WORKING PAPER SERIES

Toward an Understanding of the BDM: Predictive Validity, Gambling Effects, and Risk Attitude.

Sebastian Lehmann

Working Paper No. 1/2015

OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

FACULTY OF ECONOMICS
AND MANAGEMENT

Impressum (§ 5 TMG)

Herausgeber:
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Der Dekan

Verantwortlich für diese Ausgabe:
Sebastian Lehmann
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Postfach 4120
39016 Magdeburg
Germany

http://www.fww.ovgu.de/femm

Bezug über den Herausgeber
ISSN 1615-4274

Toward an Understanding of the BDM: Predictive Validity, Gambling Effects, and Risk Attitude.

Sebastian Lehmann

31.12.2014

Acknowledgements: I would like to thank Andreas Hahn, Thomas Hochmuth, Jessica Lawniczak, Janina D. Lettow, Marcel Lichters, Doreen Neubert, Daniel Werner, Stanislaw Werner, Tobias Winter, and David Saalfeld for supporting the experiment.

Abstract

Pricing research suggests incentive-compatible evaluations when consumers' situation-specific WTP is to be elicited. Especially, the lottery-based Becker-DeGroot-Marschak-mechanism (BDM) is recommended, as it seems to outperform other elicitation methods. In this study, the BDM was used to measure subjects' WTP for eight shopping goods in binding purchase settings. In accordance with previous studies, the validity of elicited WTP measures was checked within subjects with respect to indicators of face and criterion validity (such as interest in buying, preference ratings, and compliance rates). In addition, this study observed real purchases of a separate validation sample measured under identical circumstances, thus assessing the predictive validity of WTPs elicited with the BDM. As a result, the BDM-based WTPs reveal a sufficient degree of internal face and criterion validity. However, the external validity in terms of predictive validity between WTP-based prediction and purchases of the validation sample seems limited. Specifically, this study found a substantial overestimation of WTP, and thus in the corresponding purchase rates in the BDM. Hence, a potential bias is indicated. However, contrary to the suggestions of earlier research, individual risk attitude or loss aversion, hence a potential gambling effect, seems not to bias BDM results or the decision whether to buy or not.

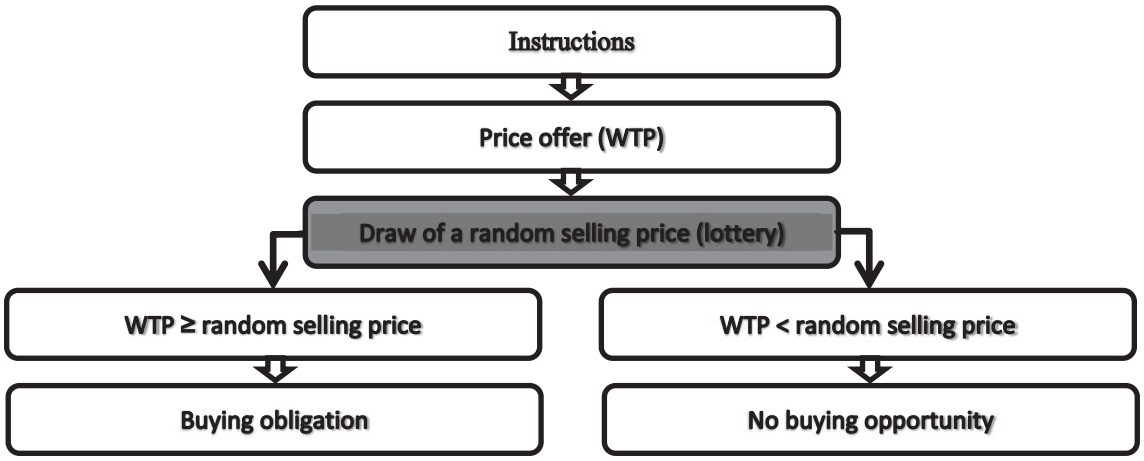
Keywords: BDM, price research, WTP, gambling effect, risk attitude

1. Introduction

The influence of the pricing policy on the bottom line of any company cannot be overestimated. The price determines whether a potential customer buys a product or chooses a competitive offer. In this way the prevailing market price has doubled impact on profits; it also affects variable production costs as well as the revenue of any company (Diller, 2007). Hence, price constitutes one of the most effective drivers of profits (Simon, 2004). Therefore, designing an optimal pricing policy is a major challenge within a company's marketing strategy. The pricing policy is necessarily based on the managers' understanding of consumers' preferences and demand behavior as these determine perceived value; that is, utility (Marshall, 1920) of a market offer for a customer. The perceived utility is the fundamental building block of customers' willingness to pay (hereafter WTP; Kalish & Nelson, 1991), and thus the basis for profit-optimizing prices. In order to determine this elementary component and thus to infer optimal decisions, market researchers need valid and feasible elicitation methods. With regard to this issue, academic literature has proposed several different approaches to measure consumers' WTP as a surrogate of perceived utility (for further insights into measurement methods of WTP see Le Gall-Ely (2009) and Völckner (2006b). The most promising approach seems to be the incentive-compatible procedures that emerged from experimental economics (Hoffman, Menkhaus, Chakravarti, Field, & Whipple, 1993; Miller, Hofstetter, Krohmer, & Zhang, 2011; Völckner, 2006a; Wertenbroch & Skiera, 2002). These methods offer some appealing properties since they combine the theoretical advantage of real transaction data and the operational advantage of survey data. In particular, these procedures are theoretically assumed to provide consumers with the incentive to truthfully reveal their WTP, as the subjects are put in a real purchase situation in which they are asked to make a binding purchase offer for the relevant product, without being able to actually directly influence the final selling price with their offer (Shogren et al., 2001; Wertenbroch & Skiera, 2002, p. 230). One of the most widely used methods is the Becker-DeGroot-Marschak mechanism (hereafter BDM, Becker, DeGroot, and Marschak, 1964). Here participants are asked to submit their WTP for a given good. The actual sales price is then randomly drawn from an (unknown) distribution of prices. Participants are required to purchase the good at a price below or equal to their bids, but they cannot buy the good at a price above their bids (Becker et al., 1964)[1]. For a graphical representation see figure 1.

[1] In the original research design of Becker, DeGroot, and Marschak a minimum selling price from the seller instead of a maximum buying price from the buyer was elicited; however, the latter is the common approach in recent price research, see e.g., Müller, Kroll, & Vogt, 2012a; Völckner, 2006a; Wertenbroch & Skiera, 2002.

Figure 1: Classical BDM procedure



It has been shown that this procedure is (theoretical) incentive compatible (Irwin, McClelland, McKee, Schulze, & Norden, 1998), but it seems unclear whether the BDM is also behaviorally incentive compatible (Kaas & Ruprecht, 2006), especially if distortions of expected utility theory arise, which is the theoretical foundation of the incentive compatibility of the BDM procedure (Horowitz, 2006).

When looking at the existing literature on price research, it seems the major concern in the evaluation of WTP elicitation methods is on incentive compatibility (see e.g., Ding, Grewal, and Liechty, 2005; Horowitz, 2006; Irwin et al., 1998; Kaas and Ruprecht, 2006; Lusk, Alexander, and Rousu, 2007; McAfee and McMillan, 1987; Miller et al., 2011; Rutström, 1998; Völckner, 2006a; Wang, Venkatesh, and Chatterjee, 2007; Wertenbroch and Skiera, 2002). However, if (price) research results are to be transferred beyond the specific research setting, their external validity must not be neglected (Ehrenberg, 1993; Hubbard & Armstrong, 1994; Winer, 1999). Thus far the external validity was assessed several times (see e.g., Horowitz, 2006; Völckner, 2006a; Wertenbroch, and Skiera, 2002) and compared to other (incentive compatible) WTP elicitation methods (see e.g., Miller et al. 2011; Müller, Voigt, and Erichson, 2010; Noussair, Robin, & Ruffieux, 2004). Even so, BDM seems to perform best in most of the comparisons among the methods tested (Völckner, 2006a; Völckner, 2006b; Miller et al., 2011) Yet, it still tends to over- (Müller and Voigt, 2010) or underestimate (Kaas Ruprecht, 2006) true WTP elicited at the POS in some occasions. In this context predictive validity as important building block of external validity plays a crucial role. Especially, as some sources question the predictive validity of BDM results. For example, Ding et al. (2005) found weaknesses in the congruence between the predicted choice and the actual choice. The BDM prediction was correct in only 15% of the cases. Yet, no compelling empirical explanation of this incongruence on the basis of personal characteristics and/or values can be found. Moreover, existing research testing the predictive validity of BDM results focuses at the comparison of lab versus field based results. Given the high context and situation specificity of WTP (Bateman, Munro, Rhodes, Starmer, & Sugden, 1997; Bettman, Luce, & Payne, 1998; Huber, Payne, & Puto, 1982; Kahneman & Tversky, 1979; Thaler, 1985) this is very likely to produce significant differences *per se*. An evaluation of whether the BDM

procedure as such might be a source of biase as it interacts with personal traits, is hard to tell, if many facorts are altered at the same time, hence in a field setting.

This research sets out to test the predictive validity of the BDM procedure and thus assesses the external validity of results with the aim of generalizing them to situations perceived to match the point of sale (hereafter POS) situation closer; that is, being perceived as more realistic than the BDM. Specifically, this study compares the predicted purchase rates (hereafter PR) from the BDM for eight shopping goods to the real PR of a matched validation sample in which the participants faced a dichotomous choice task (hereafter DCT) on whether to buy or not to buy at the presented sales prices[2]; that is, the open-ended elicitation of the willingness to buy (hereafter WTB). This benchmark is said to predict real purchase behavior better because of its closer resemblance to a real purchase situation (Müller et al., 2010; Völckner, 2006b) and its cognitive simplicity (Mitchell & Carson, 1989). To account for the high context and situation specificity of WTP/WTB an identical setting for both groups was created to equalize as many relevant background factors (Lynch, 1982; 1999) as possible. In fact, this study measured the WTP in the BDM condition and the PR of the DCT (partially) at the same time in the same laboratory. So the aim of the DCT was not to elicit true WTB, but to validate the predicted PR from the BDM, as this elicitation procedure is closer to a real purchase decision at the POS (Müller et al., 2010), and the real purchase situation is truly what the BDM is supposed to predict. If the BDM is a good predictor for WTB, one would expect small or no differences between the predicted and real PR.

As this study finds a significant difference between the PR predicted by the BDM and the real PR from the DCT, it investigates possible explanations as being the risk attitude (de Meza & Reyniers, 2013; Kaas & Ruprecht, 2006) of participants and the gambling effect (Völckner, 2006a).

This paper proceeds as follows. First, it sheds light on WTP elicitation methods with a special focus at the BDM. Then it highlights drawbacks of the procedure and potential sources of bias. After the presentation of experimental design aspects to prevent known sources of bias, this paper finally discusses its findings and limitations.

2. Aspects of Validity in WTP Elicitation Methods

If there is adequate data reliability, the quality of WTP estimates depends on their validity. Hence, in order to improve the BDM procedure, it is crucial to identify potential sources of systematic bias and analyze how to reduce bias. While external validity is significant when validating results for pricing purposes, research on the validity of WTP elicitation methods is rather limited (see Ding et al., 2005; Miller et al., 2011; Müller et al., 2010). Examining whether a method provides valid estimates of WTP is a complex task, since the WTP is a latent, and thus an unobservable, construct (Völckner, 2006a).

[2] This is a prices point chosen from the price density function of the BDM.

Moreover, it is not possible to prove external validity: It can only be assessed through other means (see the dialogue between Calder, Phillips, and Tybout (1981; 1982) and Lynch (1982; 1983; 1999). Therefore, different conceptions and means of validity have been used to test the validity of research results. Specifically, researchers most often use the concepts of face and criterion validity to check whether their results are meaningful. However, testing whether results are generalizable requires further consideration (Hubbard & Armstrong, 1994). Comparative studies that examine the extent to which results from different procedures are consistent, and therefore test convergent validity, represent a step toward generalizability[3] because they also test the predictive validity of results. However, a replication (with extension) of the original study is a true test of external validity. This can be accomplished by a careful, theory-based variation of research components (for example, the research setting or the sample composition) to test whether earlier results hold under different situations that go beyond the original research frame (robustness), or might even be generalizable (Hubbard & Armstrong, 1994; Lynch, 1999).

Literature on WTP elicitation methods provides some general aspects to keep in mind during price research. Usually, methods that elicit WTP in a binding setting, that is a real purchase situation, are assumed to provide higher validity than those that elicit hypothetical WTP. The main concern is that hypothetical methods might lead to systematically biased estimates, as participants apply less effort to their calculation and thus define their WTP inaccurately, because their responses do not have real economic consequences (Ding et al., 2005). Several empirical studies give evidence for the existence of such a hypothetical bias (Miller et al., 2011; Völckner, 2006a; Wertenbroch & Skiera, 2002). In general, elicitations in hypothetical settings seem to generate significantly higher valuations than their counterparts in real settings, across and within different methods (List & Gallet, 2001; Murphy, Allen, Stevens, & Weatherhead, 2005). However, this is not to say that hypothetical approaches are of no value in guiding pricing decisions as Miller et al. notes (2011).

In addition, there might be a potential strategic bias when direct WTP elicitation methods are used. Participants might believe their responses influence future outcomes, such as the likelihood of a new product's market entry, or future market prices (Wertenbroch & Skiera, 2002). Thus, participants in both hypothetical and binding settings[4] might have an incentive to strategically overstate or understate their true WTP in order to maximize their individual utility (Carson & Groves, 2007). In contrast to direct methods, indirect methods draw people's attention away from the pricing decision and therefore mitigate strategic response behavior (Wang et al., 2007, p. 203).

[3] However, a comparison of results from different incentive-compatible methods with each other provides only limited information in the event the results differ. In such a case, at best, only one of the methods empirically provides incentive compatibility and thus results that may be generalizable, but one cannot determine which one, if any, it is (Kaas & Ruprecht, 2006).

[4] In binding settings, however, strategic responses seem to be less likely, as participants have to consider the cost of deviating from their true WTP that occur either as a forgone purchase opportunity or an obligating purchase price above their actual WTP.

The observed context dependence of WTP indicates that a valid method should imitate the real purchase situation, with regard to the truly influential background factors affecting WTP (Lynch, 1999), as closely as possible (Hofstetter & Miller, 2009; Wertenbroch & Skiera, 2002). The theory of "bounded rationality" suggests that environmental characteristics influence consumers' choices (Simon, 1955). Furthermore, because consumers construct their preferences with regard to the choice context rather than reveal pre-defined preferences (Bettman et al., 1998), it becomes clear that the decision context is part of an individual's utility function (Kahneman & Tversky, 1979).

2.1 Evaluation of the BDM

The BDM has been tested and evaluated a number of times (see e.g., Miller et al., 2011; Müller et al., 2010; Noussair et al., 2004; Völckner 2006a; Wertenbroch & Skiera, 2002). Most of the evidence shows that the BDM has some appealing properties that make it meaningful and contribute to its broad usage in scientific research. Several studies indicate that the BDM is (theoretically) incentive compatible (Hofstetter & Miller, 2009; Irwin et al., 1998; Wertenbroch and Skiera, 2002, among others), as long as expected utility theory applies (Horowitz, 2006), which, apart from the operational advantages, is the most important benefit of the BDM. Furthermore, Wang et al. (2007) suggest a high level of internal and external validity; the findings of Miller et al. (2011), who found the BDM results to fit real demand curves best, support this suggestion. On the basis of their results, Wertenbroch and Skiera (2002) also argue in favor of high face, internal, and criterion validity. Additionally, they showed the robustness of BDM results in the case of WTP elicitation at the POS under real market conditions.

A rather disputable point is whether the BDM is easy to understand. Hofstetter and Miller (2009) and Wertenbroch and Skiera (2002) argue that the procedure is transparent and well understood by the participating subjects. However, Schreier and Werfer (2007) find the understanding to be a source of potentially biased results, and Völckner (2006b) lists several studies (Kagel & Levin, 1993, among others) that show problems with understanding – at least, understanding the dominant strategy – of the BDM

Although the BDM might be a clearly recommendable method to elicit WTP in theory, its application is limited to academic research, as it seems to play no role in the practice of market research (Steiner & Hendus, 2012). Moreover, even though the BDM is (theoretically) incentive compatible, it has some limitations, which might explain its infrequent application in market research.

First, the procedure might be applicable to existing, low-priced products only[5], as incentive compatibility is bound to the binding character of the purchase, in case there is no sellable product, or any kind of limited liquidity of the participants might lead to biased WTP results, as demand is artificially decreased (Wertenbroch & Skiera, 2002). Kaas and Ruprecht (2006) even suggest a high underestimation of WTP in the case of new and high-involvement products.

[5] Recently, however, Lieven and Lennerts (2013) successfully applied the BDM to non-existing products with the help of vouchers, which are bought on the spot and redeemable once the product is marketed.

Second, as the BDM is often applied in the laboratory, it also faces typical concerns raised in artificial environments and potentially resulting in biases therefrom (Irwin et al., 1998). Muller and Ruffieux (2011, p. 183) suggest congruency between the experimental situation and the real purchase situation is critical, as it points favorably to external validity. A less important point in this realm seems to be the presence of price tags, which are easy to include in the experimental procedure (see this study). However, in case no price tags are available to the participants, it might lead to a situation in which risk-averse decision-makers, who are uncertain about the market price, will include this uncertainty in the form of a risk premium in their WTP statement. In fact, this leads to reduced WTP, and there will be no disclosure of the true WTP (Muller & Ruffieux, 2011). Moreover, Harrison, Harstad, and Rutström (2004) show that in a binding context, either the subjectively perceived or objective market prices censor the WTP, that is, field-price censoring.

Third, the decisions in – and therefore perhaps the results of – the BDM and a real purchase situation are quite different. The BDM asks for a maximum amount a person is willing to pay for a certain product; however, at the POS the consumer has only to decide whether to buy at a given price or not.

Fourth, as a direct WTP elicitation method, the BDM might suffer from high price consciousness and potential strategic bidding behavior (Kaas & Ruprecht, 2006; Le Gall-Ely, 2009; Simon & Kucher, 1988)

Fifth, there have been concerns that the range of the prices used within the BDM procedure to determine the final selling price (distributional dependence) possibly affects the respondents and that consequently the price bids deviate from expected utility theory (Horowitz, 2006).

Other (incentive-compatible) WTP elicitation methods raise the concerns above as well, and thus it is rather a question of how the BDM is implemented, that is, the setting and presentation of stimuli. It is often possible to address these concerns in an appropriate experimental design that (i) uses existing products with real demand and prevents cash limitations of the participants, (ii) accounts for the situation specificity of the WTP by resemblance of important background factors of real POS purchases, and (iii) investigates potential influences of the BDM procedure that might give rise to bias, as the methods we use to measure our results always influence these results (Stewart, 2009). The potential interaction effects of the participants' personal traits and the way the BDM is conducted (lottery) are of particular interest. Subsection 2.2 discusses the possibility of such interactions, and the subsequent section 3 suggests an appropriate experimental design.

2.2 Potential Distortions of BDM Results Due to Gambling

For the BDM, Völckner (2006b) suggested that possible strategic underbidding or overbidding might be the result of the BDM procedure's specific setup. In fact, the procedure originated in experimental economics and has been designed to elicit the certainty equivalent of monetary lotteries (Becker et al., 1964). To prove incentive compatibility, the authors based their argument on the assumptions of expected utility theory (Becker et al., 1964). Even so, this is a traditional economic approach when dealing with risky choices such as the outcomes of lotteries (Edwards, 1954). It might be that one

cannot automatically assume bidders to have a pre-defined utility function and make risky choices in such a way that they maximize their expected utility, as predicted in theory. In general, economic literature has observed that violations of the independence axiom, which implies a preference function that is linear in probabilities, mainly harm the validity of expected utility theory empirically (Machina, 1982). Edwards (1954) proposed that bidders have probability preferences; for instance, they prefer a low probability of making large losses over a high probability of making small losses. Independence is violated mainly when people overweight the low probability of extreme outcomes, as has been frequently observed (Machina, 1982). Karni and Safra (1987) provided evidence that the BDM mechanism is not incentive compatible for eliciting the certainty equivalent of lotteries if bidders' preference relations violate the expected utility hypothesis. More recently, Horowitz (2006) claimed that this result holds true, even for non-random goods, under the assumption that the distribution of prices affects bidders' responses.

A popular approach contrary to the expected utility hypothesis is prospect theory (Kahneman & Tversky, 1979). Prospect theory states that utility is determined rather by relative changes in wealth compared with a reference point – perceived as either gains or losses – than by the absolute value of an outcome. The expected utility of a risky choice in prospect theory is shaped by the assumptions that losses loom larger than gains, and low probabilities are overweighted (Kahneman & Tversky, 1979). Especially when participants are uncertain about their explicit WTP, prospect theory – in particular loss aversion – might imply systematic underbidding in the BDM (Kaas & Ruprecht, 2006). In addition, it would be advisable to consider the idea of regret (Horowitz, 2006). People might derive disutility from forgoing the opportunity to purchase. In particular, regret might affect external validity if it implies some kind of gambling behavior that is tied to the BDM's lottery setting. Consequently, participants could be willing to pay a higher price in a BDM setting than they would actually pay in a real market.

If potential gambling effects (Völckner, 2006a, pp. 139-140) or the closely linked risk attitude (de Meza & Reyniers, 2013) play a role in WTP elicitation, it is important to remember that the classical theory of expected utility does not account for gambling effects (Diecidue, Schmidt, & Wakker, 2004; Fishburn, 1980; Luce & Marley, 2000). In fact, the utility of gambling has been "considered the main cause of deviation from expected utility in the economics literature up to 1944" (Diecidue et al., 2004, p. 242). In other words, the procedure might potentially allow interactions with specific personal traits, such as risk attitude or loss aversion. As the BDM procedure (as depicted in Figure 1) includes a lottery, such interactions seem very likely.

The bidder cannot directly influence the final selling price, which is the basis for incentive compatibility, but can influence the odds of winning the lottery. Hence, the subject's propensity for gambling might influence the results. Diecidue et al. (2004) argue that the incorporation of a risky option (e.g., in the form of uncertainty about the final selling price) will change the decision context, which leads to a different valuation function. This means the possibility of gambling potentially leads

to additional (dis)utility for respondents in a BDM procedure, as the included lottery might represent a risky situation for them. “People do perceive a categorical difference between risky and riskless. As soon as a sure outcome is changed into a risky gamble, no matter how small the risk, new emotions are triggered, and people turn to a different evaluation procedure, the one for risky choices” (Diecidue et al., 2004, p. 243). As Tversky (1967) explicitly highlighted, gambling has some kind of utility, which can affect the differences between risky and riskless utilities. In this context, the BDM as a kind of lottery might yield different utility for different kinds of people depending on the degree of individual risk aversion.

To risk loving respondents, the BDM lottery might offer a positive additional utility, as the utility of gambling overcompensates the disutility (or cost) of gambling. This might lead to an overstated WTP, as the additional utility is positively priced into the WTP bid. By contrast, in the case of risk-averse respondents, the disutility of gambling might outweigh the positive utility of gambling and thus lead to understated WTP.

The transformation function weighting these potential additional positive or negative utilities from gambling is individual and therefore different for different degrees of risk aversion (Diecidue et al., 2004). This means the BDM might systematically elicit different WTPs for different levels of risk attitude, or for people who are prone to gambling, versus those who are not. This might be the result of either the differently perceived utility of the overall situation (i.e., overall utility; which is the utility of the product plus the utility gained or reduced by the procedure (lottery)) or, as suggested, a different perception of the probabilities in the lottery, which influences the perceived utility (Diecidue et al., 2004). If there is a gambling effect, the BDM procedure, which includes a lottery, should yield different utility for different degrees of risk attitude and/or loss aversion. Consequently, this study hypothesizes that if a gambling effect influences the WTP elicited with the BDM, this WTP is significantly different for risk-averse and less risk-averse participants, and/or for different degrees of loss aversion.

2.3 Derived Hypotheses

The work above leads us to the following hypotheses:

The WTB derived from the stated WTP in the BDM procedure predicts the PR in the DCT procedure.

H1a: The estimated purchase rates from BDM and DCT are identical.

In case a significant difference exists in these estimates, this might be due to:

H1b: differences in perceived degree of realism. The bids of participants with an above-mean level of perceived degree of realism will be significantly different from those of participants scoring below the mean.

H1c: differences in risk attitude. The bids of participants with an above-mean level of risk attitude will be significantly different from those of participants scoring below the mean.

H1d: differences in loss aversion in risky choice (number of lotteries played). The bids of participants with an above-mean level of loss aversion will be significantly different from those of participants scoring below the mean.

3. Method and Study Design

3.1 Toward an Appropriate WTP Elicitation Procedure

As highlighted above, the BDM has several appealing properties. However, the BDM is not immune to potential biases. A careful setup of the procedure is mandatory, because for the most potential biases are stemming from the participants of the study, the specific purchase situation created, and the specific character of the BDM mechanism. In the following, the paper reports on these important issues for WTP elicitation (for a detailed overview of the experimental procedure, see Figure 2 in Section 2.2.).

Preventing issues linked to demand and financial matters

This study used a non-convenience sample, in accordance with Ding et al. (2005) and Müller et al. (2010). Participants were pre-screened for their interest in buying at least three of eight products, in order to ensure this study employs meaningful products for the participants of the experiment and thus avoid a negative demand bias in WTP/WTB measures.

This study implemented out-of-pocket transactions to avoid any "house-money effect," that is increased risk-seeking in the presence of a prior gain (Thaler & Johnson, 1990). The show-up fee was handed over directly, while the recruitment took place, on average, two weeks before the experiment, as an initial monetary endowment or participation fee might have affected consumers' reference point and generated biased valuations (Rosenboim & Shavit, 2012; Thaler, 1980). The time span should be large enough to avoid any mental accounting, as the fee of €10 should be spent in the meantime. Therefore, participants had to pay any purchases out of their own pocket (McClure, Laibson, Loewenstein, & Cohen, 2004; Rosenboim & Shavit, 2012)

Participants' liquidity constraints might provoke downward-biased WTP estimates (Wertenbroch & Skiera, 2002, p. 238) when they are dealing with goods that are more expensive. To circumvent this limitation, all eligible participants were advised to bring €70 in cash to the experiment.

Taking the situation specificity of WTP into account

This point addresses the parallelism between the laboratory and the real purchase situation. During a review of the empirical results, it becomes clear that the BDM results are not necessarily the best approximation of the true WTP (de Meza & Reyniers, 2013; Horowitz, 2006; Kaas & Ruprecht,

2006). A common explanation lies in the artificial environment of laboratory-based studies and the potentially perceived risk[6] resulting from some uncertainty in the decision context. We took several steps to minimize the lack of congruency between a real purchase situation and this laboratory-based study, and to limit uncertainty. This study provided all the subjects with about the same amount of information that is available at the POS, as suggested by Zhao and Kling (2004). It displayed the current market prices in each decision situation as they are a relevant background factor for purchase decisions (Carlsson, Frykblom, & Lagerkvist, 2007; Drichoutis, Lazaridis, & Nayga Jr, 2008; Muller & Ruffieux, 2011). Moreover, at each participant's place we placed a catalog containing key product characteristics (including the market price again), a short description from the producer of each of the eight products, and a product evaluation in the form of the original Amazon customer product feedback (for an example, see Appendix 1). This catalogue could be used to support any potential purchase decision. Improving people's knowledge about the products would first increase congruency and second reduce perceived uncertainty.

Control of BDM specific issues affecting WTP

When eliciting homegrown values, it is possible to improve the accuracy of bids, that is for example less systematic underbidding in the first rounds (Noussair et al., 2004) and other systematic distortions due to a lack of understanding (Völckner, 2006b, p. 51), by learning through repetition and explicit instructions or explanations of the dominant strategy (Rutström, 1998). Therefore, this study engaged participants in a training trial before the experiment started, as several researchers have suggested (Plott & Zeiler, 2005; Völckner, 2006a; Wertenbroch & Skiera, 2002). To avoid anchoring as much as possible, the current example used a different price range from that of the products offered later in the experiment. As mentioned above in Section 2.2, there is a debate on the comprehensibility of the procedure or at least the dominant strategy. To address this point, this study explained the dominant strategy (Irwin et al., 1998) to the respondents and ensured comprehension of the procedure by using check-up questions in the computer-based questionnaire, which the participants had to answer correctly in order to reach the binding purchase questions. Furthermore, we placed a separate sheet explaining the procedure on each place available to the respondents at any time. During the experiment, students were encouraged to ask questions and four interviewers per session (20 participants) handled these questions.

Behavioral literature stresses that generally preferences are not necessarily known to subjects and are rather unstable over time (Kahneman & Snell, 1992). The concept of constructive preferences even suggests consumers do not have pre-defined preferences in the first place. The assumption is that they construct valuations dependent on the context (Bettman et al., 1998), and thus known experimental price cues could serve as an anchor (Ariely, Loewenstein, & Prelec, 2003). Tymula, Woelbert, and

[6] This study does not use the classical distinction between risk and uncertainty, as suggested by decision theory.

Glimcher (2013) suggest two possible situations. In the first, consumers who face uncertainty about their true WTP might infer some information from the price distribution (e.g., price as a signal for quality). In the second, respondents might conclude that the intrinsic value of a good depends on its price. That is to say, price itself would be part of the consumer's utility. Thus, incentives for truthful bidding vary across different distributions of random prices (Lusk et al., 2007). Specifically, under a uniform distribution of prices, risk-averse bidders who are uncertain about their true WTP are assumed to underestimate their WTP (Kaas & Ruprecht, 2006), since the expected surplus would be greater than is the case for overestimation. This is in line with the predictions of Lusk et al. (2007), who argue that the degree of underbidding depends on bidders' expectations or knowledge about the random price range and is reduced with less risk aversion. Horowitz (2006) suggests distributional dependence even occurs if consumers are conscious of their true WTP but have different perceptions of disappointment regarding the outcome of the BDM lottery. Accordingly, a bidder can experience two types of ex-post disappointment, namely (1) paying a price above his or her true valuation, and (2) not receiving the good at all. In accordance with prospect theory, if a bidder weights these possible outcomes differently, price distribution could serve as a reference point to minimize ex-post regret. For example, disappointment from not receiving the good might be higher when a bidder's true WTP is relatively higher than the average or most likely price of the distribution. Thus, the bidder would likely overstate his or her true WTP to minimize ex-post regret. Consequently, at a given average WTP, Horowitz (2006) would predict that a right-skewed distribution could provoke underbidding, while a left-skewed distribution could provoke overbidding.

As Bohm, Lindén, and Sonnegård (1997) and Tymula et al. (2013) proposed, this study used a price range for each of the eight products that was unknown to the respondents, in order to avoid anchoring effects (Tversky & Kahneman, 1974) and limit uncertainty. However, participants were told that it would be a range around the current market price, which again was unknown to the participants, this study used eight different urns, each with a specific price range from which the subjects chose the final selling price. This setup was expected to increase the participants' confidence in the procedure. The range for each product was roughly between –30% and 5% around the stated market price. Within this range, we used a uniform probability density function, as suggested by Lusk et al. (2007).

3.2 Product Choice and Sample

Prior to the study, two important choices had to be made: the selection of the target population, and the determination of purchase-relevant products for the intended sample. The latter decision was based on several focus group interviews and pre-tests in which the following four product categories were identified as relevant for the defined target population, which comprises students at a major German university: electrical toothbrush, whiskey, external hard drive, and headphones. Each category consisted of two products from the same brand, to eliminate brand influences on WTB/WTP (Del Rio, Vazquez, & Iglesias, 2001; Dodds, Monroe, & Grewal, 1991). However, the products offered had different prices and quality options. In all cases this study used well-known brands, with positive user

feedback on Amazon.de in order to limit speculation on the quality of the goods. As for the prices, the cheapest available price during recruitment from Amazon.de or regional dealers was used[7].

To avoid potential confounding effects between the elicitation of WTP/WTB and the elicitation of other psychological constructs, this study also measured individual risk attitude and loss aversion (number of lotteries played) during the recruitment process. To measure the first construct, this study employed the 11-point risk attitude scale used in the German Socio-Economic Panel (SOEP), which many academic researchers have adopted (see e.g., Dohmen et al., 2011; Jaeger et al., 2010; Van Winden, Krawczyk, & Hopfensitz, 2011). Loss aversion (number of lotteries played) was also measured during the sampling process using a measure of loss aversion in risky choices. To that end, this study adopted a simple lottery choice task (drop of a coin), following Fehr and Goette (2007). In this choice task, individuals decide for each of 10 lotteries whether they want to play or reject it (and receive nothing). In each lottery, the winning price is fixed at €5.00, and only the losing price varies (between €0.00 and €1.90). At the end of the recruitment process, one lottery was randomly selected for play to secure independence between the lotteries (Cubitt, Starmer, & Sugden, 1998).

From 238 recruited students, 222 (93.28%) turned up, which is a satisfactory result. In order to prepare the data for analysis, suspicious cases were identified and eliminated: Respondents who either indicated in all eight BDM purchase situations a WTP of €0.00 or, in the case of the DCT, selected the "no-buy" option in all 16 decisions, were deleted. This was a reasonable step, as the sampling was based on the purchase relevance of these products for the participants; if a respondent indicated no interest in buying at all, he or she should not be part of the defined target population and therefore, the sample. Consequently, six respondents from the BDM and nine from the (DCT) condition were excluded. Therefore the net sample for further analysis contains 207 valid responses. For sample details, see Table 1 below.

[7] For a list of all products and corresponding market prices, see Appendix 2.

Table 1: Sample composition

Split	N (%)	Gender		Age		Origin	
		male	female	from 18 to 22	from 23 to 31	Former West Germany	Former East Germany
Overall	207 (100%)	124 (59.9%)	83 (40.1%)	107 (51.7%)	100 (48.3%)	111 (53.6%)	96 (46.4%)
BDM	107 (51.7%)	60 (56.1%)	47 (43.9%)	55 (51.4%)	52 (48.6%)	56 (52.3%)	51 (47.7%)
DCT	100 (48.3%)	64 (64.0%)	36 (36.0%)	52 (52.0%)	48 (48.0%)	55 (55.0%)	45 (45.0%)
X^2	-	1.352		0.007		0.147	
p	-	0.245		0.931		0.701	

3.3 Procedure

The laboratory-based study was run as a computer-aided interview using a between-subjects design. Before the random assignment into the experimental splits, the participants had to wait in front of the laboratory. This was the first time they saw the products included in the experiment as these were presented at a table next to the laboratory entrance. All participants had an opportunity to touch and evaluate the products, and everybody was informed about the binding character of the experiment. By drawing a seat number upon entering the laboratory, participants were randomly assigned to one of the 20 computer workstations and one of the two experimental procedures (BDM and DCT).

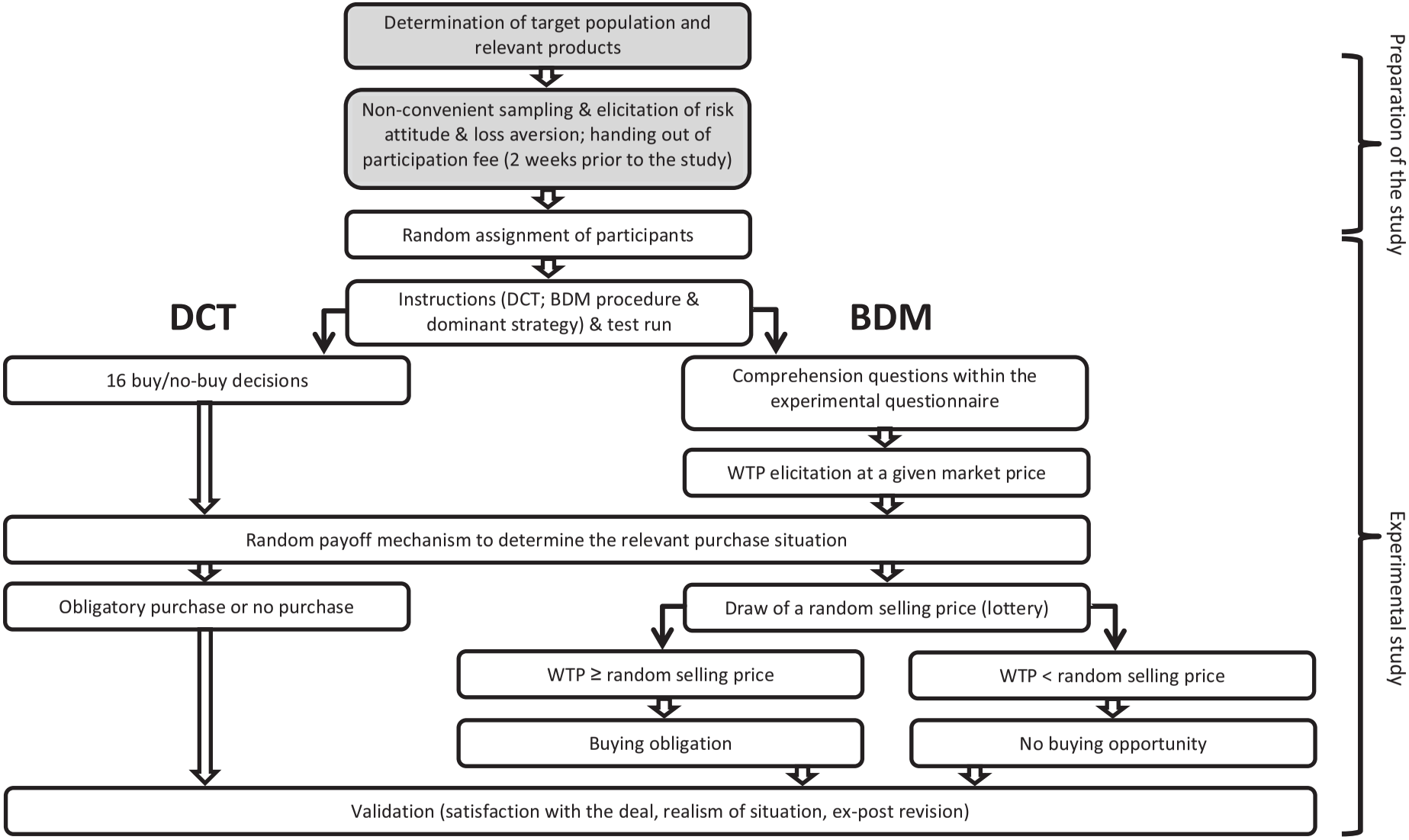
In order to introduce real choice consequences for all decisions, this study used a random payoff mechanism (Grether & Plott, 1979), because participants had to indicate in several purchase situations (BDM: 8; DCT: 16) at which price/or which product they wanted to buy, if any. Thus, participants had been instructed before the experiment started that they would face several potential purchase situations; however, the computer chose only one, at random, for each participant at the end of the study. Previous research shows that this mechanism induces independence of choices, as it prevents income or portfolio effects that are assumed to bias choices in sequential multistage decisions that a single respondent makes (Braga, Humphrey, & Starmer, 2009; Grether & Plott, 1979).

Both procedures (BDM, DCT) followed the same process. The two questionnaires consisted of three parts. The first asks for the origin (former West Germany vs. former East Germany) of the participants as well as for purchase interest, price-quality orientation, and demographics such as age, gender, body size, and level of education of both parents, which could have an impact on the general willingness to take risks (Dohmen et al., 2011). In particular, Dohmen et al. (2011) showed that women are significantly less willing to take risks than men in general and in many specific risky types of behaviors like car driving, financial matters, and health related issues. Moreover, the authors reveal that increasing age reduces willingness to take risks, and that taller individuals and children of parents with a higher degree of education are more willing to take risks.

The second part contains the binding purchase questions. In the BDM and DCT, these questions were asked both in the presence of the market price (the tag value) and with the same decision support material. In fact, the only difference was in the method used to elicit product value in terms of WTP (BDM) or WTB (DCT): Whereas the BDM asked the participants in eight purchase situations to state the maximum amount they would pay, the DCT asked in 16 situations whether the respondents would like to buy the product. Here each product was offered at two different prices. First, and in random order, the higher price option of all eight products was offered. Second, again in random order, all eight products were offered again at the lower price. To be able to compare the predicted PR derived from the BDM with the elicited PR from the DCT, this study took two prices for each product from the price density function that was derived from the first four experimental sessions in which solely WTP with the BDM was elicited (for the session schedule, see Appendix 3). Between the second and third parts of the questionnaire, the random payoff mechanism was used to determine the relevant binding situation, and in the case of the BDM, the respondents were asked to draw from an urn a price from an unknown price range with a uniform distribution around the market price for the specific product determined in the random draw just before. For a detailed overview of the experimental procedure, see Figure 2.

The third and last part asks how satisfied respondents are with the result (purchase/no purchase), what they think about the final selling price, and whether they would be willing to change their decision for the drawn situation if they had a chance to do so.

Figure 2: Detailed experimental procedure



4. Results

First, this study tested for structural differences between the BDM and DCT sample. No significant differences exist for any of the tested variables: age, gender, body size, education level of parents, price versus quality importance, risk attitude, loss aversion in risky choice (number of lotteries played), product relevance and interest in buying (all $t \leq 1.668$, $p \geq .097$; all $\chi^2 \leq 1.926$, $p \geq .165$; see Appendices 4–7). This study therefore assumes that both samples are comparable in their characteristics. Furthermore, as the price density function that determined the two selling prices in the DCT condition was derived from the first four BDM sessions, this study tested whether there is a difference between the WTP values elicited in the first four and the successive BDM sessions. The results indicate no significant differences (all $t \leq 1.087$, $p \geq .279$, see Appendix 8), which suggests that the price density functions from the first four BDM sessions can be used safely to select the selling prices for DCT.

4.1 Face and Criterion Validity

In accordance with Müller et al. (2012a), Völckner (2006a), and Wertenbroch & Skiera (2002), among others, the current study tested the meaningfulness of data by comparing the resulting estimates with demand-specific data to test face or criterion validity. As a result, BDM and DCT (WTB) seemed to lead respondents to calculate and reveal their true WTP for the specific purchase situation, since valuations were correlated with demand-specific measures. WTP (all Pearson's $r \geq .585$; $p < .001$) and WTB (all Pearson's $r \geq .279$; $p < .001$) correlated significantly with interest in buying. However, the WTB measure of the Oral-B Professional Care 3000 showed a non-significant positive correlation of $r = .106$, $p > .100$ (see Appendices 9–10).

In addition, this study followed the analysis in Wertenbroch and Skiera (2002, p. 234) and controlled for strategic deviation in both conditions. It used the same indicators of overbidding and underbidding, which it obtained from the third part of the questionnaire after the announcement of the relevant purchase situation and the individually drawn prices.

The analysis involved an examination of whether participants complied with their purchase obligation and were satisfied with their decision. The argumentation is based on the assumption that not getting the product at a price below one's WTP, as well as overpaying one's WTP, results in dissatisfaction or even a refusal to obey the purchase obligation. Since all buyers complied with their purchase obligation in the BDM and DCT condition and were mainly satisfied with their decision (on a 7-point Likert-type scale from 1 (very unsatisfied) to 7 (very satisfied), BDM: M=5.24, SD=1.85; DCT: M=5.46, SD=1.78), they did not seem to have significantly understated or overstated their true WTP.
As suggested by Kaas and Ruprecht (2006), this study additionally tested behavioral incentive compatibility in the BDM condition. However, the actual outcome for each participant (i.e., the drawn product and price) was not recorded as in Kaas and Ruprecht (2006); this study used the hypothetical

question on the willingness to revise the decision of the binding situation ex-post instead. The result indicates that 14.2% (20%) in the BDM (DCT) condition would change their WTP statement if they could. This procedure leaves us with a somewhat weaker indicator than that used by Kaas and Ruprecht (2006). However, subjects evaluated the purchase occasion at the resulting price on a 7-point Likert-type scale from "too expensive, bad deal" to "very good deal." In combination with the dichotomous choice whether participants wanted to change their decision in the drawn binding purchase situation, this study was partially able to estimate their bidding strategy. This indicator should yield valid results, as the situation closely resembles a real shopping situation in which the price is given, and subjects need to evaluate the purchase occasion at this given price (Kaas & Ruprecht, 2006, p. 42). Respondents who wanted to change their decision and simultaneously said the product was a bad deal probably had bid too much, while the respondents who thought the product was a good deal most likely had underbid. In case of a neutral evaluation of the deal, this study was not able to assume underbidding or overbidding. Our approach classifies 3.8% (17%) of the subjects in the BDM (DCT) as potential underbidders, and 7.5% (0%) of subjects as potential overbidders (see Tables 2 and 3).

Table 2: Possible underbidding and overbidding in the BDM

	Evaluation: purchase is a **good deal**	Evaluation: product is **too expensive, bad deal**	Evaluation: neutral, **neither good nor bad deal**
I want to change my decision	Possible underbidding **3.8%** of subjects	Possible overbidding **7.5%** of subjects	Not indicative **2.8%** of subjects
I do not want to change my decision	Possible truthful bidding **40.6%** of subjects	Possible truthful bidding **20.8%** of subjects	Not indicative **24.5%** of subjects

Table 3: Possible underbidding and overbidding in the DCT

	Evaluation: purchase is a **good deal**	Evaluation: product is **too expensive, bad deal**	Evaluation: neutral, **neither good nor bad deal**
I want to change my decision	Possible underbidding **17%** of subjects	Possible overbidding **0%** of subjects	Not indicative **3%** of subjects
I do not want to change my decision	Possible truthful bidding **65%** of subjects	Possible truthful bidding **5%** of subjects	Not indicative **10%** of subjects

To further test criterion validity, this study analyzed whether participants were willing to pay more for products they found to be relevant for purchase than for products they found not to be relevant. As expected, the analysis shows that WTP for relevant products is significantly higher than for non-relevant ones, with one exception: All $t \geq 3.536$, $p \leq .001$ and $\chi^2 \geq 18.290$, $p \leq .000$ show a significant difference (see Appendices 11 and 12). Only the values for the Oral-B Professional Care 3000 WTP ($t(15.653) = 2.087$, $p = .054$) and WTB ($\chi^2(1) = 3.171$, $p = .209$) do not differ significantly. However, this seems to be due to the very low number of participants willing to buy the toothbrush, as WTP was found to be more than twice as high for participants that indicated the Oral-B Professional Care 3000

to be relevant (€13.08 vs. €26.33), and there was a more than eightfold increase in WTB (1.1% vs. 9.1%).

All in all, the current research finds both measures to have satisfactory face and criterion validity. However, the BDM seems to produce fewer participants willing to revise their decision ex-post.

4.2 Predictive Validity of the BDM

In an identical setting (time, place, general setup) using a matched sample that does not show any structural differences (see Section 3), the current research found a significant difference between the results of the BDM and DCT procedure.

To compare the BDM with the DCT, this study used the predicted PR from the BDM, which is the ratio of respondents who indicated a WTP equal to or above the chosen price point[8], and the real PR from the validation sample in the DCT condition. For example, in the case of the Oral-B 500 toothbrush, one can see at the price of €19.99 a predicted PR of 21.3% in the BDM, and 16.2% in the DCT condition. Between both measures, the difference (in this example, 5.1 percentage points) for each of the eight products was calculated (see Table 3).

Table 3: Difference in PR between BDM and DCT

Product	1st selling price in DCT	PR (BDM) in % N = 108	PR (DCT) in % N = 99	GAP in pp	2nd selling price in DCT	PR (BDM) in % N = 108	PR (DCT) in % N = 99	GAP in pp
Oral-B Professional Care 500	19.99 €	21.3	16.2	**5.1**	14.99 €	35.2	24.2	**10.9**
Oral-B Professional Care 3000	39.99 €	14.8	2.0	**12.8**	29.99 €	24.1	7.1	**17.0**
Johnnie Walker Red Label	9.99 €	26.9	23.2	**3.6**	7.99 €	42.6	31.3	**11.3**
Johnnie Walker Black Label	19.99 €	25.0	19.2	**5.8**	14.99 €	36.1	27.3	**8.8**
Intenso Memory Station 320GB	34.99 €	15.7	10.1	**5.6**	24.99 €	35.2	17.2	**18.0**
Intenso Memory Station 500GB	39.99 €	16.8	24.2	**-7.4**	34.99 €	28.7	27.3	**1.4**
Sony MDR-ZX300	19.99 €	21.3	16.2	**5.1**	14.99 €	39.8	23.2	**16.6**
Sony MDR-ZX600	34.99 €	15.9	13.1	**2.8**	30.99 €	16.8	12.1	**4.7**

To test whether this difference between the BDM and DCT is significant, an independent samples *t*-test was conducted. Indeed, the PR in the BDM are significantly lower than in DCT ($t(30)= 2.48$, $p=.019$). Moreover, a non-parametric test to support the results of the *t*-test and account for the low number of observations was conducted. The Mann-Whitney U test shows similar results and predicts a

[8] In order to transfer individual WTP into PR, this study assumed that all prices up to the stated WTP are acceptable for purchase, as it is common practice in price research. Thus, leaving aside the possibility of too low prices, this might be an indication of quality issues and therefore lead to a situation where participants do not purchase the good below an individual floor price as Van Westendorp (1976) suggested.

medium-effect size for the difference between the BDM and DCT results (U= 77, z= -1.92, p= .056, r= -.34). Thus H1a can be rejected, as this research finds a significant difference between the results of both procedures.

4.3 Investigating the Gap Between BDM and DCT Estimates

There were several steps to test hypotheses H1b through H1d. First, as suggested by Müller et al. (2010, p. 118), one can see that participants in the DCT condition (M= 4.71, SD= 1.31) rate the procedure as more realistic than the ones in the BDM condition (M= 4.27, SD= 1.357) (t(205)= 2.363, p= .019). This assertion supports the intended scenario in which we like to compare the BDM results with results that are closer to a more realistic purchase situation. This difference in the perceived degree of realism might already explain the difference between BDM predictions and DCT results. However, an independent samples t-test shows no significant difference in WTP between participants scoring above the mean rating of the degree of realism (7-point Likert-scale type, M= 4.48, SD= 1.35) and those scoring below the mean (all $t\leq$ 1.239, $p\geq$ .219, see Appendix 13). Thus, H1b finds no support.

In order to investigate the difference between the BDM and DCT (hypotheses H1c and H1d), this study tested whether the gap might be explained by two attributes. As suggested by de Meza and Reyniers (2013), as well as Kaas and Ruprecht (2006), a potential source of bias in the BDM results might be the individual risk attitude. An independent samples t-test shows no significant difference in risk attitude between the BDM and DCT condition (t(204)= .858, p= .392, see Appendix 4). The same holds for loss aversion in risky choice (number of lotteries played). The comparison between BDM and DCT shows no significant difference in the number of lotteries played (t(204)= .137, p= .891, see Appendix 4). Having found no difference between the BDM and DCT for both measures, this research can analyze whether risk attitude or loss aversion in risky choice might moderate WTP statements and thus be indicative of the measured deviation between the BDM and DCT giving rise to a potential gambling effect influence on BDM results. This study tested whether the WTPs or PR of respondents with an above-average rating in risk attitude (BDM: M= 5.35, SD= 1.76; DCT: M= 5.57, SD= 1.94) lead to significantly different results than those of respondents scoring below the average. However, all WTPs were insignificant, which suggests no difference between WTPs of more risk-averse respondents and less risk-averse ones in the BDM (all $t\leq$ 1.251, $p\geq$ .221, see Appendix 14). As expected, the same holds for the PR in the DCT, which does not include a lottery (all $\chi^2 \leq$ 2.369, $p\geq$ .124, see Appendix 15). Furthermore, we analyzed whether the degree of loss aversion might interact with the BDM procedure and thus produce significantly different WTP for participants having a higher-than-average degree versus a lower-than-average degree of loss aversion (number of lotteries played). Again, all WTPs were insignificant, which suggests no difference between WTPs for different degrees of loss aversion (all $t\leq$ .867, $p\geq$ .388, see Appendix 16); the same holds for the PR in the DCT (all $\chi^2 \leq$ 3.214, $p\geq$ .073, see Appendix 17). This means hypotheses H1c and H1d can be rejected. Having found no indication for WTP differences between different degrees of risk attitude and loss

aversion in BDM results, this analysis cannot support the notion of a gambling effect interacting with the BDM procedure and therefore biasing WTP estimates.

Further, this study tested the relative explanatory power of risk attitude, loss aversion in risky choice (number of lotteries played), and interest in buying. In particular, this study investigated with logistic regressions whether the distinction between predicted buyer and predicted non-buyer could be explained using these variables. In general, risk attitude and loss aversion cannot significantly explain whether the BDM procedure classifies someone as a buyer or a non-buyer, which suggests that neither variable might influence the WTP in a BDM setting. When the dependent variable of predicted buyer (vs. non-buyer) is regressed on loss aversion in risky choices (number of lotteries played) or on risk attitude, the logit model is – with the exception of one case – insignificant (all $\chi^2(1) \leq 2.320$, $p \geq .128$, see Appendices 18a and 18d). The logit model for the Sony MDR-ZX300 at the first selling price shows a χ^2 value of 3.904 (*df= 1*) and a *p*-value of .048. A logit model with the two independent variables risk attitude and loss aversion in risky choice (number of lotteries played) yields a model χ^2 of 6.862 (*df= 2*) with a corresponding p-value of .032; for the model variable risk attitude, a significant p-value of .037. However, this might potentially be a type 1 error. Both pseudo R^2 values in the single independent variable model (Cox and Snell R^2= .036, Nagelkerke R^2= .057), as well as in the model with two independent variables (Cox and Snell R^2= .063, Nagelkerke R^2= .099), indicate very limited improvement of the full model over the intercept model. Moreover, the logit model with one single independent and the one with two independent variables, as well as the independent variable risk attitude itself, becomes insignificant for the second selling price, $\chi^2(1)$= 1.373 with p= .241 and $\chi^2(2)$= 1.584 with p= .453. Interest in buying is the only variable that significantly contributes to the distinction between predicted buyer and predicted non-buyer. The independent variable interest in buying becomes significant when it is added to the logit model. Neither risk attitude or loss aversion in risky choices (number of lotteries played) is significant in any of the logit models that include interest in buying (see Appendices 18b, 18c and 18e).

To see whether the determinants of risk attitude that have been suggested by Dohmen et al. (2011) help to explain the segmentation into (predicted) buyer and (predicted) non-buyer, this study compared the distinct sub-groups for each of the eight products with each other: real buyer versus real non-buyer in DCT (see Appendix 19a), predicted buyer versus predicted non-buyer in BDM (see Appendix 19b), predicted buyer in BDM versus real buyer in DCT (see Appendix 19c), predicted non-buyer in BDM versus real non-buyer in DCT (see Appendix 19d). In particular, this study compared these segments with regard to their structural identity using Person's χ^2 test, or Fisher's exact test if indicated. In order to gain a sufficiently large number of observations for each category in each cell of the contingency table, all metric scales were split into two groups using the median or the mean as delimiter. Three variables (gender, education level of father, and education level of mother) were used in their original form, median splits of age and body size, as well as mean splits of risk attitude, and loss aversion

(number of lotteries played). In most of the cases, the two groups in the comparison did not show significantly different results for these variables. However, in some cases, single variables indicated a structural difference between the subgroups under consideration. In 448 χ^2-based structural identity checks, 12 turned out to be significant; however caution is needed interpreting these, as at least 22 studies are expected false positives given a significance level of 5%. In particular, six incidences of a significant difference in gender, three in level of education of mother, and one in body size, level of education of father, as well as in risk attitude, were detected. The age variable did not show any significant difference between the subgroups. Nevertheless, some differences are even expected and therefore support the validity of the presented results. When comparing predicted buyer and predicted non-buyer with each other, as well as real buyer with real non-buyer, this research finds six significant differences with respect to gender in both types of whiskey. Not surprisingly, male participants are more likely to purchase whiskey and show a higher WTP (for details, see Appendices 19a and 19b). Similarly, this study expects that the significant difference and the marginally significant ($p<.1$) values in body size between the predicted buyer and predicted non-buyer of Johnnie Walker Red and Black Label at selling prices one and two also reflect the different tastes among genders, given that males on average are taller than women. In this example, in the case of a significant difference arising at selling price one between predicted buyer and predicted non-buyer, one will very likely find a significant difference in selling price two as well. This is because all predicted buyers from the higher price (selling price one) will be among the predicted buyers at selling price two (i.e., the lower price). Except for the differences in tastes for whiskey between both genders, this research did not find a meaningful pattern of structural differences between the compared groups. Thus, it cannot show that the variables, linked to risk attitude as suggested by Dohmen et al. (2011), distinguish between the subgroups of (predicted) buyer and (predicted) non-buyer.

5. Conclusion and Discussion

The expected utility hypothesis predicts that bidders reveal their true WTP in the BDM mechanism. However, empirical findings suggest that bids are indeed affected by, first, a grasp of the procedure and the dominant strategy (Kagel & Levin, 1993) and, second, potential deviations from expected utility due to perceived risk (de Meza & Reyniers, 2013), a gambling effect (Horowitz, 2006; Völckner, 2006b), or the underlying distribution of prices (Bohm et al., 1997; Horowitz, 2006; Tymula et al., 2013).

This research finds empirical evidence for a deviation of the predicted PR from BDM compared with the real PR of the DCT. The BDM was found to overstate real PR in 15 out of 16 price points, compared with a DCT situation perceived as more realistic. This overstatement of WTP is somewhat surprising, as in direct WTP elicitation methods – to which the BDM belongs – cognitive effort is

higher in order to derive one's true WTP and the price, as part of the overall purchase situation tends to be somewhat overemphasized in such settings, which should lead to an underestimation of the true WTP (Hofstetter & Miller, 2009). Nevertheless, it seems plausible to find differences in predicted versus real PR because of the dissimilarity in perceived degree of realism between both procedures. In BDM, the respondents face a setup perceived to be less realistic; however, the question remains whether this perception really results in behavior that is different at the POS. The results of the current study can give only limited support for a significant change in corresponding behavior, as additional behavioral indicators, with the exception of the WTP (WTB), were not measured.

The results of the current study and other studies (Kaas & Ruprecht, 2006; Miller et al., 2011; Wertenbroch & Skiera, 2002) seem to be contradictory regarding the direction of the deviation between the BDM and DCT results. Kaas and Ruprecht (2006) suggested a systematic underbidding due to risk aversion in their BDM results. By contrast, this study finds that the WTP, and thus the predicted PR from BDM, is too high, which is in line with the findings of Müller and Voigt (2010). The current findings might also be supported on theoretical terms by Horowitz (2006). He suggests that the possibility of ex-post regret from not receiving the good at all might lead to an overstatement of WTP to minimize ex-post regret.

Given the results run contrary to earlier research, it is crucial to consider possible explanations that might address the inconsistency. A first salient difference concerns the experimental setting. The current study implemented the BDM in a laboratory setting with students who consider themselves a relevant target group of the presented product, while Kaas and Ruprecht (2006) and Miller et al. (2011) conducted their experiments in class with university students, without reporting a pre-selection process. Obviously, different samples might influence the results; even samples from the same population, for example students, do not consistently react to experimental stimuli in the same way (Peterson & Merunka, 2014). Furthermore, this study used eight rather medium-priced shopping goods, instead of the low-priced, fast-moving consumer goods other studies have used. There is evidence suggesting that in this case respondents will make a more deliberate choice (Deshpande & Hoyer, 1983; Holton, 1958; Hoyer, 1984; Kaish, 1967). This study did not track whether this has been the case, and one cannot be sure a more careful decision process would influence WTP bids significantly. However, all studies were set in a monopolistic context, participants were told about the dominant strategy of the BDM mechanism, and results were compared with real transaction data, albeit elicited at different selling point settings (field vs. laboratory). Nevertheless, bidders in the former settings seemed to deviate downward, whereas bidders in the latter setting seemed to deviate upward. It is possible that, in a real POS setting, participants might define their WTP more carefully, and thus risk-averse subjects in particular would rather underestimate their WTP (Kaas & Ruprecht, 2006). By contrast, less risk-averse subjects might not deviate from true WTP or even be willing to bid above their actual WTP. However, this research found no significant difference in WTP and WTB for any risk type of bidder (risk attitude and loss aversion) influencing BDM estimates and as-expected

DCT results; see Appendices 14-17. Even so, it did not find any support for an influence of the individual risk attitude or loss aversion (number of lotteries played) on WTP and WTB measures; this study cannot tell whether the *perceived* risk might have such an effect. It might well be that the degree of perceived risk between the BDM and DCT procedures differs significantly between both methods and therefore mitigates the differences in (predicted) PR.

While the gambling effect was also suggested as a potential reason for biased BDM results (Völckner, 2006a; Müller & Voigt, 2010), this study could not find empirical support for this contention. According to the literature from decision theory, this is not surprising, as the gambling effect is based on the availability of a risky and simultaneously non-risky option within the same choice set (Bleichrodt & Schmidt, 2002). However, the classical BDM offers no such non-risky option. It is debatable whether a bid of €0.00 defines a riskless option. Similarly, this study cannot exclude the possibility that participants of the study did not consider an outside option during the elicitation of the WTP, which was not controlled for in the experiment. Thus, it still might be that participants considered the possibility of buying the products outside the laboratory as a riskless reference situation. Nevertheless, the extensive analysis of the potential influence of risk attitude and loss aversion in risky choice on WTP (WTB) within this study cannot completely rule out the potential influence of these factors. However, the presented results suggest a very limited influence. Nevertheless, if perceived risk does play a role, it might be possible to prevent the results from being biased. A "cheap talk design," suggested by Cummings and Taylor (1999), might lead to a situation in which respondents price this perceived risk into their decisions. Research in the realm of priming supports the fruitfulness of this approach (Schwarz & Clore, 1983).

In light of several studies finding under-prediction and over-prediction, it seems plausible that an underlying yet unknown motive moderates these two aspects in the BDM procedure. Kaas and Ruprecht (2006) suggest that this might be because of product familiarity. However, and more plausible, a part of the respondent's underlying goal function, as it directly enters the bidding strategy of respondents. Precisely what is the respondent looking for during the elicitation of WTP? It might be a good bargain or the actual product, and Appendix 20 suggests such a moderating motive. Further research is invited to investigate this possibility in depth.

6. Limitations

The current study has several limitations. First, even though the respondents rated the elicitation method of WTP (WTB) as rather realistic, it might say nothing about their corresponding behavior. So, the question of how realistic they perceived the situation to be is probably not as good as the question whether they behaved similarly in the experimental situation as they would in a purchase situation at the POS.

Second, this study used two price points in the DCT condition; even so, the product order was random and all the higher prices of each product were shown first for each product. This might, however, have

anchored the respondents. It might be that in order to be consistent and avoid cognitive dissonance, the subjects recalled whether they had bought the product at the first offered price and then stuck to that decision without an independent new consideration of the current situation. However, this might work for purchases and non-purchases, and therefore cancels out.

Third, this study elicited WTP and WTB in monopolistic situations due to the used monadic product offers. Even so, this was true for both procedures and should therefore not matter when comparing both. However, it is not completely clear how the WTP/WTB changes in relation to each other in competitive designs. Research findings from experimental economics suggest that a subject's preference order between options can change significantly when competitive or monadic elicitation procedures are used (see preference reversals, Lichtenstein & Slovic, 1971; Müller, Kroll, & Vogt, 2012b).

Fourth, the existing literature on binding (incentive compatible) WTP/WTB concentrates on rather low-value, fast-moving consumer goods (e.g., food). The usage of moderately priced shopping goods has not been considered so far, giving rise to the question of whether earlier results apply to the current setting in which customers are supposed to enter into a more demanding decision process in which they carefully consider quality and price (AMA; Kaish, 1967, p. 29).

Fifth, consumers might simply not be able to determine their true WTP if their preferences are rather constructive than pre-defined (Bettman et al., 1998). In general, one needs to be careful in accessing the external validity of any WTP elicitation method as WTP is a latent construct and thus not directly observable. It is not clear which procedure is finally closer to the true WTP. The information content of the current findings, therefore, is limited. Having said that, it becomes clear that besides pure method comparisons (convergent validity), additional real transaction data from POS is needed to evaluate external validity. The approach used in the current study sits in the middle of these attempts of assessing external validity as it uses two different elicitation methods in an, even in terms of time, identical setting. However, these results should be further tested and replicated in order to test the generalizability of BDM results (Hubbard & Armstrong, 1994) to support scientific knowledge base updates (Lehmann & Bengart, 2015).

Finally, it is worth mentioning that the usage of students has its limitations, and therefore, results need further replications before they might be generalized beyond the research setting (Peterson, 2001). However, in the light of time and budget constraints and to be consistent with prior research on the validity of the BDM (see e.g, de Meza & Reyniers, 2013; Müller et al., 2010; Wertenbroch & Skiera, 2002), this study used students as well, although knowing that the generalization to other student populations might be questionable (Peterson & Merunka, 2014). Nevertheless, the defined target population can reasonably be considered a regular and profitable buyer segment within the general target group of the eight products being tested.

7. Bibliography

AMA. (Ed.) Dictionary of the American Marketing Association. American Marketing Association.

Ariely, D., Loewenstein, G., and Prelec, D. (2003). "Coherent Arbitrariness": Stable Demand Curves Without Stable Preferences. *The Quarterly Journal of Economics, 118*(1), 73-106.

Bateman, I., Munro, A., Rhodes, B., Starmer, C., and Sugden, R. (1997). A test of the theory of reference-dependent preferences. *The Quarterly Journal of Economics,* 479-505.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9*(3), 226-232.

Bettman, J. R., Luce, M. F., and Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research, 25*(3), 187-217.

Bleichrodt, H., and Schmidt, U. (2002). A context-dependent model of the gambling effect. *Management Science, 48*(6), 802-812.

Bohm, P., Lindén, J., and Sonnegård, J. (1997). Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanism vs. Markets. *The Economic Journal, 107*(443), 1079-1089.

Braga, J., Humphrey, S. J, and Starmer, C. (2009). Market experience eliminates some anomalies—and creates new ones. *European Economic Review, 53*(4), 401-416.

Calder, B. J., Phillips, L. W., and Tybout, A. M. (1981). Designing research for application. *Journal of Consumer Research, 8*(2), 197-207.

Calder, B. J., Phillips, L. W., and Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research, 9*(3), 240-244.

Carlsson, F., Frykblom, P., and Lagerkvist, C. J. (2007). Preferences with and without prices-does the price attribute affect behavior in stated preference surveys? *Environmental and Resource Economics, 38*(2), 155-164.

Carson, R. T., and Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics, 37*(1), 181-210.

Cubitt, R. P., Starmer, C., and Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics, 1*(2), 115-131.

Cummings, R. G., and Taylor, L. O. (1999). Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review, 89*(3), 649-665.

de Meza, D., and Reyniers, D. (2013). Debiasing the Becker–DeGroot–Marschak valuation mechanism. *Economics Bulletin, 33*(2), 1446-1456.

Del Rio, A. B., Vazquez, R., and Iglesias, V. (2001). The effects of brand associations on consumer response. *Journal of Consumer Marketing, 18*(5), 410-425.

Deshpande, R., and Hoyer, W. D. (1983). *Consumer decision making: strategies, cognitive effort and perceived risk.* Paper presented at the AMA Educators' Proceedings.

Diecidue, E., Schmidt, U., and Wakker, P. P. (2004). The utility of gambling reconsidered. *Journal of Risk and Uncertainty, 29*(3), 241-259.

Diller, H. (2007). *Preispolitik*. W. Kohlhammer Verlag. Stuttgart.

Ding, M., Grewal, R., and Liechty, J. (2005). Incentive-aligned conjoint analysis. *Journal of Marketing Research, 42*(1), 67-82.

Dodds, W. B., Monroe, K. B., and Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research, 28*(3), 307.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association, 9*(3), 522-550.

Drichoutis, A. C., Lazaridis, P., and Nayga Jr, R. M. (2008). The role of reference prices in experimental auctions. *Economics Letters, 99*(3), 446-448.

Edwards, W. (1954). The Theory of Decision Making. *Psychological Bulletin, 51*(4), 380.

Ehrenberg, A. S. C. (1993). Even the Social Sciences Can have Laws. *Nature, 365*, 385.

Fehr, E., and Goette, L. (2007). Do workers work more if wages are high? Evidence from a randomized field experiment. *The American Economic Review, 91*(1), 298-317.

Fishburn, Peter C. (1980). A simple model for the utility of gambling. *Psychometrika, 45*(4), 435-448.

Grether, D. M., and Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review, 69*(4), 623-638.

Harrison, G. W, Harstad, R. M, and Rutström, E. E. (2004). Experimental methods and elicitation of values. *Experimental economics, 7*(2), 123-140.

Hoffman, E., Menkhaus, D. J., Chakravarti, D., Field, R. A., and Whipple, G. D. (1993). Using laboratory experimental auctions in marketing research: A case study of new packaging for fresh beef. *Marketing science, 12*(3), 318-338.

Hofstetter, R., and Miller, K. (2009). Bessere Preisentscheidungen durch Messung der Zahlungsbereitschaft. *Marketing Review St. Gallen, 26*(5), 32-37.

Holton, R. H. (1958). The distinction between convenience goods, shopping goods, and specialty goods. *Journal of Marketing, 23*(3), 53-56.

Horowitz, J. K. (2006). The Becker-DeGroot-Marschak mechanism is not necessarily incentive compatible, even for non-random goods. *Economics Letters, 93*(1), 6-11.

Hoyer, W. D. (1984). An examination of consumer decision making for a common repeat purchase product. *Journal of Consumer Research, 11*(3), 822-829.

Hubbard, R., and Armstrong, J. S. (1994). Replications and extensions in marketing: Rarely published but quite contrary. *International Journal of Research in Marketing, 11*(3), 233-248.

Huber, J., Payne, J. W., and Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research, 9*(1), 90-98.

Irwin, J. R., McClelland, G., McKee, M., Schulze, W. D, and Norden, N. E. (1998). Payoff dominance vs. cognitive transparency in decision making. *Economic Inquiry, 36*(2), 272-285.

Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., and Bonin, H. (2010). Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics, 92*(3), 684-689.

Kaas, K. P., and Ruprecht, H. (2006). Are the vickrey auction and the BDM mechanism really incentive compatible?- empirical results and optimal bidding strategies in cases of uncertain willingness-to-pay. *Schmalenbach Business Review, 58*(1), 37-55.

Kagel, J. H, and Levin, D. (1993). Independent private value auctions: Bidder behaviour in first-, second-and third-price auctions with varying numbers of bidders. *The Economic Journal, 103*(419), 868-879.

Kahneman, D., and Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making, 5*(3), 187-200.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica,* 47(2), 263-291.

Kaish, S. (1967). Cognitive dissonance and the classification of consumer goods. *Journal of Marketing, 31*(4), 28-31.

Kalish, S., and Nelson, P. (1991). A comparison of ranking, rating and reservation price measurement in conjoint analysis. *Marketing Letters, 2*(4), 327-335.

Karni, E., and Safra, Z. (1987). "Preference reversal" and the observability of preferences by experimental methods. *Econometrica, 55*(3), 675-685.

Le Gall-Ely, M. (2009). Definition, Measurement and Determinants of the Consumer's Willingness to Pay: A Critical Synthesis and Avenues for Further Research. *Recherche et Applications en Marketing, 24*(2), 91-112.

Lehmann, S., and Bengart, P. (2015). Replications hardly possible: Reporting practice in top-tier marketing journals. *Journal of Modelling in Management,* forthcomming. doi: 10.1108/JM2-04-2014-0030

Lichtenstein, S., and Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*(1), 46.

Lieven, T., and Lennerts, S. (2013). Measuring willingness to pay by means of the trade-off between free available cash and specific-purpose vouchers. *BuR-Business Research, 6*(2), 154-171.

List, J. A, and Gallet, C. A. (2001). What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics, 20*(3), 241-254.

Luce, R D., and Marley, A. AJ. (2000). On elements of chance. *Theory and Decision, 49*(2), 97-126.

Lusk, J. L., Alexander, C., and Rousu, M. C. (2007). Designing experimental auctions for marketing research: The effect of values, distributions, and mechanisms on incentives for truthful bidding. *Review of Marketing Science, 5*(1).

Lynch, J. G. (1982). On the external validity of experiments in consumer research. *Journal of Consumer Research, 9*(3), 225-239.

Lynch, J. G. (1983). The role of external validity in theoretical research. *Journal of Consumer Research, 10*(1), 109-111.

Lynch, J. G. (1999). Theory and external validity. *Journal of the Academy of Marketing Science, 27*(3), 367-376.

Machina, M. J. (1982). "Expected Utility" analysis without the independence axiom. *Econometrica, 50*(2), 277-323.

Marshall, A. (1920). *Principles of Economics* (8 ed.). London: Macmillan and Co.

McAfee, R. P., and McMillan, J. (1987). Auctions and bidding. *Journal of Economic Literature, 25*(2), 699-738.

McClure, S. M, Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503-507.

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

Mitchell, R. C., and Carson, R. T. (1989). *Using surveys to value public goods: the contingent valuation method*: Routledge.

Müller, H., Kroll, E. B., and Vogt, B. (2012a). Do real payments really matter? A re-examination of the compromise effect in hypothetical and binding choice settings. *Marketing Letters, 23*(1), 73-92.

Müller, H., Kroll, E. B., and Vogt, B. (2012b). Violations of procedure invariance — The case of preference reversals in monadic and competitive product evaluations. *Journal of Retailing and Consumer Services, 19*(4), 406-412.

Müller, H., Voigt, S., and Erichson, B. (2010). Ermittlung von zahlungsbereitschaften mittels monadischer preis- und kaufabfragen: Neue empirische erkenntnisse. *Marketing ZFP: Journal of Research and Management, 32*(2), 117-127.

Muller, L, and Ruffieux, B. (2011). Do price-tags influence consumers' willingness to pay? On the external validity of using auctions for measuring value. *Experimental Economics, 14*(2), 181-202.

Murphy, J. J., Allen, P. G., Stevens, T. H., and Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics, 30*(3), 313-325.

Noussair, C., Robin, S., and Ruffieux, B. (2004). Revealing consumers' willingness-to-pay: A comparison of the BDM mechanism and the vickrey auction. *Journal of Economic Psychology, 25*(6), 725-741.

Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second order meta analysis. *Journal of Consumer Research, 28*(3), 450-461.

Peterson, R. A., and Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research, 67*(5), 1035-1041.

Plott, C. R, and Zeiler, K. (2005). The Willingness to Pay–Willingness to Accept Gap, the. *The American Economic Review, 95*(3), 530-545.

Rosenboim, M., and Shavit, T. (2012). Whose money is it anyway? Using prepaid incentives in experimental economics to create a natural environment. *Experimental Economics, 15*(1), 145-157.

Rutström, E. E. (1998). Home-grown values and incentive compatible auction design. *International Journal of Game Theory, 27*(3), 427-441.

Schreier, M., and Werfer, J. (2007). Auktionen versus lotterien: Ein empirischer vergleich zur messung von zahlungsbereitschaften. *Die Betriebswirtschaft: DBW*, *67*(1), 22-40.

Schwarz, N., and Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of personality and social psychology, 45*(3), 513.

Shogren, J. F, Cho, S., Koo, C., List, J., Park, C., Polo, P., and Wilhelmi, R. (2001). Auction mechanisms and the measurement of WTP and WTA. *Resource and Energy Economics, 23*(2), 97-109.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*(1), 99-118.

Simon, H. (2004). Ertragssteigerung durch effektivere pricing-prozesse. *Zeitschrift für Betriebswirtschaft, 11*, 1083-1102.

Simon, H., and Kucher, E. (1988). Die bestimmung empirischer preisabsatzfunktionen : methoden, befunde, erfahrungen. *Journal of Business Economics, 58*(1), 171-183.

Steiner, M., and Hendus, J. (2012). How consumers' willingness to pay is measured in practice: An empirical analysis of common approaches' relevance. *Working Paper Series University Münster*.

Stewart, D. W. (2009). The role of method: some parting thoughts from a departing editor. *Journal of the Academy of Marketing Science, 37*(4), 381-383.

Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization, 1*(1), 39-60.

Thaler, R. H. (1985). Mental accounting and consumer choice. *Marketing Science*, *4*(3), 199-214.

Thaler, R. H, and Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science, 36*(6), 643-660.

Tversky, A. (1967). Additivity, utility, and subjective probability. *Journal of Mathematical psychology, 4*(2), 175-201.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Tymula, A., Woelbert, E., and Glimcher, P. (2013). Flexible valuations for consumer goods as measured by the Becker-DeGroot-Marschak mechanism. *University of Sydney, School of Economics Working Paper Series*, 2013-20, 1-23.

van Westendorp, P. H. (1976). *NSS price sensitivity meter (PSM)–A new approach to study consumer perception of prices.* Paper presented at the ESOMAR Congress (25 th) Venice.

Van Winden, F., Krawczyk, M., and Hopfensitz, A. (2011). Investment, resolution of risk, and the role of affect. *Journal of Economic Psychology, 32*(6), 918-939.

Völckner, F. (2006a). An empirical comparison of methods for measuring consumers' willingness to pay. *Marketing Letters, 17*(2), 137-149.

Völckner, F. (2006b). Methoden zur Messung individueller Zahlungsbereitschaften: Ein Überblick zum State of the Art. *Journal für Betriebswirtschaft*, *56*(1), 33-60.

Wang, T., Venkatesh, R, and Chatterjee, R. (2007). Reservation price as a range: An incentive-compatible measurement approach. *Journal of Marketing Research, 44*(2), 200-213.

Wertenbroch, K., and Skiera, B. (2002). Measuring consumers' willingness to pay at the point of purchase. *Journal of Marketing Research, 39*(2), 228-241.

Winer, R. S. (1999). Experimentation in the 21st century: The importance of external validity. *Journal of the Academy of Marketing Science, 27*(3), 349-358.

Zhao, J., and Kling, C. L. (2004). Willingness to pay, compensating variation, and the cost of commitment. *Economic Inquiry, 42*(3), 503-517.

8. Appendix

1) Product catalogue page, example external hard drive
(full material available on request)

Intenso Memory Station 500 GB

Speicherkapazität	500 Gigabyte
Anschluss	USB 2.0
Formgröße	2,5 Zoll
Üblicher Marktpreis (Amazon)	52,99 €

Produktbeschreibung:

:: Memory Station 2,5"
Kapazität:
500 GB
Formfaktor:
2,5"
Farbe:
Schwarz | Weiß | Grün | Pink
Material:
Kunststoff
Stromversorgung:
High Speed USB-Y Anschluss (2 Anschlüsse notwendig) - keine externe Stromversorgung notwendig
Max. Leserate:
33,00 MB/s (221x)
Max. Schreibrate:
35,50 MB/s (238x)
Zubehör:
USB-Y Kabel

2) Product setup and selling prices in DCT

Product group	Product	Market price (in €)	1st selling price in DCT (in €)	2nd selling price in DCT (in €)
1. toothbrush	Oral-B Professional Care 500	30.90 (Amazon.de)	19.99	14.99
	Oral-B Professional Care 3000	67.16 (Amazon.de)	39.99	29.99
2. whiskey	Johnnie Walker Red Label	12.99 (local retail)	9.99	7.99
	Johnnie Walker Black Label	29.45 (local retail)	19.99	14.99
3. external hard drive	Intenso Memory Station 320GB	46.39 (Amazon.de)	34.99	24.99
	Intenso Memory Station 500GB	52.99 (Amazon.de)	39.99	34.99
4. headphones	Sony MDR-ZX300	24.44 (Amazon.de)	19.99	14.99
	Sony MDR-ZX600	44.99 (Amazon.de)	34.99	30.99

3) Experimental Session Schedule

Time/day	Monday	Wednesday	Friday
8 am	BDM	DCT	BDM & DCT
10 am	BDM	DCT	BDM & DCT
12 am	BDM	DCT	BDM & DCT
2 pm	BDM	DCT	BDM & DCT

4) Structural identity test BDM versus DCT, metric variables

Indicator	Overall		BDM		DCT		Difference BDM vs. DCT		
N	207		107		100				
	M	SD	M	SD	M	SD	t	df	p
Age	22.75	2.56	22.70	2.49	22.81	2.65	.305	205	.761
Body size	177.00	9.74	176.39	9.85	177.64	9.62	.921	205	.358
Risk attitude*	5.46	1.85	5.35	1.76	5.57	1.94	.858	204	.392
Loss aversion*	6.98	3.10	7.01	3.13	6.95	3.08	.137	204	.891
Pricequality importance: toothbrush	4.36	1.41	4.34	1.35	4.39	1.49	.272	205	.786
Pricequality importance: external hard drive	3.95	1.57	3.89	1.59	4.01	1.56	.557	205	.578
Pricequality importance: whiskey	4.68	1.58	3.72	1.80	3.62	1.76	.403	205	.688
Pricequality importance: headphones	3.67	1.78	4.50	1.54	4.87	1.61	1.668	205	.097

*Overall N =206 (106 BDM; 100 DCT)

5) Structural identity test BDM versus DCT, nominal variables

Indicator	Difference BDM vs. DCT		
	X^2	df	p
Gender	1.352	1	.245
Education level of father	.051	205	.959
Education level of mother	.045	205	.964

6) Structural identity test BDM versus DCT: Differences in product relevance

Product	Product relevance			Difference BDM vs. DCT		
	Overall	BDM	DCT	χ^2	df	p
N	207	107	100	-	-	-
Oral-B Professional Care 500	67 (32.4%)	36 (33.6%)	31 (31.0%)	.165	1	.684
Oral-B Professional Care 3000	26 (12.6%)	15 (14.0%)	11 (11.0%)	.429	1	.513
Johnnie Walker Red Label	108 (52.2%)	55 (51.4%)	53 (53.0%)	.053	1	.818
Johnnie Walker Black Label	64 (30.9%)	34 (31.8%)	30 (30.0%)	.076	1	.782
Intenso Memory Station 320GB	56 (27.1%)	26 (24.3%)	30 (30.0%)	.851	1	.356
Intenso Memory Station 500GB	81 (39.1%)	37 (34.6%)	44 (44.0%)	1.926	1	.165
Sony MDRZX300	92 (44.4%)	47 (43.9%)	45 (45.0%)	.024	1	.876
Sony MDRZX600	61 (29.5%)	36 (33.6%)	25 (25.0%)	1.859	1	.173

7) Structural identity test BDM versus DCT: Differences in interest in buying

Product		Interest in buying			Difference BDM vs. DCT		
		Overall	BDM	DCT	*t*	df	p
	N	207	107	100			
Oral-B Professional Care 500	M	2.22	2.13	2.32	.816	205	.416
	SD	1.67	1.61	1.72			
Oral-B Professional Care 3000	M	1.59	1.70	1.48	1.396*	179.434	.164
	SD	1.16	1.38	.86			
Johnnie Walker Red Label	M	2.83	2.86	2.80	.221	205	.825
	SD	1.94	2.00	1.88			
Johnnie Walker Black Label	M	2.52	2.56	2.47	.344	205	.732
	SD	1.90	1.92	1.88			
Intenso Memory Station 320GB	M	2.46	2.41	2.51	.388	205	.698
	SD	1.83	1.81	1.85			
Intenso Memory Station 500GB	M	2.87	2.76	3.00	.811	205	.418
	SD	2.15	2.03	2.28			
Sony MDRZX300	M	2.93	2.92	2.94	.088	205	.930
	SD	1.96	2.04	1.89			
Sony MDRZX600	M	2.62	2.53	2.71	.667	205	.499
	SD	1.88	1.84	1.93			

* Welch's *t*-test

8) Mean WTP check between the first four sessions and successive sessions

Product	Mean WTP (SD)		Difference		
	first	successive	t	df	p
Oral-B Professional Care 500	9.60 (8.15)	11.81 (10.46)	1.087*	52.365	.282
Oral-B Professional Care 3000	14.24 (15.64)	16.44 (18.12)	0.643	105	.522
Johnnie Walker Red Label	6.37 (4.61)	6.50 (4.83)	0.137	105	.891
Johnnie Walker Black Label	10.10 (8.58)	12.17 (10.37)	1.087	105	.279
Intenso Memory Station 320GB	16.20 (13.28)	17.51 (15.50)	0.451	105	.653
Intenso Memory Station 500GB	19.69 (15.72)	21.40 (19.11)	0.454*	54.591	.651
Sony MDR-ZX300	10.79 (8.16)	11.56 (8.15)	0.452	105	.652
Sony MDR-ZX600	17.85 (13.69)	16.76 (13.13)	0.385	104	.701

* Welch's *t*-test

9) Correlations between interest in buying and WTP

Product	Pearson's r	Bootstrapped confidence intervals[a]
Oral-B Professional Care 500	.649*	[.753 ; .508]
Oral-B Professional Care 3000	.635*	[.765 ; .435]
Johnnie Walker Red Label	.710*	[.785 ; .613]
Johnnie Walker Black Label	.705*	[.798 ; .590]
Intenso Memory Station 320GB	.596*	[.722 ; .447]
Intenso Memory Station 500GB	.665*	[.768 ; .551]
Sony MDRZX300	.689*	[.782 ; .580]
Sony MDRZX600	.649*	[.754 ; .515]

[a] [upper ; lower]

* p < 0.01; ** p < 0.05; *** p <0.1.

10) Correlations between interest in buying and WTB

Product	Selling price	Pearson's r	Bootstrapped confidence intervals[a]
Oral-B Professional Care 500	1st	.593*	[.707 ; .449]
Oral-B Professional Care 3000	1st	.106	[.358 ; -.111]
Johnnie Walker Red Label	1st	.531*	[.660 ; .380]
Johnnie Walker Black Label	1st	.695*	[.790 ; .567]
Intenso Memory Station 320GB	1st	.479*	[.603 ; .323]
Intenso Memory Station 500GB	1st	.683*	[.784 ; .567]
Sony MDRZX300	1st	.531*	[.668 ; .355]
Sony MDRZX600	1st	.597*	[.710 ; .459]
Oral-B Professional Care 500	2nd	.622*	[.744 ; .496]
Oral-B Professional Care 3000	2nd	.279*	[.495 ; .002]
Johnnie Walker Red Label	2nd	.697*	[.785 ; .580]
Johnnie Walker Black Label	2nd	.713*	[.811 ; .584]
Intenso Memory Station 320GB	2nd	.507*	[.636 ; .364]
Intenso Memory Station 500GB	2nd	.680*	[.787 ; .551]
Sony MDRZX300	2nd	.583*	[.710 ; .450]
Sony MDRZX600	2nd	.573*	[.691 ; .430]

[a] [upper ; lower]
* p < 0.01; ** p < 0.05; *** p <0.1.

11) WTP for relevant versus non relevant products

Product	Mean WTP (SD)		Difference		
	not relevant	relevant	t	df	p
Oral-B Professional Care 500	7.36 (6.85)	16.09 (9.85)	4.766*	52.693	.000
Oral-B Professional Care 3000	13.08 (14.18)	26.33 (23.92)	2.087*	15.643	.054
Johnnie Walker Red Label	3.78 (3.76)	8.90 (4.04)	6.773	105	.000
Johnnie Walker Black Label	7.58 (7.44)	17.59 (8.95)	6.067	105	.000
Intenso Memory Station 320GB	13.44 (12.45)	26.49 (14.02)	4.509	105	.000
Intenso Memory Station 500GB	16.06 (15.37)	28.13 (16.74)	3.747	105	.000
Sony MDRZX300	1.75 (1.10)	4.40 (1.80)	12.393	143	.000
Sony MDRZX600	14.15 (11.35)	24.28 (14.96)	3.536*	53.938	.001

* Welch's *t*-test

12) Purchase rates difference between relevant versus non relevant products for 1st selling price in DCT

Product	Purchase rates		Difference		
	not relevant	relevant	χ^2	df	p
Oral-B Professional Care 500	1.4%	48.4%	35.064	1	.000*
Oral-B Professional Care 3000	1.1%	9.1%	3.171	1	.209*
Johnnie Walker Red Label	0.0%	43.4%	26.489	1	.000
Johnnie Walker Black Label	2.9%	56.7%	39.509	1	.000
Intenso Memory Station 320GB	1.4%	30.0%	19.048	1	.000*
Intenso Memory Station 500GB	1.8%	52.3%	34.433	1	.000
Sony MDRZX300	1.8%	33.3%	18.290	1	.000
Sony MDRZX600	4.0%	40.0%	21.485	1	.000*

* Fisher's Exact Test.

13) WTP differences between below versus above the mean degree of realism.

Product	Mean WTP (SD)		Difference		
	low	high	t	df	p
Oral-B Professional Care 500	9.54 (8.53)	11.23 (9.47)	.968	105	.335
Oral-B Professional Care 3000	13.13 (14.60)	17.17 (18.32)	1.239*	88.889	.219
Johnnie Walker Red Label	6.20 (4.39)	6.67 (5.01)	.513	105	.609
Johnnie Walker Black Label	10.32 (8.44)	11.30 (10.09)	.537*	91.664	.592
Intenso Memory Station 320GB	15.40 (13.24)	18.10 (14.81)	.995	105	.322
Intenso Memory Station 500GB	19.03 (15.33)	21.71 (18.51)	.804*	91.105	.423
Sony MDRZX300	11.01 (7.68)	11.07 (8.73)	.040	105	.968
Sony MDRZX600	16.10 (12.53)	19.19 (14.46)	1.177	104	.242

* Welch's *t*-test

14) WTP of respondents below versus above the mean risk level

Product	Mean WTP in € (SD)		Difference		
	below mean	above mean	t	df	p
Oral-B Professional Care 500	9.88 (9.25)	10.31 (8.45)	.248	104	.804
Oral-B Professional Care 3000	13.01 (15.73)	15.99 (16.77)	.946	104	.346
Johnnie Walker Red Label	6.22 (4.80)	6.56 (4.62)	.377	104	.707
Johnnie Walker Black Label	10.20 (8.79)	11.08 (9.58)	.492	104	.623
Intenso Memory Station 320GB	17.30 (14.00)	15.62 (13.78)	.620	104	.537
Intenso Memory Station 500GB	20.70 (16.28)	19.39 (17.18)	.402	104	.689
Sony MDRZX300	12.00 (8.60)	10.05 (7.64)	1.231	104	.221
Sony MDRZX600	17.80 (13.82)	16.92 (13.17)	.331	103	.741

15) WTB of respondents below versus above the mean risk level

Product (selling price)	PR in %		Difference		
	below mean	above mean	X^2	df	p
Oral-B Professional Care 500 (1)	11.90	18.97	.904	1	.342
Oral-B Professional Care 3000 (1)	2.38	1.72	.054*	1	1.00
Johnnie Walker Red Label (1)	28.58	18.97	1.269	1	.260
Johnnie Walker Black Label (1)	11.90	24.14	2.369	1	.124
Intenso Memory Station 320GB (1)	7.14	12.07	.657*	1	.513
Intenso Memory Station 500GB (1)	21.42	25.86	.263	1	.608
Sony MDRZX300 (1)	11.9	18.97	.904	1	.342
Sony MDRZX600 (1)	7.14	17.24	2.196	1	.138
Oral-B Professional Care 500 (2)	16.67	29.31	2.135	1	.144
Oral-B Professional Care 3000 (2)	9.52	6.90	.228*	1	.717
Johnnie Walker Red Label (2)	30.95	31.03	.000	1	.993
Johnnie Walker Black Label (2)	21.42	31.03	1.140	1	.286
Intenso Memory Station 320GB (2)	14.29	18.97	.378	1	.539
Intenso Memory Station 500GB (2)	26.19	27.59	.024	1	.877
Sony MDRZX300 (2)	19.05	25.86	.639	1	.424
Sony MDRZX600 (2)	7.14	15.51	1.618	1	.203

* Fisher's exact test

16) WTP of respondents below versus above the mean level of loss aversion (number of lotteries played)

Product	Mean WTP in € (SD)		Difference		
	below mean	above mean	t	df	p
Oral-B Professional Care 500	10.30 (9.73)	9.97 (8.11)	.191	104	.849
Oral-B Professional Care 3000	13.41 (15.52)	15.50 (16.64)	.655	104	.514
Johnnie Walker Red Label	6.40 (4.86)	6.41 (4.59)	.016	104	.988
Johnnie Walker Black Label	9.91 (8.49)	11.23 (9.70)	.016	104	.988
Intenso Memory Station 320GB	16.48 (14.51)	16.33 (13.44)	.057	104	.954
Intenso Memory Station 500GB	18.26 (16.88)	21.21 (16.61)	.867	104	.388
Sony MDRZX300	10.97 (8.58)	10.94 (7.83)	.021	104	.984
Sony MDRZX600	18.53 (14.15)	16.47 (12.91)	.774	103	.441

17) WTB of respondents below versus above the mean level of loss aversion (number of lotteries played)

Product (selling price)	PR in %		Difference		
	below mean	above mean	X^2	df	p
Oral-B Professional Care 500 (1)	11.11	20.00	1.455	1	.228
Oral-B Professional Care 3000 (1)	0.00	3.64	1.670*	1	.500
Johnnie Walker Red Label (1)	22.22	21.18	.028	1	.867
Johnnie Walker Black Label (1)	24.44	14.55	1.576	1	.209
Intenso Memory Station 320GB (1)	15.55	5.45	2.806*	1	.108
Intenso Memory Station 500GB (1)	24.44	23.64	.009	1	.925
Sony MDRZX300 (1)	20.00	12.73	.973	1	.324
Sony MDRZX600 (1)	13.33	12.73	.974	1	.324
Oral-B Professional Care 500 (2)	22.22	25.45	.142	1	.707
Oral-B Professional Care 3000 (2)	4.44	10.91	1.405*	1	.289
Johnnie Walker Red Label (2)	31.11	30.91	.000	1	.983
Johnnie Walker Black Label (2)	26.66	27.27	.005	1	.946
Intenso Memory Station 320GB (2)	24.44	10.91	3.214	1	.073
Intenso Memory Station 500GB (2)	26.66	27.27	.005	1	.946
Sony MDRZX300 (2)	28.88	18.18	1.602	1	.206
Sony MDRZX600 (2)	11.11	12.73	.061	1	.805

***** Fisher's Exact test

18a) Logistic regression: Predicted Buyer in open-ended WTP (DV) regressed on risk attitude (IV)

Product	p-values for P1	N	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2	χ^2	df	Sig.
Oral-B Professional Care 500	.611	106	105.266	.002	.004	.263	1	.608
Oral-B Professional Care 3000	.985	106	82.746	.000	.000	.000	1	.985
Johnnie Walker Red Label	.913	106	124.388	.000	.000	.012	1	.913
Johnnie Walker Black Label	.527	106	117.698	.004	.006	.406	1	.524
Intenso Memory Station 320GB	.608	106	86.170	.002	.004	.261	1	.609
Intenso Memory Station 500GB	.644	106	93.126	.002	.003	.217	1	.642
Sony MDR-ZX300	.051	106	101.624	.036	.057	3.904	1	.048
Sony MDR-ZX600	.827	106	89.912	.000	.001	.048	1	.826
Product	**p-values for P2**	**N**	**-2 Log likelihood**	**Cox & Snell R^2**	**Nagelkerke R^2**	**χ^2**	**df**	**Sig.**
Oral-B Professional Care 500	.326	106	134.862	.009	.013	.985	1	.321
Oral-B Professional Care 3000	.207	106	111.732	.016	.024	1.668	1	.197
Johnnie Walker Red Label	.320	106	144.087	.009	.013	1.006	1	.316
Johnnie Walker Black Label	.438	106	137.730	.006	.008	.610	1	.435
Intenso Memory Station 320GB	.444	106	135.259	.006	.008	.587	1	.443
Intenso Memory Station 500GB	.816	106	124.345	.001	.001	.055	1	.815
Sony MDR-ZX300	.244	106	14.093	.013	.017	1.373	1	.241
Sony MDR-ZX600	.644	106	93.126	.002	.003	.217	1	.642

18b) Logistic regression: Predicted Buyer in open-ended WTP (DV) regressed on risk attitude & interest in buying (IVs)

Product	p-values for P1 Risk attitude	Interest in buying	N	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2	χ^2	df	Sig.
Oral-B Professional Care 500	.771	.000	106	75.520	.247	.391	3.008	2	.000
Oral-B Professional Care 3000	.839	.000	106	53.138	.244	.450	29.608	2	.000
Johnnie Walker Red Label	.660	.000	106	92.695	.259	.374	31.704	2	.000
Johnnie Walker Black Label	.832	.000	106	8.637	.298	.443	37.466	2	.000
Intenso Memory Station 320GB	.865	.000	106	6.691	.216	.387	25.740	2	.000
Intenso Memory Station 500GB	.321	.000	106	65.930	.228	.389	27.413	2	.000
Sony MDR-ZX300	.252	.000	106	82.135	.198	.314	23.393	2	.000
Sony MDR-ZX600	.457	.000	106	62.521[a]	.228	.399	27.439	2	.000
Product	**p-values for P2** Risk attitude	Interest in buying	**N**	**-2 Log likelihood**	**Cox & Snell R^2**	**Nagelkerke R^2**	**χ^2**	**df**	**Sig.**
Oral-B Professional Care 500	.494	.000	106	105.182	.251	.348	3.665	2	.000
Oral-B Professional Care 3000	.138	.001	106	96.292	.149	.227	17.109	2	.000
Johnnie Walker Red Label	.160	.000	106	9.646	.402	.539	54.446	2	.000
Johnnie Walker Black Label	.803	.000	106	103.915	.277	.380	34.425	2	.000
Intenso Memory Station 320GB	.486	.000	106	105.987	.245	.340	29.860	2	.000
Intenso Memory Station 500GB	.508	.000	106	91.083	.270	.390	33.317	2	.000
Sony MDR-ZX300	.851	.000	106	107.591	.274	.371	33.874	2	.000
Sony MDR-ZX600	.312	.000	106	64.434	.239	.408	28.909	2	.000

18c) Logistic regression: Predicted Buyer in open-ended WTP (DV) regressed on risk attitude & interest in buying & lotteries played (IVs)

Product	p-values for P1			N	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2	χ^2	df	Sig.
	Risk attitude	Interest in buying	Lotteries							
Oral-B Professional Care 500	.778	.000	.087	106	72.580	.267	.424	32.948	3	.000
Oral-B Professional Care 3000	.837	.000	.301	106	52.094	.251	.463	3.653	3	.000
Johnnie Walker Red Label	.657	.000	.737	106	92.582	.259	.375	31.818	3	.000
Johnnie Walker Black Label	.834	.000	.986	106	8.637	.298	.443	37.467	3	.000
Intenso Memory Station 320GB	.806	.000	.638	106	6.463	.217	.390	25.968	3	.000
Intenso Memory Station 500GB	.309	.000	.747	106	65.826	.229	.391	27.516	3	.000
Sony MDR-ZX300	.155	.000	.082	106	78.762	.223	.354	26.766	3	.000
Sony MDR-ZX600	.465	.000	.806	106	62.460	.229	.399	27.500	3	.000
Product	**p-values for P2**			**N**	**-2 Log likelihood**	**Cox & Snell R^2**	**Nagelkerke R^2**	**χ^2**	**df**	**Sig.**
	Risk attitude	Interest in buying	Lotteries							
Oral-B Professional Care 500	.496	.000	.819	106	105.129	.252	.348	3.717	3	.000
Oral-B Professional Care 3000	.135	.001	.750	106	96.191	.150	.228	17.209	3	.001
Johnnie Walker Red Label	.154	.000	.667	106	9.463	.403	.540	54.630	3	.000
Johnnie Walker Black Label	.802	.000	.983	106	103.915	.277	.380	34.425	3	.000
Intenso Memory Station 320GB	.520	.000	.539	106	105.612	.248	.344	3.234	3	.000
Intenso Memory Station 500GB	.508	.000	.994	106	91.083	.270	.390	33.317	3	.000
Sony MDR-ZX300	.856	.000	.280	106	106.419	.282	.382	35.047	3	.000
Sony MDR-ZX600	.315	.000	.921	106	64.424	.239	.408	28.918	3	.000

18d) Logistic regression: Predicted Buyer in open-ended WTP (DV) regressed on lotteries played (IV)

Product	p-values for P1	N	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2	χ^2	df	Sig.
Oral-B Professional Care 500	.474	106	105.024	.005	.008	.505	1	.477
Oral-B Professional Care 3000	.936	106	82.74	.000	.000	.006	1	.936
Johnnie Walker Red Label	.590	106	124.104	.003	.004	.296	1	.587
Johnnie Walker Black Label	.397	106	117.358	.007	.010	.745	1	.388
Intenso Memory Station 320GB	.382	106	85.613	.008	.014	.818	1	.366
Intenso Memory Station 500GB	.621	106	93.091	.002	.004	.252	1	.616
Sony MDR-ZX300	.148	106	103.209	.022	.034	2.320	1	.128
Sony MDR-ZX600	.920	106	89.951	.000	.000	.010	1	.920
Product	**p-values for P2**	**N**	**-2 Log likelihood**	**Cox & Snell R^2**	**Nagelkerke R^2**	**χ^2**	**df**	**Sig.**
Oral-B Professional Care 500	.407	106	135.141	.007	.009	.705	1	.401
Oral-B Professional Care 3000	.723	106	113.273	.001	.002	.128	1	.721
Johnnie Walker Red Label	.978	106	145.092	.000	.000	.001	1	.978
Johnnie Walker Black Label	.413	106	137.655	.006	.009	.684	1	.408
Intenso Memory Station 320GB	.930	106	135.839	.000	.000	.008	1	.930
Intenso Memory Station 500GB	.376	106	123.588	.008	.011	.812	1	.368
Sony MDR-ZX300	.592	106	141.180	.003	.004	.286	1	.593
Sony MDR-ZX600	.855	106	93.309	.000	.001	.033	1	.855

18e) Logistic regression: Predicted Buyer in open-ended WTP (DV) regressed on lotteries played & interest in buying (IVs)

Product	p-values for P1		N	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2	χ^2	df	Sig.
	Lotteries	Interest in buying							
Oral-B Professional Care 500	.087	.000	106	72.661	.267	.423	32.868	2	.000
Oral-B Professional Care 3000	.301	.000	106	52.14	.251	.463	3.610	2	.000
Johnnie Walker Red Label	.741	.000	106	92.780	.258	.373	31.620	2	.000
Johnnie Walker Black Label	.970	.000	106	8.681	.297	.443	37.423	2	.000
Intenso Memory Station 320GB	.662	.000	106	6.523	.217	.389	25.908	2	.000
Intenso Memory Station 500GB	.818	.000	106	66.898	.221	.377	26.445	2	.000
Sony MDR-ZX300	.120	.000	106	8.809	.208	.330	24.719	2	.000
Sony MDR-ZX600	.777	.000	106	63.006	.225	.393	26.955	2	.000
Product	**p-values for P2**		**N**	**-2 Log likelihood**	**Cox & Snell R^2**	**Nagelkerke R^2**	**χ^2**	**df**	**Sig.**
	Lotteries	Interest in buying							
Oral-B Professional Care 500	.811	.000	106	105.599	.248	.344	3.247	2	.000
Oral-B Professional Care 3000	.810	.001	106	98.589	.130	.199	14.811	2	.001
Johnnie Walker Red Label	.725	.000	106	92.608	.391	.524	52.485	2	.000
Johnnie Walker Black Label	.998	.000	106	103.978	.277	.380	34.362	2	.000
Intenso Memory Station 320GB	.503	.000	106	106.025	.245	.339	29.822	2	.000
Intenso Memory Station 500GB	.965	.000	106	91.528	.267	.386	32.871	2	.000
Sony MDR-ZX300	.280	.000	106	106.452	.281	.382	35.014	2	.000
Sony MDR-ZX600	.880	.000	106	65.470	.231	.395	27.873	2	.000

18f) Logistic regression: Predicted Buyer in open-ended WTP (DV) regressed on risk attitude & lotteries played (IVs)

Product	p-values for P1 Risk attitude	Lotteries	N	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2	χ^2	df	Sig.
Oral-B Professional Care 500	.578	.452	106	104.710	.008	.012	.818	2	.664
Oral-B Professional Care 3000	.989	.937	106	82.740	.000	.000	.007	2	.997
Johnnie Walker Red Label	.942	.594	106	124.099	.003	.004	.301	2	.860
Johnnie Walker Black Label	.562	.418	106	117.018	.010	.015	1.086	2	.581
Intenso Memory Station 320GB	.560	.361	106	85.277	.011	.019	1.155	2	.561
Intenso Memory Station 500GB	.666	.641	106	92.902	.004	.007	.440	2	.802
Sony MDR-ZX300	.037	.105	106	98.666	.063	.099	6.862	2	.032
Sony MDR-ZX600	.821	.909	106	89.899	.001	.001	.061	2	.970
Product	**p-values for P2** Risk attitude	Lotteries	**N**	**-2 Log likelihood**	**Cox & Snell R^2**	**Nagelkerke R^2**	**χ^2**	**df**	**Sig.**
Oral-B Professional Care 500	.350	.440	106	134.254	.015	.021	1.593	2	.451
Oral-B Professional Care 3000	.213	.780	106	111.654	.016	.025	1.747	2	.418
Johnnie Walker Red Label	.318	.925	106	144.078	.010	.013	1.015	2	.602
Johnnie Walker Black Label	.468	.440	106	137.122	.011	.016	1.217	2	.544
Intenso Memory Station 320GB	.446	.971	106	135.258	.006	.008	.589	2	.745
Intenso Memory Station 500GB	.862	.383	106	123.558	.008	.011	.842	2	.656
Sony MDR-ZX300	.257	.645	106	139.882	.015	.020	1.584	2	.453
Sony MDR-ZX600	.635	.831	106	93.081	.002	.004	.262	2	.877

19a) Structural identity test: Real buyer versus real non-buyer in DCT

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Oral-B Professional Care 500	χ^2	2.460	.138	.519	.978	2.214	.904	1.455
	df	1	1	1	5	4	1	1
	p	.117	.710	.471	.964	.696	.342	.228
	N	100	100	100	100	100	100	100
Oral-B Professional Care 3000	χ^2	.174	2.211	.003	2.705	.543	.054	1.670
	df	1	1	1	5	4	1	1
	p	1.000*	.137	.954	.745	.969	.817	.196
	N	100	100	100	100	100	100	100
Johnnie Walker Red Label	χ^2	.402	.869	.245	7.447	1.743	1.269	.028
	df	1	1	1	5	4	1	1
	p	.526	.351	.621	.189	.783	.260	.867
	N	100	100	100	100	100	100	100
Johnnie Walker Black Label	χ^2	9.618	.327	.1.170	4.286	6.143	2.369	1.576
	df	1	1	1	5	4	1	1
	p	.002	.568	.279	.509	.189	.124	.209
	N	100	100	100	100	100	100	100
Intenso Memory Station 320GB	χ^2	1.235	.641	.641	2.329	3.447	.657	2.806
	df	1	1	1	5	4	1	1
	p	.323*	.423	.541*	.802	.486	.418	.094
	N	100	100	100	100	100	100	100
Intenso Memory Station 500GB	χ^2	.640	.051	.481	1.955	2.246	.263	.009
	df	1	1	1	5	4	1	1
	p	.424	.822	.488	.052	.691	.608	.925
	N	100	100	100	100	100	100	100
Sony MDR-ZX300	χ^2	.497	.031	.519	4.548	5.757	.904	.974
	df	1	1	1	5	4	1	1
	p	.481	.861	.471	.473	.218	.342	.324
	N	100	100	100	100	100	100	100

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX600	χ^2	1.083	.205	.020	8.187	4.577	2.196	.008
	df	1	1	1	5	4	1	1
	p	.367*	.651	.886	.146	.333	.138	.929
	N	100	100	100	100	100	100	100
Oral-B Professional Care 500 [MP2]	χ^2	.640	1.395	1.351	4.298	2.640	2.135	.142
	df	1	1	1	5	4	1	1
	p	.424	.238	.245	.507	.620	.144	.707
	N	100	100	100	100	100	100	100
Oral-B Professional Care 3000 [MP2]	χ^2	2.084	2.540	.014	7.758	2.343	.228	1.405
	df	1	1	1	5	4	1	1
	p	.253*	.111	1.000*	.170	.673	.633	.236
	N	100	100	100	100	100	100	100
Johnnie Walker Red Label [MP2]	χ^2	.947	.003	2.820	8.469	2.191	.000	.000
	df	1	1	1	5	4	1	1
	p	.331	.959	.093	.132	.701	.993	.983
	N	100	100	100	100	100	100	100
Johnnie Walker Black Label [MP2]	χ^2	7.205	.220	3.188	4.337	5.143	1.140	.005
	df	1	1	1	5	4	1	1
	p	.007	.639	.074	.502	.273	.286	.946
	N	100	100	100	100	100	100	100
Intenso Memory Station 320GB [MP2]	χ^2	1.382	.382	.007	3.079	2.099	.378	3.214
	df	1	1	1	5	4	1	1
	p	.240	.536	.932	.688	.718	.539	.073
	N	100	100	100	100	100	100	100
Intenso Memory Station 500GB [MP2]	χ^2	.114	.000	.846	13.060	2.893	.024	.005
	df	1	1	1	5	4	1	1
	p	.735	.986	.358	.023	.576	.877	.946
	N	100	100	100	100	100	100	100

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX300 [MP2]	χ^2	1.813	.208	.869	7.001	8.506	.639	1.602
	df	1	1	1	5	4	1	1
	p	.178	.648	.351	.221	.075	.424	.206
	N	100	100	100	100	100	100	100
Sony MDR-ZX600 [MP2]	χ^2	.716	.022	1.175	5.346	5.613	1.618	.061
	df	1	1	1	5	4	1	1
	p	.529*	.882	.278	.375	.230	.203	.805
	N	100	100	100	100	100	100	100

* Fisher's Exact Test

ms = mean split; mds = median split

19b) Structural identity test: Predicted buyer versus predicted non-buyer in BDM

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Oral-B Professional Care 500	χ^2	.102	.392	.643	3.588	6.809	.120	.286
	df	1	1	1	5	6	1	1
	p	.749	.531	.423	.610	.339	.729	.593
	N	107	107	107	107	107	106	106
Oral-B Professional Care 3000	χ^2	.627	.157	.795	1.631	6.956	.092	.001
	df	1	1	1	5	6	1	1
	p	.428	.692	.373	.897	.325	.761	.947
	N	107	107	107	107	107	106	106
Johnnie Walker Red Label	χ^2	6.324	.156	3.488	4.291	5.493	.067	.334
	df	1	1	1	5	6	1	1
	p	.012	.693	.062	.508	.482	.795	.563
	N	107	107	107	107	107	106	106
Johnnie Walker Black Label	χ^2	4.750	.003	.921	1.631	8.883	.000	.866
	df	1	1	1	5	6	1	1
	p	.029	.957	.337	.897	.180	.993	.352
	N	107	107	107	107	107	106	106
Intenso Memory Station 320GB	χ^2	.315	.177	.000	2.746	8.827	.001	.595
	df	1	1	1	5	6	1	1
	p	.574	.674	.988	.739	.184	.971	.441
	N	107	107	107	107	107	106	106
Intenso Memory Station 500GB	χ^2	.324	2.019	2.291	4.320	1.229	.973	.014
	df	1	1	1	5	6	1	1
	p	.569	.155	.130	.504	.115	.324	.907
	N	107	107	107	107	107	106	106
Sony MDR-ZX300	χ^2	.026	.392	.102	1.774	6.628	2.590	.891
	df	1	1	1	5	6	1	1
	p	.871	.531	.749	.879	.357	.108	.345
	N	107	107	107	107	107	106	106

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX600	χ^2	.081	.019	1.728	5.757	6.375	.046	.013
	df	1	1	1	5	6	1	1
	p	.777	.890	.189	.331	.382	.829	.909
	N	107	107	107	107	107	106	106
Oral-B Professional Care 500 [MP2]	χ^2	.851	.159	.851	5.455	5.368	.456	.897
	df	1	1	1	5	6	1	1
	p	.356	.690	.356	.363	.498	.500	.343
	N	107	107	107	107	107	106	106
Oral-B Professional Care 3000 [MP2]	χ^2	.204	.151	.204	2.642	5.938	.950	.008
	df	1	1	1	5	6	1	1
	p	.651	.697	.651	.755	.430	.330	.929
	N	107	107	107	107	107	106	106
Johnnie Walker Red Label [MP2]	χ^2	1.424	.413	5.198	3.549	4.777	.247	.035
	df	1	1	1	5	6	1	1
	p	.001	.520	.023	.616	.573	.619	.852
	N	107	107	107	107	107	106	106
Johnnie Walker Black Label [MP2]	χ^2	6.157	2.645	1.348	3.040	11.895	.053	1.647
	df	1	1	1	5	6	1	1
	p	.013	.104	.246	.694	.064	.818	.199
	N	107	107	107	107	107	106	106
Intenso Memory Station 320GB [MP2]	χ^2	.851	1.470	.011	4.850	13.524	.941	.508
	df	1	1	1	5	6	1	1
	p	.356	.225	.918	.434	.035	.332	.476
	N	107	107	107	107	107	106	106
Intenso Memory Station 500GB [MP2]	χ^2	.006	.463	.892	6.747	13.223	.031	.334
	df	1	1	1	5	6	1	1
	p	.939	.496	.345	.240	.040	.859	.563
	N	107	107	107	107	107	106	106

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX300 [MP2]	χ^2	.048	.912	.048	1.593	6.569	1.486	.414
	df	1	1	1	5	6	1	1
	p	.826	.340	.826	.902	.363	.223	.520
	N	107	107	107	107	107	106	106
Sony MDR-ZX600 [MP2]	χ^2	.002	.149	.986	6.585	7.324	.208	.176
	df	1	1	1	5	6	1	1
	p	.961	.699	.321	.253	.292	.649	.675
	N	107	107	107	107	107	106	106

ms = mean split; mds = median split

19c) Structural identity test: Predicted buyer in BDM versus real buyer in DCT

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Oral-B Professional Care 500	χ^2	2.596	.268	.321	2.900	3.654	.782	.782
	df	1	1	1	4	4	1	1
	p	.107	.605	.571	.657*	.587*	.374	.376
	N	39	39	39	39	39	38	38
Oral-B Professional Care 3000	χ^2	.028	1.800	.281	2.250	1.041	.008	1.236
	df	1	1	1	3	2	1	1
	p	1.000*	.477*	1.000*	.706*	1.000*	1.000*	.515
	N	18	18	18	18	18	17	17
Johnnie Walker Red Label	χ^2	.091	.222	.000	4.978	2.391	.025	.252
	df	1	1	1	5	3	1	1
	p	.763	.637	.992	.512*	.525*	.875	.615
	N	53	53	53	53	53	53	53
Johnnie Walker Black Label	χ^2	3.972	.086	.793	3.349	3.006	2.232	2.738
	df	1	1	1	4	3	1	1
	p	.064*	.770	.373	.632*	.432*	.135	.098
	N	47	47	47	47	47	46	46
Intenso Memory Station 320GB	χ^2	.885	.650	.035	1.937	6.795	.694	3.232
	df	1	1	1	3	2	1	1
	p	.420*	.420	1.000*	.723*	.041*	.678*	.111*
	N	26	26	26	26	26	25	25
Intenso Memory Station 500GB	χ^2	1.893	1.167	1.422	4.144	6.288	.021	.088
	df	1	1	1	3	3	1	1
	p	.169	.280	.233	.312*	.088*	.885	.767
	N	42	42	42	42	42	41	42
Sony MDR-ZX300	χ^2	.063	.083	.083	4.298	3.868	3.886	2.268
	df	1	1	1	4	4	1	1
	p	.802	.894	.773	.431*	.414*	.049	.132
	N	39	39	39	39	39	38	38

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX600	χ^2	2.306	.267	2.162	3.671	4.939	1.824	.074
	df	1	1	1	3	3	1	1
	p	.129	.605	.141	.334*	.171*	.259*	.785
	N	31	31	31	31	31	30	30
Oral-B Professional Care 500 [MP2]	χ^2	.683	.316	.144	4.066	5.003	1.228	.264
	df	1	1	1	5	5	1	1
	p	.409	.574	.704	.618*	.477*	.268	.607
	N	62	62	62	62	62	61	61
Oral-B Professional Care 3000 [MP2]	χ^2	2.380	1.551	.336	5.779	1.140	.248	.589
	df	1	1	1	3	2	1	1
	p	.210*	.257*	.689*	.059*	.674*	.659*	.678*
	N	34	34	34	34	34	33	33
Johnnie Walker Red Label [MP2]	χ^2	.017	.054	.653	7.311	.256	.057	.052
	df	1	1	1	5	3	1	1
	p	.895	.817	.419	.223*	1.000*	.811	.820
	N	78	78	78	78	78	78	78
Johnnie Walker Black Label [MP2]	χ^2	2.045	.208	1.824	3.351	1.568	1.085	.837
	df	1	1	1	4	3	1	1
	p	.153	.648	.177	.639*	.815*	.298	.360
	N	67	67	67	67	67	66	66
Intenso Memory Station 320GB [MP2]	χ^2	1.073	.002	.441	4.456	6.796	1.416	1.416
	df	1	1	1	4	3	1	1
	p	.300	.965	.507	.297*	.068*	.234	.234
	N	54	54	54	54	54	53	53
Intenso Memory Station 500GB [MP2]	χ^2	.600	.133	.357	7.594	4.260	.095	.245
	df	1	1	1	3	3	1	1
	p	.439	.716	.550	.052*	.169*	.757	.621
	N	57	57	57	57	57	56	57

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX300 [MP2]	χ^2	.040	.519	.019	4.375	5.101	2.108	.611
	df	1	1	1	5	5	1	1
	p	.841	.471	.891	.546*	.385*	.147	.434
	N	65	65	65	65	65	64	64
Sony MDR-ZX600 [MP2]	χ^2	1.551	.185	3.656	4.506	4.446	1.172	.023
	df	1	1	1	3	3	1	1
	p	.274*	.667	.056	.209*	.219*	.442*	.880
	N	31	31	31	31	31	30	30

* Fisher's Exact Test.

ms = mean split; mds = median split

19d) Structural identity test: Predicted non-buyer in BDM versus real non-buyer in DCT

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Oral-B Professional Care 500	χ^2	1.244	.024	.214	3.749	2.474	.146	.058
	df	1	1	1	6	6	1	1
	p	.265	.877	.643	.711	.871	.703	.810
	N	168	168	168	168	168	167	167
Oral-B Professional Care 3000	χ^2	1.339	.038	.673	1.979	2.054	.670	.004
	df	1	1	1	5	5	1	1
	p	.247	.846	.412	.852	.842	.413	.950
	N	189	189	189	189	189	188	188
Johnnie Walker Red Label	χ^2	1.346	.234	2.105	3.833	3.311	.656	.026
	df	1	1	1	6	6	1	1
	p	.246	.628	.147	.699	.769	.418	.871
	N	154	154	154	154	154	154	154
Johnnie Walker Black Label	χ^2	.377	.096	.238	4.192	2.635	.297	.559
	df	1	1	1	6	6	1	1
	p	.539	.756	.626	.651	.853	.586	.455
	N	160	160	160	160	160	159	159
Intenso Memory Station 320GB	χ^2	.295	.455	.267	2.346	4.218	.022	.090
	df	1	1	1	6	6	1	1
	p	.587	.500	.605	.885	.647	.881	.764
	N	181	181	181	181	181	180	180
Intenso Memory Station 500GB	χ^2	.154	.084	.390	2.966	3.561	.521	.497
	df	1	1	1	6	6	1	1
	p	.695	.772	.532	.813	.736	.471	.481
	N	165	165	165	165	165	164	165
Sony MDR-ZX300	χ^2	.898	.024	1.168	2.941	3.034	.160	.060
	df	1	1	1	6	6	1	1
	p	.343	.877	.280	.816	.805	.689	.806
	N	168	168	168	168	168	167	167

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX600	χ^2	1.231	.021	1.451	2.963	3.165	.292	.296
	df	1	1	1	6	6	1	1
	p	.267	.885	.228	.814	.788	.589	.587
	N	176	176	176	176	176	175	175
Oral-B Professional Care 500 [MP2]	χ^2	.959	.176	.353	2.536	2.767	.624	1.172
	df	1	1	1	6	5	1	1
	p	.328	.675	.552	.864	.736	.429	.279
	N	145	145	145	145	145	144	144
Oral-B Professional Care 3000 [MP2]	χ^2	2.133	.121	.761	2.056	1.051	.998	.074
	df	1	1	1	5	4	1	1
	p	.144	.728	.383	.841	.902	.318	.785
	N	173	173	173	173	173	172	172
Johnnie Walker Red Label [MP2]	χ^2	.616	.389	.908	3.316	1.287	.549	.030
	df	1	1	1	6	5	1	1
	p	.433	.533	.341	.768	.936	.459	.862
	N	129	129	129	129	129	129	129
Johnnie Walker Black Label [MP2]	χ^2	.577	.741	.489	1.975	2.470	.320	.022
	df	1	1	1	5	5	1	1
	p	.447	.389	.484	.853	.781	.572	.881
	N	140	140	140	140	140	139	139
Intenso Memory Station 320GB [MP2]	χ^2	.150	.485	.077	2.144	5.490	.000	.018
	df	1	1	1	5	6	1	1
	p	.698	.486	.781	.829	.483	.986	.892
	N	153	153	153	153	153	152	152
Intenso Memory Station 500GB [MP2]	χ^2	.538	.110	.411	2.649	4.825	.372	.048
	df	1	1	1	6	6	1	1
	p	.463	.741	.522	.851	.566	.542	.826
	N	150	150	150	150	150	149	149

Product		Gender	Age (mds)	Body size (mds)	level of education of father	level of education of mother	Risk attitude (ms)	Lotteries played (ms)
Sony MDR-ZX300 [MP2]	χ^2	1.297	.087	.993	2.937	4.173	.006	.104
	df	1	1	1	6	6	1	1
	p	.255	.768	.319	.817	.653	.940	.747
	N	142	142	142	142	142	141	141
Sony MDR-ZX600 [MP2]	χ^2	1.518	.091	1.455	2.764	3.074	.457	.457
	df	1	1	1	6	6	1	1
	p	.218	.763	.228	.838	.799	.499	.499
	N	176	176	176	176	176	175	175

ms = mean split; mds = median split

20) WTP differences depending on buying motive (bargain seeking below vs. above the mean).

Product	Mean WTP (SD)		Difference		
	below mean	above mean	t	df	p
Oral-B Professional Care 500	7.21 (7.40)	12.54 (9.38)	3.282*	104.222	.001
Oral-B Professional Care 3000	9.46 (13.82)	18.92 (17.10)	3.162*	103.719	.002
Johnnie Walker Red Label	5.54 (4.45)	7.04 (4.75)	1.662	105	.100
Johnnie Walker Black Label	8.63 (8.14)	12.31 (9.65)	2.080	105	.040
Intenso Memory Station 320GB	15.00 (13.15)	17.79 (14.51)	1.022	105	.309
Intenso Memory Station 500GB	16.80 (15.65)	22.72 (17.28)	1.819	105	.072
Sony MDR-ZX300	9.06 (7.44)	12.47 (8.36)	2.180	105	.032
Sony MDR-ZX600	14.87 (12.73)	19.44 (13.76)	1.744	104	.084

* Welch's *t*-test

Otto von Guericke University Magdeburg
Faculty of Economics and Management
P.O. Box 4120 | 39016 Magdeburg | Germany

Tel.: +49 (0) 3 91 / 67-1 85 84
Fax: +49 (0) 3 91 / 67-1 21 20

www.fww.ovgu.de/femm

ISSN 1615-4274