800, 9700 AV Groningen, The Netherlands
[b] TU Darmstadt, Department of Information Systems (Electronic Markets), Darmstadt, Germany
[c] Goethe-University Frankfurt, Department of Marketing, Faculty of Business and Economics, Frankfurt, Germany
[d] BASF AG, Ludwigshafen, Germany

ARTICLE INFO

Article history:
Received 17 December 2010
Accepted 2 January 2012
Available online 12 January 2012

Keywords:
Choice-based conjoint analysis
Willingness to pay
Marketing research

ABSTRACT

The increasing consideration of behavioral aspects in operations management models has prompted greater use of choice-based conjoint (CBC) studies in operations research. Such studies can elicit consumers' willingness to pay (WTP), a core input for many optimization models. However, optimization models can yield valid results only if consumers' WTP is estimated accurately. A simulation study and two field studies show that extreme response behavior in CBC studies, such that consumers always or never choose the no-purchase option, harms the validity of WTP estimates. Reporting the share of consumers who always and never select the no-purchase option allows for detecting extreme response behavior. This study suggests an individually adapted design that avoids extreme response behavior and thus significantly improves WTP estimation accuracy.

1. Introduction

Many critical business decisions occur at the intersection of operations and marketing as these functions strive to resolve the inherent tension between operational complexity and product differentiation (Karniouchina et al., 2009). Marketing aims to meet consumers' preferences by offering differentiated products; operations management must ensure business operations are efficient and use as few resources as possible. Increasing awareness of this critical intersection, prompted by the emerging market orientation of many firms, emphasizes the need to incorporate marketing aspects into operations management models (Bendoly et al., 2006; Karniouchina et al., 2009; Schlereth et al., 2010). In turn, choice-based conjoint (CBC) analysis, a popular market research method, recently has received more attention in operations management research (Halme and Kallio, 2011). Such analyses elicit consumers' preferences for product attributes and price by asking consumers to select products or no purchase repeatedly. The elicited information allows estimating consumers' utility function and provides insights into their preferences for product attributes and price, as well as their willingness to pay (WTP), defined as the price at which a consumer is indifferent between purchasing and not purchasing (Moorthy et al., 1997).

Optimization models in operations research use consumers' preferences and WTP estimates as critical inputs for product line and assortment decisions (e.g., Chen and Hausman, 2000; Kraus and Yano, 2003; Rusmevichientong et al., 2010; Scholl et al., 2005; Tarasewich and McMullen, 2001), the design of profit-maximizing products (e.g., Albers and Brockhoff, 1977; Easton and Pullman, 2001; Kohli and Krishnamurti, 1989; Verma et al., 2001), pricing decisions (e.g., Day and Venkataramanan, 2006; Green and Krieger, 1992), and the realignment of service operations (e.g., Coltman et al., 2010). Yet optimization models can yield valid results only if the input is valid. Thus valid estimates of consumers' WTP are crucial for the validity of optimization models. In particular, CBC studies require that the considered prices correspond with consumers' WTP. Expert interviews, pretests, and consumer feedback can indicate a reasonable price range, but they cannot guarantee sufficient overlap with the entire range of consumers' maximum and minimum WTP.

To understand the consequences of a lack of overlap between consumers' WTP and prices in CBC studies, consider a situation in which the prices in a CBC study are too high. In an extreme case, consumers always choose the no-purchase option, because this alternative has the highest consumer surplus. The researcher cannot identify consumers' actual purchase threshold (i.e., price at which consumers would buy a product), because extrapolation to lower prices is prone to error (Orme, 2002). Ultimately, predicted demand based on considered prices is zero, with no information about a profit-maximizing price. In the opposite situation, the offered prices in a

* Corresponding author. Tel.: +31 50 363 9086; fax: +31 50 363 8252.
E-mail address: s.gensler@rug.nl (S. Gensler).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2012.01.002

CBC study are too low, so consumers never choose the no-purchase option, because consumer surplus is greatest from choosing a product alternative. The researcher in this case cannot identify the price at which consumers stop buying. Extrapolation again should be avoided, so the derived price is the highest price in the CBC study, even though it may not be the profit-maximizing price. These two scenarios clearly are extremes; however, the prices in CBC studies easily can be too high or too low for some consumers. Thus, only a particular share of consumers always or never chooses the no-purchase option. We refer to both these options as extreme response behavior.

Recent CBC applications suggest the frequency of extreme response behavior. Parker and Schrift (2011), Sonnier et al. (2007) and Natter et al. (2008) identify 63.5%, 49.7%, and 22.8% of consumers, respectively, who never select the no-purchase option. No information about when these consumers stop buying a product is available, even though such information is critical for optimization models; skimming the purchases of these large groups of consumers can improve profits. Gilbride et al. (2008) also report that 35.1% of their consumers always selected the no-purchase option, which meant no information was available about whether these customers were simply not interested in the product or if they would buy the product at a lower price. In the latter case, offering the customers a lower price could improve a firm's profit, if the respective costs were low enough.

These authors offer no discussion of the effects of extreme response behavior on their results or how to address the potential effects. The same silence marks previous CBC literature that has focused on developing efficient designs based on predefined price levels (e.g., Kessels et al., 2006; Toubia et al., 2007), the role of the no-purchase option (e.g., Gunasti and Ross, 2009; Haaijer et al., 2001), specifications of the utility function (e.g., Meijer and Rouwendal, 2006; Sonnier et al., 2007), or algorithms to estimate the parameters of the utility function (e.g., Halme and Kallio, 2011; Karniouchina et al., 2009).

Spurred by this deficiency, we pursue a twofold goal. First, we show that extreme response behavior in CBC studies leads to invalid estimates of consumers' WTP, even if only a partial share of consumers exhibits this behavior. Second, we develop an individually adapted design for CBC studies that can capture heterogeneity in consumers' WTP but remains easy to use and implement. The proposed approach results in valid estimates for consumers' WTP, which then can inform optimization models in operations research, such as those focused on product lines and prices. Our work contributes to research at the intersection of marketing and operations research: Researchers using CBC studies to derive consumers' WTP should check for extreme response behavior in their data and report two measures that reveal extreme response behavior: the shares of consumers who never and who always choose the no-purchase option. Thus operations management and marketing researchers can use our findings to improve their CBC analyses of consumers' WTP and, ultimately, the prediction of demand.

2. Impact of consumers' response behavior on estimated WTP

2.1. WTP estimation with CBC analysis

In a CBC analysis, a repeated-design choice experiment, consumers repeatedly select one product out of a set or choose the no-purchase option (see Fig. 1 for an example).[1] The underlying assumption is that consumers choose the most preferred alternative, with the highest net utility and consumer surplus (Louviere and Woodworth, 1983). The no-purchase option allows them to indicate that no product alternative is acceptable (Parker and Schrift, 2011).

A logit model formulation describes consumers' choice behavior. That is, the probability that consumer h chooses product i in choice set a is equal to:

$$\begin{aligned} P_{h,i,a} &= \frac{\exp(u_{h,i})}{\exp(u_{h,NP}) + \sum_{i' \in I_a} \exp(u_{h,i'})} \\ &= \frac{\exp\left(\sum_{j \in J}\sum_{m \in M_j} \beta_{h,j,m} \cdot x_{i,j,m} - \beta_{h,\text{price}} \cdot p_i\right)}{\exp(\beta_{h,NP}) + \sum_{i' \in I_a} \exp\left(\sum_{j \in J}\sum_{m \in M_j} \beta_{h,j,m} \cdot x_{i',j,m} - \beta_{h,\text{price}} \cdot p_{i'}\right)}, \\ &\forall h \in H,\ i \in I,\ a \in A, \end{aligned} \tag{1}$$

and the probability that consumer h chooses the no-purchase option is

$$P_{h,NP,a} = \frac{\exp(\beta_{h,NP})}{\exp(\beta_{h,NP}) + \sum_{i' \in I_a} \exp\left(\sum_{j \in J}\sum_{m \in M_j} \beta_{h,j,m} \cdot x_{i',j,m} - \beta_{h,\text{price}} \cdot p_{i'}\right)} \quad \forall h \in H,\ i \in I,\ a \in A. \tag{2}$$

In these equations,

- $P_{h,i,a}$ probability that consumer h chooses product i in choice set a
- $P_{h,NP,a}$ probability that consumer h chooses the no-purchase option in choice set a
- $u_{h,i}$ utility of product i for consumer h
- $u_{h,NP}$ utility of no-purchase option for consumer h
- $\beta_{h,j,m}$ parameter of the level m of attribute j for consumer h (assuming part-worth model)
- $x_{i,j,m}$ binary variable indicating whether product i features level m of attribute j
- $\beta_{h,price}$ price parameter for consumer h (assuming vector model)
- p_i price of product i
- $\beta_{h,NP}$ parameter (utility) for the no-purchase option for consumer h
- A index set of choice sets
- H index set of consumers
- I index set of products
- I_a index set of products in choice set a (not including the no-purchase option)
- J index set of attributes without price
- M_j index set of levels for attribute j

Because WTP differs across consumers, the parameters in Eqs. (1) and (2) can vary too, so latent class and hierarchical Bayesian methods can be used to derive individual parameters (Andrews et al., 2002; Karniouchina et al., 2009; Natter and Feurstein, 2002). In a latent class multinomial logit (MNL) model that maximizes the likelihood function, we can use information about the posterior segment membership probability to derive individualized parameters based on the estimated segment-specific parameters (Wedel and Kamakura, 2000). By deriving individual parameters for the product attributes ($\beta_{h,j,m}$), price ($\beta_{h,price}$), and no-purchase option ($\beta_{h,NP}$), we can gain information about a consumer's WTP, or the price at which consumer h is indifferent between purchasing and not purchasing product i (Moorthy et al., 1997):

$$\sum_{j \in J} \sum_{m \in M_j} \beta_{h,j,m} \cdot x_{i,j,m} - \beta_{h,price} \cdot WTP_{h,i} = \beta_{h,NP} \quad \forall h \in H,\ i \in I. \tag{3}$$

Rearranging Eq. (3) yields:

$$WTP_{h,i} = \frac{1}{\beta_{h,\text{price}}} \cdot \left(\sum_{j \in J} \sum_{m \in M_j} \beta_{h,j,m} \cdot x_{i,j,m} - \beta_{h,NP} \right) \quad \forall h \in H,\ i \in I. \tag{4}$$

[1] The following discussion holds for other types of choice experiments; it is not restricted to CBC analysis.

Product 1	Product 2	Product 3	No-purchase option
Connection to integrated DVD burner	No possibility to archive recordings	Connection to home network	I would not buy any of the offered products.
No access to online video rental service	Access to online video rental service	Access to online video rental service	
Recommendation system based on experts	Recommendation system based on own behavior	Recommendation system based on other users' behavior	
News, erotic and sport	Movies and documentaries	TV series and comics	
9.95 EUR	14.95 EUR	4.95 EUR	
☐	☐	☐	☐

Fig. 1. Example choice set in a CBC study (digital video recorder).

2.2. Effect of extreme response behavior on consumers' WTP

When designing a CBC study, the researcher aims for well-balanced response behavior; consumers should select product alternatives in some choice sets and the no-purchase option in others. Such response behavior ensures sufficient information to estimate the parameters of the utility function accurately. However, if the price levels do not overlap with a consumer's WTP range, consumers instead exhibit extreme response behavior.

We distinguish two types of extreme response behavior: consumers *always* choose the no-purchase option, or consumers *never* choose the no-purchase option. Imagine again the scenario that all consumers always choose the no-purchase option because the price levels in the choice experiment are higher than all consumers' WTP. So the estimated parameter for the no-purchase option is relatively large and positive, and we would predict choices of the no-purchase option (Eq. (2)). The estimated WTP is very small and perhaps even negative (Eq. (4)). Moreover, the estimated parameters for the attribute levels and price are neither accurate nor efficient, because we have limited information to estimate these parameters (Brazell et al., 2006). Imagine the scenario that all consumers always choose a product, and the estimated parameter for the no-purchase option is relatively large and negative, with a nearly zero probability of choosing the no-purchase option (Eq. (2)). The estimated parameter for the no-purchase option is neither accurate nor efficient in this case. In addition, if the price levels are too low, consumers do not consider price a critical product attribute that affects their choices. The estimated price parameter is relatively small and perhaps even positive, leading to a high estimated WTP (Eq. (4)), likely outside the range of considered price levels in the CBC study.

Although both these types of extreme response behavior result in inefficient and invalid parameter estimates, we recognize the extremity of these scenarios, so we use a simulation study to demonstrate that the accuracy of the parameter estimates is affected even if only a share of consumers exhibits extreme response behavior. A CBC study with extreme response behavior thus leads to poor solutions in operation research models.

2.3. Simulation study to illustrate effect of consumers' response behavior on WTP

2.3.1. Data generation

We generated data by simulating the response behavior of 300 consumers who belonged to one of three underlying segments of equal size (100 consumers each), to account for heterogeneity in the market (Andrews et al., 2002). The design consisted of 12 choice sets with three products and a no-purchase option; each product featured four attributes (three attributes had three levels and one attribute had two levels) and the price. Moreover, we considered three holdout choice sets to assess predictive validity.

The average WTP values of the different segments (WTP_s) were 15.00€, 25.00€, and 35.00€, respectively. Each segment assigned a different weight to the four attributes: $w_1 = \{0.2, 0.4, 0.2, 0.2\}$, $w_2 = \{0.6, 0.2, 0.1, 0.1\}$, and $w_3 = \{0.3, 0.2, 0.4, 0.1\}$. Therefore, the average WTP for every segment and attribute, $WTP_{s,j}$, equals $w_{s,j} \times WTP_s$. We generated a WTP for every attribute level from left-censored normal distributions with a mean equal to $WTP_{s,j}$ and a variance of 2.00€, which allows for within-class heterogeneity. The WTP for the attribute levels were independent, normally distributed variables; their sum was normally distributed with means of 15.00€, 25.00€, and 35.00€, respectively, and a variance of 8.00€. We added an error variance of 4% to $WTP_{h,i}$ to simulate potential response errors (Andrews et al., 2002).

We manipulated the overlap between the considered price range and consumers' WTP range to influence the share of consumers who show extreme response behavior by allotting three distinct levels to the price variable: 100% overlap, 50% above or below, or 100% above or below. With 100% overlap, the price range (highest to lowest price level) and consumers' WTP range correspond, whereas in the 50% above or below conditions, only half of the WTP range overlaps with the price range, and the other half is higher or lower than that range. Finally, in the 100% above or below condition, the WTP range and price range do not overlap at all. For the price ranges, we established a low price that extends from 7.50€ to 25.00€; medium prices from 22.50€ to 35.00€; and high prices from 25.00€ to 45.00€. These levels generated data sets with different shares of consumers who exhibited extreme response behavior (Table 2), similar to existing field studies (e.g., Gilbride et al., 2008; Natter et al., 2008; Sonnier et al., 2007), including sets in which all consumers show extreme response behavior.

By applying a common random numbers variance reduction technique and using the same population of consumers in all scenarios (i.e., within-subject design), we ensured that changes in the variables of interest reflect the effect of the overlap only. We used one replication, similar to Andrews et al. (2002). The simulation was written in C# under the.net framework; the source code for the simulation is available on request.

2.3.2. Estimation and performance measures

We estimated a latent class MNL model with three segments using the Expectation–Maximization algorithm (Wedel and Kamakura, 2000). We then derived individual WTP estimates on the basis of the posterior probabilities of segment membership. We use a latent class MNL model to derive the individual parameter estimates; previous research has shown that the performance of hierarchical Bayesian methods depends on the parameterization of the objective function (Sonnier et al., 2007). Applying a latent

Table 1
Overlap of considered prices and consumers' WTP range: estimation accuracy and prediction accuracy.

	Consumers who always choose product (%)	Consumers who always choose no-purchase (%)	MAPE (WTP) (%)	First choice hit rate (%)
100% overlap	1.00	0.00	17.57	86.73
50% above	0.00	34.67	37.03	97.40
100% above	0.00	100.00	1,891.03	87.60
50% below	43.67	0.00	24.10	83.30
100% below	89.33	0.00	100.00	100.00
Mean			413.95	91.01

Notes: MAPE = mean absolute percentage error.

class mitigates this problem (Meijer and Rouwendal, 2006). For our performance measures, we included WTP estimation accuracy, measured by the mean absolute percentage error (MAPE), and prediction accuracy, measured by the first choice hit rate. All measures were calculated on the basis of the products in the three holdout choice sets.

2.3.3. Results

As we show in Table 1, MAPE (WTP) is strongly affected by the overlap. When consumers' WTP and price range completely overlapped, the MAPE (WTP) was 17.57%; when this overlap decreased, the MAPE (WTP) increased substantially (37.03% and 24.10 % for 50% above and below, respectively). Consumers' response behavior affects estimated WTP, even if only some consumers exhibit extreme response behavior.

Furthermore, the overlap between prices and consumers' WTP range influences prediction accuracy, though this relationship is less clear. Prediction accuracy equaled 100% in the 100% below scenario: Consumers' WTP was lower than the prices listed for all product alternatives, so they always chose the no-purchase option what is easy to predict. Thus prediction accuracy is not a good measure for assessing model performance when extreme response behavior occurs.

3. CBC research and extreme response behavior

Since Louviere and Woodworth (1983) first introduced CBC analysis to marketing literature, many marketing studies have applied the methodology to elicit consumers' preferences. Its popularity as a measure of consumers' preferences also has been driven by new developments for estimating the parameters of the utility function at segment and individual levels with latent class and Bayesian methods (Halme and Kallio, 2011). Applications of CBC analysis in operations research also have increased, due to the attention to behavioral aspects in operations management models (e.g., Coltman et al., 2010; Day and Venkataramanan, 2006; Rusmevichientong et al., 2010).

Other studies aim to develop CBC analysis further. The research areas most closely related to consumers' response behavior are

- development of efficient designs,
- role of no-purchase option, and
- specification and estimation of utility function.

Regarding *efficient designs* for CBC studies, the central question is how to design choice sets so that the resulting parameter estimates are statistically efficient. Creating efficient designs requires some knowledge about the parameters to be estimated. Research therefore has used managers' prior beliefs about parameter estimates (Sandor and Wedel, 2001). Other studies have suggested various optimization criteria such as D-, G-, M- or V-optimality (Arora and Huber, 2001; Huber and Zwerina, 1996; Kanninen, 2002; Kessels et al., 2006; Kuhfeld et al., 1994; Toubia and Hauser, 2007; Vermeulen et al., 2008). These optimization criteria rely on predefined attribute levels and thus an a priori selection of appropriate price levels. Offering a wide price range and more price levels seemingly should help address the problem of extreme response behavior, but such considerations can influence consumers' behavior by decreasing the utility balance in the choice sets (Huber and Zwerina, 1996). Moreover, the designs developed by these optimization criteria are adapted across consumers and do not capture consumer heterogeneity.

Instead, recent studies propose individualized (adaptive) designs. Toubia et al. (2007) suggest framing choice questions so that they reveal more information about consumers' preferences. The set of choice questions and their corresponding answers define a polyhedron, or the feasible parameter estimates for consumers' preferences that fit previous observations. Each choice narrows the range of feasible parameter estimates until the range equals a single set. The method works well when the predefined prices overlap with consumers' WTP range, but if the predefined price levels are too low or too high, consumers still show extreme response behavior. In response, Eggers and Sattler (2009) propose asking consumers directly about their WTP and then defining the price levels accordingly. They ask consumers to identify the best and worst level for each attribute, then ask them to state their WTP for two products: the best possible and the worst possible. The price levels thus are adapted to each consumer's WTP range. However, their approach can result in a large WTP range for an individual consumer that increases the difficulty of defining the price levels in the choice tasks. For example, utility balance may not arise in some choice sets, which would affect the efficiency of the parameter estimates. Overall, existing approaches to determine efficient designs cannot circumvent extreme response behavior effectively.

Another stream of research focuses on the *no-purchase option*, whose presence makes the experimental setting more realistic, helps eliminate statistical biases, and improves demand estimates (Parker and Schrift, 2011). Including the no-purchase option is especially crucial for determining consumers' WTP, because it can estimate the purchase threshold (Eq. (4)). Previous studies mainly investigate factors that result in the choice of the no-purchase option (e.g., Dhar, 1997; Gunasti and Ross, 2009) or discuss the best way to parameterize it (Haaijer et al., 2001). No studies have examined the effect of choosing or not choosing a no-purchase option on parameter estimates.

Finally, several studies compare different ways to *specify or estimate the parameters of the utility function* (e.g., Andrews et al., 2002; Evgeniou et al., 2007; Sonnier et al., 2007). Sonnier et al. (2007) recognize that extreme values for consumers' WTP and misleading results occur in random coefficient models (such as hierarchical Bayesian models), because the price parameter approaches zero or becomes positive. They compare direct and indirect specifications of the distribution of equalization prices,[2] allowing for random parameters for the attribute and price estimates, and assert that a parameterization of the likelihood function that directly identi-

[2] The equalization price is the maximum amount a consumer will pay for level changes in product attributes.

fies equalization prices is preferable, because it avoids invalid equalization price estimates when the distribution of the price parameter has a mass at zero. However, they do not question the reason for this mass, which occurs primarily when consumers rarely or never choose the no-purchase option (i.e., 49.7% of consumers in their study). The prices seem too low for these consumers and exert no influence on their choice decisions. Specifying the utility function based on equalization prices may retrieve reasonable WTP estimates with a random coefficient model (Meijer and Rouwendal, 2006), but it cannot address extreme response behavior. Furthermore, studies that focus on estimating the parameters of the utility function usually compare different optimization methods and use predictive validity to evaluate their performance (Evgeniou et al., 2007; Halme and Kallio, 2011; Karniouchina et al., 2009). However, as we show in Table 1, prediction accuracy is not an appropriate measure of model performance when the aim is to gain insights into consumers' WTP. The accuracy of the estimated parameters is influenced not only by the estimation procedure but even more substantially by consumers' response behavior.

As this review of existing literature indicates, extreme response behavior has received no notable attention. The consequences of extreme response behavior remain poorly understood, and we lack approaches to overcoming it. We thus suggest an individually adapted design for CBC studies to avoid extreme response behavior; it does not use predefined price levels but rather adapts prices individually to each consumer's WTP range, such that the accuracy of consumers' WTP estimates improves.

4. Using individually adapted prices in CBC studies

4.1. Individually adapted CBC analysis

The premise behind an individual adaptation of the price levels is to identify a consumer's WTP and let the prices of the product alternatives in the choice sets oscillate around this level. We systematically vary prices in the choice sets depending on a consumer's behavior, without limiting the prices to a predefined price range. When a consumer selects a product alternative, his or her WTP must be equal to or higher than the actual price of at least one product alternative in the choice set. When a consumer selects the no-purchase option, his or her WTP must be lower than the actual prices of all product alternatives in the choice set. We therefore continuously adapt prices upward as a consumer selects a product alternative and downward as he or she selects the no-purchase option. The upward and downward adaptation of prices follows an algorithm that steadily reduces the latitude of price changes to identify a consumer's individual WTP. Using this approach, we strongly reduce the likelihood of extreme response behavior. We call this approach individually adapted choice-based conjoint (IACBC) analysis.

To implement this individually adapted design, we first generated an efficient randomized design based on a complete enumeration strategy that aimed to produce a nearly orthogonal design. The product alternatives in each choice task were as different as possible to ensure minimal attribute-level overlap. We used placeholders for the price levels (e.g., low, medium, high) when generating the choice sets, then replaced them in the first choice set with predefined starting prices (which reflect the researcher's best knowledge about consumers' WTP). In each subsequent choice set, we adjusted the prices according to the decision in the preceding choice set by multiplying or dividing them by some factor $f(n)$, where n is the number of shifts in direction between purchase and no-purchase decisions. The function $f(n)^z$ converges to 1 and calibrates the amplitude of oscillation and the pace of convergence. At the start, z equals 1 and remains 1 if the previous decision is a purchase decision, not the no-purchase option ($d_{h,NP,a} = 0$). Instead, z moves to -1 if the previous decision is no-purchase ($d_{h,NP,a} = 1$). We then divided (multiplied) the prices by $f(n)$ if the consumer chose the no-purchase (product) option. This process mimics a binary search algorithm, which makes progressively better guesses and eventually converges to the sought value (Ahuja et al., 1999): Our algorithm employs progressively better guesses about a consumer's WTP to calculate prices around his or her WTP, with an alternating elicitation of prices below and above a recent guess about this consumer's WTP.

Fig. 2 depicts a flowchart of the algorithm. Conditionals, which appear as diamonds, typically contain true/false tests. Assume, for example, a function $f(n)^z = (1 + 2/(n + 1))^z$. The selection of a product alternative in the first choice set causes the starting prices to be multiplied by 2 for the following choice set, whereas the selection of the no-purchase option causes the starting prices to be divided by 2. The first shift in direction decreases the multiplier or divisor to 1.66, the second shift in direction to 1.5, and so forth. In addition to generating purchase and no-purchase decisions for each consumer, the declining latitude of price changes around the individual WTP leads to more challenging choice decisions as the utility balance among product alternatives and the no-purchase option increases (Hauser and Toubia, 2005).

The individual adaptation of prices likely increases awareness of price and product attributes in a choice set. Therefore, it should encourage consumers to better think about their individual preferences (Wathieu and Bertini, 2007). The induced well-balanced response behavior also should improve the efficiency of the price and no-purchase parameter (Park et al., 2008). By using all observations to estimate the parameters, this approach circumvents an endogeneity bias in the parameter estimates (Liu et al., 2007).

4.2. Simulation study

The setup of our simulation study is similar to the previous one: We created a market with 300 consumers who belonged to one of three underlying segments. We used 12 choice sets with three products and a no-purchase option plus three holdout choice sets for assessing predictive validity. The price featured three distinct levels, and each segment prefers the four attributes according to its own weights. Three attributes again had three levels, and one attribute had two levels. However, in contrast with our first simulation study, we individually adapted prices, in an attempt to (1) assess the WTP estimation accuracy and prediction accuracy of the proposed approach for different options and (2) compare the WTP estimation accuracy, prediction accuracy, and parameter efficiency of the individually adapted design against these values in a design with predefined prices.

The starting prices are the same as those we used in the traditional fixed design, but in the subsequent choice tasks, the prices get individually adapted according to consumers' behavior. Table 2 illustrates the adaption of the price levels for three different scenarios (100% overlap, 50% above, and 50% below) when we use the following algorithm: $f(n)^z = (1 + 2/(n + 1))^z$ for an individual consumer. The average actual WTP for the three segments was 15 EUR, 25 EUR, and 35 EUR. In the 100% overlap scenario, for example, the consumer chose product alternative 2 in the first choice set, so the shown prices doubled in the second choice set. Next the consumer chose the no-purchase option, which decreased the prices in the third choice set, and so forth.

In this simulation study, we also tested whether different options for adapting the prices affect WTP estimation accuracy and prediction accuracy. We manipulated the number of choice tasks (9, 12, and 18), the number of choice sets in which the prices are individually adapted (3, 6, and all), and the algorithm used to adapt prices ($f(n) = 1 + 1/n$, $f(n) = 1 + 2/(n + 1)$, and $f(n) = 1 + 4/(n + 3)$). If not all prices are individually adapted, the prices in the static

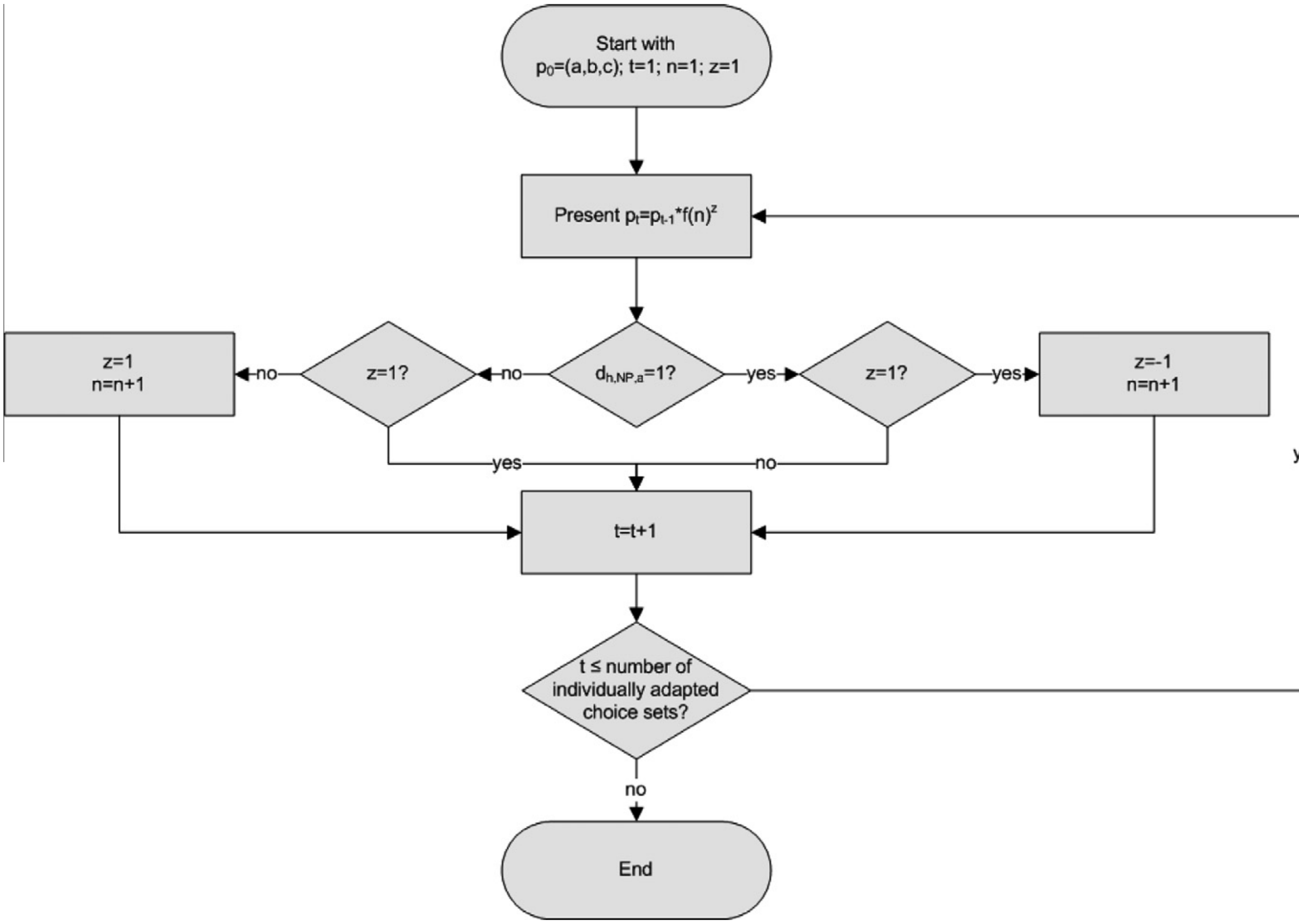

Fig. 2. Flowchart for adapting prices individually. Notes: p_0 is a vector with starting price levels a, b, c; $p_{t(t-1)}$ is a vector with price levels in choice set $t(t-1)$.

Table 2
Illustration of IACBC approach for three scenarios (100% overlap, 50% above, and 50% below).

Set	Alt	A1	A2	A3	A4	100% Overlap		50% Above		50% Below	
						Price	Choice	Price	Choice	Price	Choice
1	1	L3	L2	L2	L1	35.00	0	22.50	0	45.00	0
1	2	L2	L3	L3	L2	15.00	0	7.50	1	25.00	0
1	3	L1	L1	L1	L2	25.00	0	15.00	0	35.00	0
1	None						1		0		1
2	1	L3	L3	L1	L1	7.50	1	15.00	0	12.50	0
2	2	L2	L1	L2	L2	17.50	0	45.00	0	22.50	0
2	3	L1	L2	L3	L1	12.50	0	30.00	0	17.50	0
2	None						0		1		1
3	1	L3	L3	L3	L2	20.83	0	18.00	0	8.75	0
3	2	L2	L1	L1	L1	29.17	0	27.00	0	11.25	0
3	3	L1	L2	L2	L2	12.50	1	9.00	1	6.25	1
3	None						0		0		0
4	1	L1	L1	L3	L1	20.83	0	13.50	1	10.42	1
4	2	L3	L2	L1	L2	48.61	0	40.50	0	18.75	0
4	3	L2	L3	L2	L1	34.72	0	27.00	0	14.58	0
4	None						1		0		0
5	1	L2	L2	L3	L2	23.15	0	40.50	0	24.31	0
5	2	L1	L3	L1	L2	32.41	0	60.75	0	31.25	0
5	3	L3	L1	L2	L1	13.89	1	20.25	0	17.36	1
5	None						0		1		0
...											

Notes: Set = choice set, Alt = alternative (product or no-purchase option), A1–A4 = attributes 1 through 4, and L1–L3 = generic levels of attributes.

choice sets remain at the levels at which the last purchase decision occurred.

We expect that both the accuracy of the WTP estimates and prediction accuracy improve with increasing numbers of individually adapted choice sets and choice sets. The different algorithms for adapting the prices start with a multiplier of 2 and asymptotically converge to 1, but the pace of convergence varies. The oscillation amplitude converges rather quickly with $f(n) = 1 + 1/n$ but relatively slowly with $f(n) = 1 + 4/(n + 3)$. We expect that a faster convergence improves the accuracy of WTP estimation and prediction accuracy, because prices oscillate faster around a consumer's WTP.

We estimated a latent class MNL model with three segments and then derived individual WTP estimates on the basis of the posterior probabilities of segment membership. As performance measures, we again considered WTP estimation accuracy, measured by MAPE (WTP), and prediction accuracy, measured by the first choice hit rate. In Table 3, we provide these results. Noting the subtle differences in MAPE (WTP), we used an analysis of variance to assess which factors significantly affect WTP estimation accuracy. The algorithm for determining prices had no effect on MAPE (WTP) ($p = .347$), but all other factors significantly affected WTP estimation accuracy. Therefore, prices in all choice sets should be individually adapted. Using more choice sets also improves WTP estimation accuracy. However, the difference between 12 and 18 choice sets is small (13.84% vs. 13.44%), so 12 choice sets seems sufficient and reduces the burden on consumers. Regarding the overlap between consumers' WTP and the starting prices, the proposed approach performs best when the starting prices correspond to or are higher than consumers' WTP, though the differences again are small. Prediction accuracy (hit rate) depends on the overlap ($p = .01$), but the other factors do not affect prediction accuracy ($p > .05$).

When we compare the individually adapted design and the traditional design, we find no significant differences in prediction accuracy (88.23% vs. 91.01%, $p = .339$), but the individually adapted design significantly improves WTP estimation accuracy (14.14% vs. 413.95%, $p = .001$). The improvement afforded by the individually adapted design is important because valid WTP estimates are critical inputs for optimization models.

Because the simulated consumers are identical in both simulated data sets, we can compare parameter efficiency directly by looking at the parameters' standard deviations (Louviere et al., 2008). In Table 4, we report the results of a paired-sample t-test comparing standard deviations of the parameter estimates obtained from the individually adapted design and the traditional design. We took three different scenarios into account using nine choice sets for each consumer: (1) 100% overlap of the price and WTP range; (2) 50% above; and (3) 50% below. When there was a significant negative difference between the standard deviations of the parameters obtained from the individually adapted design and the traditional design, the parameters from the individually adapted design became more efficient.

Table 3
Effects of Algorithms and Overlap on WTP Estimation Accuracy and Prediction Accuracy.

	MAPE (WTP) (%)	First choice hit rate (%)
Overlap		
100% Overlap	13.70	83.00
50% Above	14.00	95.33
100% Above	12.00	86.07
50% Below	16.41	76.74
100% Below	14.59	100.00
Number of choice sets		
9 Choice sets	15.13	87.80
12 Choice sets	13.84	88.02
18 Choice sets	13.44	88.87
Number of adapted choice sets		
3 Choice sets	18.22	86.62
6 Choice sets	12.71	89.29
All (9, 12, 18) choice sets	11.49	88.78
Algorithm		
$f(n) = 1 + 1/n$	13.60	87.91
$f(n) = 1 + 2/(n + 1)$	14.31	88.71
$f(n) = 1 + 4/(n + 3)$	14.51	88.07
Mean	14.14	88.23

Notes: Measures are calculated on the basis of the product alternatives in the three holdout choice sets. MAPE = mean absolute percentage error.

Table 4
Simulation study: differences in standard deviations (individually adapted – predefined prices).

		Difference in standard deviations		
		100% Overlap	50% Above	50% Below
Attribute 1	Level 1	**.24**	−**.36**	−**.83**
	Level 2	−**.46**	−**.24**	−**.63**
	Level 3	**.06**	−**.79**	−**1.11**
Attribute 2	Level 1	**.23**	.08	−**.99**
	Level 2	−.02	−**.51**	−**1.13**
	Level 3	.02	−**.26**	−**.89**
Attribute 3	Level 1	**.70**	−**.92**	−**1.06**
	Level 2	**.30**	−**.52**	−**.90**
	Level 3	**.48**	−**.55**	−**1.02**
Attribute 4	Level 1	**.30**	−**1.19**	−**.91**
	Level 2	**.30**	−**1.19**	−**.91**
Price		−**.06**	−**.15**	−**.30**
No-purchase		−**.67**	−**34.56**	−**7.35**

Notes: Paired sample t-test; significant ($p < .05$) differences in standard deviations in **bold** (standard deviation parameter estimate [individually adapted prices] – standard deviation parameter estimate [predefined prices]).

With 100% overlap, the individually adapted design resulted in significantly lower standard deviations for the price and no-purchase parameter estimate, whereas the standard deviation for the other parameter estimates increased slightly (Table 4). With 50% (above or below) overlap, the individually adapted design led to lower standard deviations for all parameter estimates; the standard deviation of the no-purchase parameter decreased in particular. Using an individually adapted design generally improves the efficiency of the parameter estimates.

We next conducted two field studies to demonstrate the usefulness of the individually adapted design in real-world settings and confirm its validity. Although we can observe preferences in real-world settings, we no longer know consumers' actual WTP. Therefore, we depend on reliability, face, convergent, predictive, and external validity to assess the performance of the individually adapted design approach—congruent with the evaluation criteria used in other studies in similar settings (Barrot et al., 2010; Sonnier et al., 2007; Wertenbroch and Skiera, 2002). We also collected data with traditional design (CBC) to compare the results.

5. Field study 1: membership in a supporters club

In our first empirical study, we investigated consumers' WTP for membership in the official supporters club of the German national soccer team. The WTP for this product varies widely, because it is driven by consumers' emotional bond to the team.

5.1. Study design

We conducted an extensive search for potential attributes and attribute levels among existing soccer supporter clubs. We also investigated the fees charged by supporter clubs in other countries to select the appropriate price levels. In a pilot study, we asked 53 existing and potential club members to provide importance ratings for six major product attributes and 59 attribute levels on a seven-point scale. After selecting the four most important attributes, we discussed the results with sport marketing experts from media, marketing agencies, and the national soccer federation, asking them for a realistic range of prices and attribute levels. Managers of the German supporters club believed that, with the offered benefits, supporters' WTP would be approximately 20.00€/year; a pretest among potential members confirmed this assessment. This fee is also comparable to that of other soccer supporter clubs (Club

Table 5
Attributes and attribute levels: field study 1.

Attribute	Level of attribute
Starter package	• Shirt + poster • Cap + scarf • Highlight DVD + key ring + foam hand • Calendar + flag
Multimedia	• Game highlight reports on the Internet • Pre- and after-game reports on the Internet • Mobile phone downloads • Headlines on the mobile phone
Tickets/travels	• Right of first refusal for tickets • Extra ticket contingent for World Cup/European Cup • Budget travel packages for away games • Loyalty program: Rewards for frequent game attendants
Miscellaneous	• Magazine • 2 Free tickets for a friendly match of a youth team • 10% Fiscount on all products in the fan shop • "Fan-tastic" moments (tickets to unique events)
Price	• 15€/year • 20€/year • 25€/year (For CBC and starting values for IACBC)

Oranje and the Scottish supporters club both charge 25€/year). Table 5 lists the final attributes and their levels.

The design consisted of 12 choice sets, from which we estimated the parameters, and 3 holdout sets to assess predictive validity. Each choice set contained three membership alternatives and a no-purchase option. The holdout sets were the same for CBC and IACBC. For IACBC, we individually adapted the prices in all 12 choice sets based on an oscillation amplitude of $f(n) = 1 + 2/(n + 1)$. Every consumer also stated his or her WTP directly for a specific product (base product), according to an open-ended contingent valuation question (Wertenbroch and Skiera, 2002). We used this directly stated WTP for the base product to assess convergent validity.

We tested this design in a pilot study with 25 existing and potential members. We then placed a link to the online questionnaires on the website of the supporters' club and randomly assigned participants to either the CBC or IACBC design. This procedure ensured that consumers were generally interested in the product. We obtained 490 IACBC and 513 CBC completed surveys. Again, we employed a latent class MNL model to estimate the parameters. To determine the number of latent classes, we used the Bayesian information criterion (BIC) (Andrews et al., 2002).

5.2. Results

In Table 6, we provide the descriptive statistics for the sample, the WTP estimates, and the measures to assess the reliability and face, convergent, and predictive validity of the WTP estimates. We do not report the estimated parameters because they are not of primary interest; rather we aim to assess the ability of the different approaches to elicit consumers' WTP for membership and thus consumers' demand for this product.

In the IACBC results, the best model (according to the BIC) contains five segments. We used each consumer's segment membership probability to derive individual parameter estimates, which then enabled us to calculate each consumer's WTP. To assess whether the WTP estimates derived from IACBC are reliable, we split the sample in half and compared the WTP estimates. They were not significantly different (Δ = 1.98€, p = .648). The estimated mean WTP for the base product was 60.90€, and we found a minimum WTP of .00€ (maximum = 131.73€). The average WTP across all product combinations was 24.86€, slightly higher than the current yearly membership fee of 20.00€ but certainly plausible. For 73.88% of consumers, the estimated average WTP across all attribute combinations fell within the considered price range (Table 6). Thus, the WTP estimates demonstrate face validity. To assess the convergent validity, we computed the correlation between the directly stated and estimated WTP for the base product; the correlation was relatively high (.33). Finally, we evaluated predictive validity by the first choice hit rate in a validation sample. To calculate the hit rate, we used a validation sample of 50 randomly selected consumers who were not considered when estimating the parameters, along with the three holdout sets. We derived the individual parameter estimates by computing consumers' probability of class membership, according to their response behavior. The first choice hit rate of 56.46% is good and much higher than the 25% chance criterion.

Consumers might react to the individually adapted prices by "playing around" with them. To address this concern, we looked at the number of no-purchase decisions per consumer. If consumers are playing around with the individually adapted prices, they should select the no-purchase option relatively frequently (testing lower bound) or infrequently (testing upper bound). Only 5% of the consumers selected the no-purchase option more than eight times, and only 3% did so less than four times; therefore, it appears that consumers respond adequately.

Table 6
Field study 1: sample description, WTP estimates, and validity measures.

	IACBC (N = 490)	CBC I (N = 513)	CBC II (N = 213)
Number of latent classes	5	4	3
Number of parameters that needed to be estimated	74	59	44
Share of never selecting no-purchase option (%)	0.00	58.48	0.00
Share of always selecting no-purchase option (%)	0.00	0.00	0.00
Share of no-purchase decisions (%)	46.92	12.69	30.56
Split-half reliability (€)[a]	1.98	103.64	31.99
(p-value)	(0.648)	(0.043)	(0.001)
Mean WTP for base product (€)	60.90	1,439.30	50.43
Median WTP for base product (€)	44.57	633.95	65.50
Minimum WTP for base product (€)	0.00	14.7525,6	3.80
Maximum WTP for base product (€)	131.73	17.91	71.99
Mean average WTP across attribute combinations (€)	24.86	565.89	16.10
Share of consumers where average WTP is within price range (%) [b]	73.88	0.20	0.94
Correlation estimated and directly stated WTP	0.33	0.15	0.21
Predictive validity (validation sample)			
First choice hit rate (%)	56.46	58.00	60.00
Log-likelihood value	−149.88	−143.25	−150.62

Notes: The measures were computed on the basis of individual parameters.
[a] Difference in means for the estimated WTP in the two independent samples.
[b] Based on prices shown in choice sets 2–12.

For the CBC data set, we computed the share of consumers who always or never selected the no-purchase option and found that 58.48% never chose the no-purchase option, whereas 0% always chose it (Table 6). The share of consumers who never selected the no-purchase option seems quite high, which implies that consumers' WTP was higher than the considered price levels. The directly stated WTP for the membership fee varies between 0.00€ and 200.00€, illustrating both the heterogeneity in consumers' WTP and the difficulty of capturing the entire WTP range with predefined price levels. Although the share of consumers who never selected the no-purchase option seems quite high, we believe that such response behavior is not unusual. For example, Parker and Schrift (2011) and Sonnier et al. (2007) find that 63.55% and 49.7% of consumers, respectively, never selected the no-purchase option. To eliminate possible effects of extreme response behavior, we also analyzed a subset of the CBC sample that contains only consumers who select the no-purchase option at least once (CBC II, $N = 213$).

For CBC I and CBC II, the WTP estimates are significantly different when we test split-half reliability ($\Delta = 103.64$, $p = .043$; $\Delta = 31.99$, $p = .001$). For CBC I, mean WTP for the base product was 1439.30€ per year, which translates into a monthly membership fee of almost 120€. Perhaps only one segment drives this result. But for the four segments, we also derived the following WTP values for the base product: 780.88€ (segment 1), 3663.39€ (segment 2), 16.46€ (segment 3), and 85.64€ (segment 4). The WTPs in Segments 1 and 2 are unreasonably high. It is also important to note that the consumers who showed extreme response behavior were not assigned to any specific segment (52% in segment 1, 33% in segment 2, and 15% in segment 4). The mean for the average WTP across all attribute combinations was also very high (565.89€) and showed no face validity.

For CBC II, the mean WTP for the base product equaled 50.43€ (62.00€ in segment 1, 71.67€ in segment 2, and 4.56€ in segment 3). The WTP estimates for the base product offer greater face validity than those for CBC I; furthermore, the average WTP of 16.10€ across all attribute combinations seems plausible (Table 6). Yet, the consumers who never selected the no-purchase option are no longer considered, so we gain no information about their WTP, even though these "hot prospects" remain an interesting target group for the club. For only .94% of the consumers does the estimated average WTP lie within the price range. Thus an extrapolation would be necessary, which is prone to error and should be avoided (Orme, 2002).

The correlation between the directly stated and estimated WTP was low for CBC I and CBC II (.15 and .21), and significantly lower than for IACBC. The predictive validity of CBC I and CBC II were slightly higher than for IACBC, but the differences were not substantial. Overall, the hit rates were comparable with those from similar field studies (e.g., Iyengar et al., 2008; Jedidi and Zhang, 2002; Schlereth and Skiera, forthcoming). These results support the predictive validity of CBC, though as we already noted, hit rates are not useful when researchers' interest is in estimating consumers' WTP.

5.3. Summary

We determined the price levels for this study using expert and consumer interviews, together with a pretest. Nevertheless, the predefined price levels did not capture the entire range of consumers' WTP, so we observed a large share of consumers showing extreme response behavior when we used CBC: 58.48% of consumers never selected the no-purchase option, indicating that price levels were too low and that problems with the validity of WTP estimates exists. Although consumers showing extreme response behavior by always selecting an alternative are an interesting target group, no plausible WTP estimates can be derived for this group with CBC. In contrast, IACBC captures the entire range of consumers' WTP, resulting in plausible WTP estimates for all consumers. Moreover, IACBC allows us to make demand predictions for membership fees varying between 3.00€ and 200.00€ (minimum and maximum price in choices sets across consumers)—not just between 15.00€ and 25.00 €. Such information may prove valuable as input for optimization models in operations research, if the goal is to consider behavioral aspects.

6. Field study 2: digital video recorder

In our second empirical study, we investigated consumers' WTP for digital video recorders. In this study, we can also assess the external validity of the suggested IACBC.

6.1. Study design

We investigated consumers' WTP for a digital video recorder (DVR). We searched for relevant attributes and attribute levels among potential DVR buyers, reviewed current offers, and interviewed potential buyers to determine price levels (for a similar approach, see Jedidi et al., 2003). Table 7 lists the relevant attributes and their levels.

We cooperated with a market research agency to obtain a representative sample. The study design consisted of 12 choice sets and 3 holdout sets to assess predictive validity. Each choice set contained three different DVRs and a no-purchase option. Every consumer also stated his or her WTP directly for a specific product (base product), according to an open-ended contingent valuation question. We used the directly stated WTP as a measure of convergent validity. We obtained 263 completed surveys for the traditional design and 163 for the individually adapted design.[3] We applied a latent class MNL model to estimate the parameters and BIC to determine the number of latent classes.

6.2. Results

When guided by predefined price levels (CBC), 30.67% of the consumers never selected the no-purchase option, and 15.13% always did (Table 8). The share of consumers who exhibited extreme response behavior is thus comparable to percentages from recently published studies (e.g., Gilbride et al., 2008; Natter et al., 2008). In Table 8, we also provide the descriptive statistics regarding the WTP estimates and measures of face, convergent, and predictive validity. For the individually adapted design (IACBC), we estimated a mean WTP for the base product of 7.99€ (maximum = 23.99€), which seems plausible. The mean average WTP across all attribute combinations was 6.43€, which also seems plausible. The correlation between the estimated and directly stated WTP also was high (.42). To assess predictive validity, we again considered the first choice hit rate in a validation sample, according to the three holdout choice sets and a selection of 50 consumers not used to calibrate the model. The first choice hit rate of 58.50% was much higher than the 25% chance criterion.

When using predefined price levels (CBC), the mean WTP for the base product was 10.38€, but the maximum WTP of 83.90€ appears quite high. The mean average WTP across all attribute combinations was 11.55€, higher than that for IACBC. The .10 correlation between estimated and directly stated WTP also was significantly lower than that for IACBC. These results suggest a higher face and convergent validity for the WTP estimates derived from IACBC.

[3] We collected more responses for the traditional design because we wanted to analyze the effects of extreme response behavior in more detail.

Table 7
Attributes and attribute levels: field study 2.

Attribute	Level of attribute
Possibility to archive recordings	• No possibility to archive recordings • Connection to integrated DVD burner • Connection to home network • External hard drive
Access to online video rental service	• Yes • No
Recommendation system	• No recommendation system available • Recommendation system based on experts • Recommendation system based on other users' behavior • Recommendation system based on own behavior
Additional TV channels	• Movies and documentaries • TV series and comics • News, erotic and sport
Price	• 4.95€/month • 9.95€/month • 14.95€/month (For CBC and starting values for IACBC)

Table 8
Field Study 2: Sample Description, WTP Estimates, and Validity Measures.

	IACBC (N = 163)	CBC (N = 263)
Number of latent classes	5	8
Number of parameters that needed to be estimated	59	95
Share of never selecting no-purchase option (%)	0.00	30.67
Share of always selecting no-purchase option (%)	0.00	15.13
Share of no-purchase decisions (%)	50.97	37.11
Mean WTP for base product (€)	7.99	10.38
Median WTP for base product (€)	8.23	5.55
Minimum WTP for base product (€)	0.00	0.00
Maximum WTP for base product (€)	23.99	83.90
Mean average WTP across all attribute combinations (€)	6.43	11.55
Correlation estimated and directly stated WTP	0.42	0.10
Share of consumers where average WTP within price range (%) [a]	72.39	18.49
Model validity (validation sample)		
First choice hit rate (%)	58.50	61.90
Log-likelihood	−172.62	−147.13

Notes: Measures computed on the basis of individual parameters.
[a] Based on prices shown in choice sets 2–12.

Only with respect to predictive validity does the traditional design perform slightly better—an unsurprising result considering the substantial share of consumers who selected the no-purchase option often (which also applies to the validation sample). Therefore, we can predict their choices with high accuracy, but we gain little information about their preferences for certain attributes or WTP.

To determine the external validity of the two approaches, we compared predicted and actual market penetration. Information about the actual market penetration of DVRs comes from a report by TNS Infratest (2008). The consumer panel data forecast that market penetration of DVRs would reach 16.7% in 2009. The existing DVR offers are comparable in their attributes and levels (i.e., integrated DVD burner, no access to online video rental, no recommendation system, additional televised programs), and the average price is about 14.95€. With these specifications as a base, we predicted the market penetration. With individually adapted prices, we predict a market penetration of 17.9%, whereas the traditional approach (using predefined prices) estimates a market penetration of 34.6%. Thus, the individually adapted prices resulted in very high external validity.

6.3. Summary

In this second field study, we determined the price levels using consumer interviews and existing market prices. Nevertheless, the predefined price levels did not capture the entire range of consumers' WTP, and we thus observed extreme response behavior in the CBC. In particular, 30.67% of the consumers never selected the no-purchase option, and 15.13% always selected the no-purchase option. The predefined price levels in the CBC study do not capture consumers' entire WTP range, which already creates problems for deriving valid WTP estimates. However, IACBC can capture that range and offer plausible WTP estimates for all consumers. Moreover, IACBC allows us to make demand predictions for a large range of prices (between .05 EUR/month and 46.05 EUR/month) and exhibits high external validity. Optimization models in operations research thus may benefit from this enhanced knowledge.

7. Summary and conclusions

Operations management models increasingly employ behavioral aspects, such as consumers' preferences and WTP, to identify profit-maximizing product designs, product lines, and prices. This development has led to an emerging reliance on CBC analysis in operations research. However, such research can benefit from including consumers' preferences and WTP into models only if they are accurately estimated. Otherwise, the "optimal" solutions produced will be less than optimal in truth.

The validity of the estimates of consumers' preferences and WTP rely heavily on consumers' response behavior in a CBC study; response behavior is affected by the design of the CBC study, especially the price levels provided by the study. If the prices do not sufficiently overlap with consumers' WTP, extreme response behavior occurs: Consumers always or never choose the no-purchase option. We show in a simulation study and two field studies that both types of extreme response behavior result in invalid WTP estimates, even if only a small share of consumers exhibit extreme response behavior. Accordingly, CBC studies should start reporting these two shares regularly.

We suggest an adaptive approach, IACBC, that substantially improves the elicitation of valuable information about consumers' WTP. The proposed individually adapted design significantly improves WTP estimation accuracy and increases the input quality for optimization models in operations research. The method has particular relevance for products for which no strong reference price exists and heterogeneity across consumers' WTP is likely. In those cases, the predefined price levels of nonadaptive designs likely will not cover the entire range of WTP, leaving a significant knowledge gap about the maximum or minimum price that consumers would be willing to pay.

The proposed approach offers other considerable benefits too. Firms are often interested in not only accurate WTP estimates but also demand predictions based on market simulations. With individualized price levels, we can predict demand for a continuous and large price range, which offers more detailed insights into potential price thresholds, which in turn can pave the way for a sophisticated market segmentation strategy. Furthermore, the proposed approach has greater external validity than traditional designs. Finally, the proposed approach neither induces significant changes to existing practice nor places an extreme burden on consumers, which makes it easy to implement and use.

However, our proposed approach suffers some limitations that further research might address. For example, prices sometimes feature a lower limit (e.g., marginal costs of a product), so consumers with a WTP lower than this limit are of no interest to the firm. We did not implement lower price limits, but considering them would be relatively easy; additional research could explore explicitly whether implementing price limits is worthwhile. We also adapted prices individually, but consumers' preferences for certain (metric) attributes could be quite heterogeneous, which makes it difficult to determine the levels a priori. Further research therefore should investigate whether adapting (metric) attributes other than price might improve the parameter and WTP estimates. However, in this effort, researchers must keep in mind that many attributes feature a relatively small range of technologically possible levels, and offering a larger range of attribute levels increases operational complexity and thus production costs. Careful consideration of the potential benefits of adapting other attributes thus is required.

Acknowledgements

We thank Rosanna Garcia, Robert Rooderkerk and Thomas Otter for providing valuable feedback that greatly strengthened the manuscript. We also thank Garret Sonnier, Thomas Otter, and Martin Natter for providing their CBC data sets.

References

Ahuja, R., Orlin, J., Sechi, G., Zuddas, P., 1999. Algorithms for the simple equal flow problem. Management Science 45 (10), 1440–1455.
Albers, S., Brockhoff, K., 1977. A procedure for new product positioning in an attribute space. European Journal of Operations Research 1 (4), 230–238.
Andrews, R., Ainslie, A., Currim, I., 2002. An empirical comparison of logit models with discrete versus continuous representations of heterogeneity. Journal of Marketing Research 39 (4), 479–487.
Arora, N., Huber, J., 2001. Improving parameter estimates and model prediction by aggregate customization in choice experiments. Journal of Consumer Research 28 (September), 273–283.
Barrot, C., Albers, S., Skiera, B., Schäfers, B., 2010. Vickrey vs. eBay: Why second-price sealed-bid auctions lead to more realistic price-demand functions. International Journal of Electronic Commerce 14 (4), 7–38.
Bendoly, E., Donohue, K., Schultz, K., 2006. Behavior in operations management: Assessing recent findings and revisiting old assumptions. Journal of Operations Management 24 (6), 737–752.
Brazell, J., Diener, C., Karniouchina, E., Moore, W., Séverin, V., Uldry, P.-F., 2006. The no-choice option and dual response choice designs. Marketing Letters 17 (4), 255–268.
Chen, K., Hausman, W., 2000. Technical note: Mathematical properties of the optimal product line selection problem using choice-based conjoint analysis. Management Science 46 (2), 327–332.
Coltman, T., Gattorna, J., Whiting, S., 2010. Realigning service operations strategy at DHL Express. Interfaces 40 (3), 175–183.
Day, J., Venkataramanan, M.A., 2006. Profitability in product line pricing and composition with manufacturing commonalities. European Journal of Operational Research 175 (3), 1782–1797.
Dhar, R., 1997. Consumer preference for the no-choice option. Journal of Consumer Research 24, 215–231.
Easton, F., Pullman, M., 2001. Optimizing service attributes: The seller's utility problem. Decision Sciences 32 (2), 251–275.
Eggers, F., Sattler, H., 2009. Hybrid individualized two-level choice-based conjoint (HIT-CBC): A new method for measuring preference structures with many attribute levels. International Journal of Research in Marketing 26 (2), 108–118.
Evgeniou, T., Pontil, M., Toubia, O., 2007. A convex optimization approach to modeling consumer heterogeneity in conjoint estimation. Marketing Science 26, 805–818.
Gilbride, T., Guiltinan, J., Urbany, J., 2008. Framing effects in mixed price bundling. Marketing Letters 19 (2), 125–139.
Green, P.E., Krieger, A.M., 1992. Modeling competitive pricing and market share: Anatomy of a decision support system. European Journal of Operational Research 60 (1), 31–44.
Gunasti, K., Ross, W., 2009. How inferences about missing attributes decrease the tendency to defer choice and increase purchase probability. Journal of Consumer Research 35, 823–839.
Haaijer, R., Kamakura, W., Wedel, M., 2001. The 'no-choice' alternative in conjoint choice experiments. International Journal of Market Research 43 (1), 93–106.
Halme, M., Kallio, M., 2011. Estimation methods for choice-based conjoint analysis of consumer preferences. European Journal of Operational Research 214 (1), 160–167.
Hauser, J., Toubia, O., 2005. The impact of utility balance and endogeneity in conjoint analysis. Marketing Science 24 (3), 498–507.
Huber, J., Zwerina, K., 1996. The importance of utility balance in efficient choice designs. Journal of Marketing Research 33 (3), 307–317.
Iyengar, R., Jedidi, K., Kohli, R., 2008. A conjoint approach to multi-part pricing. Journal of Marketing Research 45 (2), 195–210.
Jedidi, K., Jagpal, S., Manchanda, P., 2003. Measuring heterogeneous reservation prices for product bundles. Marketing Science 22 (1), 107–130.
Jedidi, K., Zhang, J., 2002. Augmenting conjoint analysis to estimate consumer reservation price. Management Science 48 (10), 1350–1368.
Kanninen, B., 2002. Optimal design for multinomial choice experiments. Journal of Marketing Research 39 (2), 214–227.
Karniouchina, E., Moore, W.L., van der Rhee, B., Verma, R., 2009. Issues in the use of ratings-based versus choice-based conjoint analysis in operations management research. European Journal of Operational Research 197 (1), 340–348.
Kessels, R., Goos, P., Vandebroek, M., 2006. A comparison of criteria to design efficient choice experiments. Journal of Marketing Research 43 (3), 409–419.
Kohli, R., Krishnamurti, R., 1989. Optimal product design using conjoint analysis: Computational complexity and algorithms. European Journal of Operational Research 40 (2), 186–195.
Kraus, U., Yano, C.A., 2003. Product line selection and pricing under a share-of-surplus choice model. European Journal of Operational Research 150 (3), 653–671.
Kuhfeld, W., Tobias, R., Garratt, M., 1994. Efficient experimental design with marketing applications. Journal of Marketing Research 31 (4), 545–557.
Liu, Q., Otter, T., Allenby, G., 2007. Investigating endogeneity bias in marketing. Marketing Science 26 (5), 642–650.
Louviere, J., Islam, T., Wasi, T., Street, D., Burgess, L., 2008. Designing discrete choice experiments: Do optimal designs come at a price? Journal of Consumer Research 35 (2), 360–375.
Louviere, J., Woodworth, G., 1983. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. Journal of Marketing Research 20, 350–367.
Meijer, E., Rouwendal, J., 2006. Measuring welfare effects in models with random coefficients. Journal of Applied Econometrics 21 (2), 227–244.
Moorthy, S., Ratchford, B., Talukdar, D., 1997. Consumer information search revisited: Theory and empirical analysis. Journal of Consumer Research 23 (4), 263–277.
Natter, M., Feurstein, M., 2002. Real world performance of choice-based conjoint models. European Journal of Operational Research 137 (2), 448–458.
Natter, M., Mild, A., Wagner, U., Taudes, A., 2008. Planning new tariffs at tele.ring—The application and impact of an integrated segmentation, targeting and positioning tool. Marketing Science 27 (4), 600–609.
Orme, B., 2002. Formulating attributes and levels in conjoint analysis. Sawtooth Software Research Paper Series. <www.sawtoothsoftware.com/download/techpap/formatt.pdf> (last accessed 11.10.10).
Rusmevichientong, P., Shen, Z.-J., Shmoys, D., 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. Operations Research 58 (6), 1666–1680.
Park, Y.H., Ding, M., Rao, V., 2008. Eliciting preference for complex products: A web-based upgrading method. Journal of Marketing Research 45 (5), 562–574.
Parker, J., Schrift, R., 2011. Rejectable choice sets: How seemingly irrelevant no-choice options affect consumer decision processes. Journal of Marketing Research 48 (5), 840–854.
Sandor, Z., Wedel, M., 2001. Designing conjoint choice experiments using managers' prior beliefs. Journal of Marketing Research 38 (4), 430–444.
Schlereth, C., Skiera, B., forthcoming. Measurement of consumer preferences for bucket pricing plans with different service attributes. International Journal of Research in Marketing.
Schlereth, C., Stepanchuk, T., Skiera, B., 2010. Optimization and analysis of the profitability of tariff structures with two-part tariffs. European Journal of Operational Research 206 (3), 691–701.
Scholl, A., Manthey, L., Helm, R., Steiner, M., 2005. Solving multiattribute design problems with analytic hierarchy process and conjoint analysis: An empirical comparison. European Journal of Operational Research 164 (3), 760–777.
Sonnier, G., Ainslie, A., Otter, T., 2007. Heterogeneity distributions of willingness-to-pay in choice models. Quantitative Marketing and Economics 5 (3), 313–331.
Tarasewich, P., McMullen, P.R., 2001. A pruning heuristic for use with multisource product design. European Journal of Operational Research 128 (1), 58–73.
TNS Infratest, 2008. Personal Video Recorder – Besitz im Trend (Personal Video Recorder – Trends). <http://www.tns-infratest.com/presse/pdf/Presse/20080528_TNS_Infratest_SES_PVRcharts.pdf> (last accessed 05.08.11).
Toubia, O., Hauser, J., 2007. On managerially efficient experimental designs. Marketing Science 26 (6), 851–858.
Toubia, O., Hauser, J., Garcia, R., 2007. Probabilistic polyhedral methods for adaptive choice-based conjoint analysis: Theory and application. Marketing Science 26 (5), 596–613.
Verma, R., Thompson, G.M., Moore, W.L., Louviere, J.J., 2001. Effective design of products/services: An approach based on integration of marketing and operations management decisions. Decision Sciences 32 (1), 165–194.
Vermeulen, B., Goos, P., Vandebroek, M., 2008. Models and optimal designs for conjoint choice experiments including a no-choice option. International Journal of Research in Marketing 25 (2), 94–103.
Wathieu, L., Bertini, M., 2007. Price as a stimulus to think: The case for willful overpricing. Marketing Science 26 (1), 118–129.
Wedel, M., Kamakura, W., 2000. Market Segmentation. Conceptual and Methodological Foundations. Kluwer Academic Publishers, Boston.
Wertenbroch, K., Skiera, B., 2002. Measuring consumers' willingness to pay at the point of purchase. Journal of Marketing Research 39 (2), 228–241.