

Lieven, Theo; Lennerts, Silke

## Article

# Measuring Willingness to Pay by Means of the Trade-off between Free Available Cash and Specific-Purpose Vouchers

BuR - Business Research

## Provided in Cooperation with:

VHB - Verband der Hochschullehrer für Betriebswirtschaft, German Academic Association of Business Research

*Suggested Citation:* Lieven, Theo; Lennerts, Silke (2013) : Measuring Willingness to Pay by Means of the Trade-off between Free Available Cash and Specific-Purpose Vouchers, BuR - Business Research, ISSN 1866-8658, VHB - Verband der Hochschullehrer für Betriebswirtschaft, German Academic Association of Business Research, Göttingen, Vol. 6, Iss. 2, pp. 154-171, <https://doi.org/10.1007/BF03342747>

This Version is available at:

<https://hdl.handle.net/10419/103724>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



# Measuring Willingness to Pay by Means of the Trade-off between Free Available Cash and Specific-Purpose Vouchers

Theo Lieven, Center for Customer Insight, University of St. Gallen, Switzerland, E-mail: theo.lieven@unisg.ch

Silke Lennerts, Center for Customer Insight, University of St. Gallen, Switzerland, E-mail: silke.lennerts@unisg.ch

## Abstract

Primarily because of the hypothetical character of interview situations, respondents are often unable to state their true willingness to pay (WTP). This inability results in the so-called hypothetical bias. To address this bias, incentive-compatible methods have been proposed, but such methods are applicable only to real products. We propose a new method for measuring WTP based on disposable cash and specific-purpose vouchers that are earmarked for particular goods that, contrary to incentive-compatible methods, can be used for hypothetical products and services that are less affected by the hypothetical bias. Empirical studies show that the new procedure provides WTP results that are equal to the results of alternative incentive-compatible elicitation procedures, such as the Vickrey auction and the Becker–DeGroot–Marschak (BDM) procedure.

JEL-classification: M31

Keywords: willingness to pay, hypothetical bias, incentive compatibility, stated preferences

Manuscript received August 7, 2012, accepted by Sönke Albers (Marketing) July 6, 2013.

## 1 Introduction

A considerable percentage of marketing effort is expended to elicit consumers' willingness to pay (WTP). If the actual price of the good is less than or equal to this WTP or the so-called reservation price, a consumer will demand one unit of the good (Varian 1992). However, the results from surveys in which the WTP is elicited may not reflect real values. Direct methods in which consumers are asked to explicitly state their WTP by means of open-ended questioning are admittedly easy for survey participants to follow and understand. However, particularly because of the hypothetical interview situation, respondents are often unable to state their WTP and instead provide a price that comes to mind more easily, such as the price of a competitive product or service (Blatter, Miller, Hofstetter, and Krohmer 2009). In these cases, we encounter the so-called hypothetical bias (Ajzen, Brown, and Car-

vajal 2004; Murphy, Allen, Stevens, and Weatherhead 2005). The hypothetical bias results in an overestimation of WTP (Barrot, Albers, Skiera, and Schäfers 2010; Völckner 2006). Little and Berrens (2004) and List and Gallet (2001) found the stated WTP to be as much as three times higher than the actual WTP. Thus, inferences that are made from WTP elicitation can be extremely misleading. Thus far, researchers and practitioners have made several attempts to avoid the resulting hypothetical bias. So-called incentive-compatible auctions, such as the Vickrey auction (Vickrey 1961) and the BDM procedure (Becker, DeGroot, and Marschak 1964), have been proposed but require real products that must be sold and real money, as the winners are obliged to pay for the products. With other methods, consumers make choices from a set of several product/service/price combinations (conjoint analysis, CA). All of these methods are rather complicated.



Enhancements that are based on recent research have made the procedures both more realistic and more complex.

This paper proposes a new method that overcomes some of the shortcomings of the other methods and can thus be considered an alternative. This new method has proven to be as effective as the existing valid techniques. However, with only two monetary attributes, the new procedure is less complex and therefore more easily applicable. The new method is derived from the principles of CA; however, the data are collected in a much simpler setting, with only two attributes: cash and voucher. The procedure is hypothetical and does not require a realistic setting with real products or real money. Because of its hypothetical nature, there is potential for the outcome to yield unrealistic results. However, although the method is not incentive compatible, there is a key difference from the open-ended question method. To answer an open-ended question requires only the pointing of a finger at an amount that is assumed to suitably represent WTP. By contrast, solving C/V trade-offs requires cognitive skills. Consequently, the results may be more realistic than those derived from an open-ended question.

The remainder of this article is structured as follows. First, we present a literature review that focuses on alternative WTP elicitation methods. Subsequently, the underlying principles of the proposed cash/voucher-WTP elicitation method (C/V method) are described. Next, we discuss the validity of this procedure on a theoretical basis. Thereafter, the viability of the method is analyzed empirically in an initial study by testing the criterion validity against the Vickrey and BDM procedures. A second study analyzes the effect of different sets of cash/voucher (C/V) alternatives in both within- and between-subject designs. A third study compares the results from two C/V methods covering different price ranges. Furthermore, the outcomes of two different stated preference methods – one ranking and one choice procedure – are analyzed. Finally, the results are summarized, and limitations, implications, and further research are discussed.

## 2 Review of the Literature on WTP Elicitation Methods

Several WTP elicitation methods have been proposed to avoid or mitigate the hypothetical bias. With indirect questioning, such as CA (Green and

Rao 1971; Green and Srinivasan 1978; for a complete overview see Teichert and Shehu 2010), the price is not the direct focal point, as it is ranked together with various product attributes; therefore, it is difficult for a respondent to infer the consequences of the provided preferences. Researchers have identified the lack of an incentive to answer relevant questions truthfully as the main reason for WTP overstatement (Dasgupta, Hammond, and Maskin 1979). Because there are no consequences, such as an obligation to buy a good, not taking sufficient care when answering or thoughtlessness may result in incorrect WTP statements. Therefore, procedures have been designed to develop incentive-compatible WTP assessments. Techniques such as the Vickrey auction (Vickrey 1961), first-price sealed-bid auctions (Milgrom and Weber 1982), or a comparison of a bid and a random price in the BDM mechanism at least theoretically guarantee incentive compatibility (Becker, DeGroot, and Marschak 1964; Miller, Hofstetter, Krohmer, and Zhang 2011). With all of these techniques, the survey participant is (or may be) obliged to buy a good, thus making him or her more aware and truthful. Incentive-compatible procedures have frequently been validated and have been approved as more closely reflecting real WTP (Miller, Hofstetter, Krohmer, and Zhang 2011; Wertenbroch and Skiera 2002). Recently, several combinations of elicitation techniques have enhanced existing procedures with incentive-compatibility components. A combination of CA and BDM was refined over several steps (Ding 2007; Ding, Grewal, and Liechty 2005; Dong, Ding, and Huber 2010). Considering the worldwide proliferation of Internet and online auctions, Barrot, Albers, Skiera, and Schäfers (2010) executed online auctions with the second-price sealed bid procedure and found them to be better suited for estimating realistic price-demand functions. Another approach to web-based CA has been proposed by Park, Ding, and Rao (2008), who employed a combination that incorporates a BDM procedure in which respondents can upgrade complex products and state their WTP for this improvement. The most recent developments no longer treat WTP as a price point but treat it as a range (Dost and Wilken 2012; Wang, Venkathesh, and Chatterjee 2007). Because these methods are designed to be incentive compatible, they have been combined with a BDM procedure. The aim of all of these procedures is to approximate true WTP values and to obtain a realistic price–



demand curve. However, all of these procedures have disadvantages. For example, CA is a strong tool for product designers. However, price is not the only attribute, and when comparing the part-worths of several characteristics (for example, the engine power, color, and price of a car), different and incompatible measuring scales compete with one another, particularly when they are not aligned. This conflict may result in cognitive dissonance, and individuals may refrain from making any choice at all or proceed to simple heuristics (Herrmann, Heitmann, Morgan, Henneberg, and Landwehr 2009). Furthermore, the choice of attributes and the number of attribute variations have a particularly strong influence on the measurement of price reactions and price-response functions (Green, Helsen, and Shandler 1988; Wittink, Krishnamurthi, and Reibstein 1989). Measures to reduce the number of choices in CA, such as the “Latin Square,” are not unrestrictedly accepted and appear to be – if not arbitrary – at least heuristic (Hamlin 2005). Therefore, it may be problematic to induce WTP from a CA.

Auctions such as Vickrey or BDM have different shortcomings. In contrast to real-life purchase situations, it is the bidders (not the products) who compete at an auction or lottery. This problem can lead to behavior that diverges from a realistic buying situation. Because of the complexity of these techniques, respondents are also often cognitively over-extended (Müller, Voigt, and Erichson 2009). Furthermore, to provide a credible incentive, a real-world product must be presented. Because the winner of an auction must pay the price of this good, he or she must be in possession of a means of payment (e.g., cash, credit card). Consequently, most of the research implementing Vickrey or BDM auctions has been conducted with inexpensive items, such as cans of coke (Wertenbroch and Skiera 2002). A solution for this problem is the implementation of such surveys in e-commerce, in which transactions with higher prices are common (Barrot, Albers, Skiera, and Schäfers 2010).

Two other properties of auctions may violate the validity of the elicited WTP values. The incentive compatibility becomes effective only in the upper price range. A participant who is not interested in buying a product may bid any price up to the mid-range to be “safe” from winning. Thus, only a few participants will bid zero. This bidding behavior may distort the shape of the price–demand curve.

Furthermore, those who would never buy the good do not participate in an auction. Thus, in situations in which some survey participants are not obliged to bid, a self-selection bias could exist (Heckman 1990), preventing results from being generalizable. Summarizing the overview above, several valid methods may be used to accurately elicit WTP. However, methods that are increasingly modified to be as close to reality as possible tend to be more complicated. The contribution of this article is to present a new method with reduced complexity that does not need real products and real money.

### 3 The Basic Principle of the C/V Method

Because a voucher’s value is “not more than, and may be less than the value of an equal-dollar cash transfer” (Moffitt 1989), a voucher’s face value is the upper limit of its monetary value. The sum of a cash amount of €5.50 plus a voucher with a face value of €2.40 is worth a maximum of €7.90. This sum is inferior to a combination of €6.00 cash and a voucher of €2.00, which equals €8.00. Thus, when the cash amount in one alternative is lower than in another, this alternative is inferior if the sum of both cash and voucher is also lower. However, in cases in which the cash amount is lower than in the other option but the sum of cash and voucher is not, the higher voucher amount could eventually compensate for the lower cash amount. This compensation depends on the WTP for a voucher’s underlying good. When the WTP is equal to zero, the voucher has no monetary value. However, as soon as the WTP is higher than zero, the redemption of the voucher frees up cash money.

The preference for a certain alternative depends on the remaining cash after spending an amount that is equivalent to the WTP for a good. Assume that a consumer has a WTP of €2.00 for a specific commodity. Regarding a combination of €6.00 in cash and €2.00 as a voucher for this good (alternative 1) and another combination of €5.50 cash and a €4.00 voucher (alternative 2), the voucher will be redeemed in both cases. Either €2.00 of the voucher in alternative 2 forfeits or – in the case in which the good can be acquired for only €4 – €2 of the voucher is wasted. In any case, €5.50 remain as cash. In alternative 1, €6.00 remain, which gives this alternative a higher preference. However, as soon as the WTP reaches €2.50, both alternatives are

**Table 1: Dominances and equivalences depending on four conditions**

	Conditions	Dominance/ Equivalence	Examples					
			Alternative 1		Alternative 2		WTP in cases of equivalence	Any WTP higher than in cases of equivalence
			Cash1	Voucher1	Cash2	Voucher2		
Cash2 v Cash1	Cash2 + Voucher2 < Cash1 + Voucher1	A2 < A1	€6.00	€2.00	€5.50	€2.40	n.a.	n.a.
	Cash2 + Voucher2 ≥ Cash1 + Voucher1	WTP of equivalence = Cash1 + Voucher1 - Cash2	€6.00	€2.00	€5.50	€4.00	€2.50	A2 > A1
Cash2 ^ Cash1	Cash2 + Voucher2 > Cash1 + Voucher1	A2 > A1	€6.00	€2.00	€7.00	€1.50	n.a.	n.a.
	Cash2 + Voucher2 ≤ Cash1 + Voucher1	WTP of equivalence = Cash2 + Voucher2 - Cash1	€6.00	€2.00	€7.00	€.50	€1.50	A1 > A2

**Table 2: Principle WTP values in which two out of four C/V alternatives are equivalent**

	Alternative 1		Alternative 2		Alternative 3		Alternative 4	
	Cash	Voucher	Cash	Voucher	Cash	Voucher	Cash	Voucher
		4.50	12.20	6.00	10.50	8.50	6.50	9.00
	<b>Cash</b>	<b>Voucher</b>	<b>↓Sum→</b>	16.70	16.50	15.00	10.50	
Alternative 1	4.50	12.20	16.70	-	<b>12.00</b>	<b>10.50</b>	<b>6.00</b>	
Alternative 2	6.00	10.50	16.50		-	<b>9.00</b>	<b>4.50</b>	
Alternative 3	8.50	6.50	15.00			-	<b>2.00</b>	
Alternative 4	9.00	1.50	10.50					-

Principle WTP values are in bold.

equivalent. In alternative 2, again, €5.50 remain, and in alternative 1 – because a consumer with a WTP of €2.50 will spend extra cash in the amount of €0.50, which must be deducted from the €6.00 – €5.50 also remain. Thus, for any WTP that is higher than €2.50, alternative 2 is superior. Where two alternatives are equivalent, we call these amounts principle WTP values because they are derived logically from the behavior of a rational decision maker. The following equation is used to calculate the point of equivalence:

$$\text{Principle WTP}_{1-2} = \text{Cash}_{\text{Alternative 1}} + \text{Voucher}_{\text{Alternative 1}} - \text{Cash}_{\text{Alternative 2}} \quad (1)$$

with alternative 1 having a higher cash amount than alternative 2 but with alternative 2 having a higher sum of both cash and voucher. For the described example above, this sum equates to €6.00 + €2.00 – €5.50 = €2.50. The rules for all conditions are shown in Table 1.

The above choice between alternative 1 and alternative 2 resulted in one principle WTP. When alternative 3 was added, two more WTP values emerged: one between alternative 1 and 3 and the other between 2 and 3. Table 2 shows an example with four alternatives resulting in six principle WTP values. These alternatives have been assessed in pairwise comparisons by implementing the rules from Table 1 and equation (1), respectively (e.g., between alternative 1 and 4: 9.00 + 1.50 - 4.50 = 6.00).

For each of the depicted principle WTP values, the two respective alternatives are equivalent. After a ranking procedure of the four alternatives at these points, two of the preference scores swapped positions. In cases in which a participant is not interested in a voucher's underlying good, the ranking of the combinations in Table 2 suggests that alternative 4 is number one because of the highest cash amount and that the lowest rank is for alternative 1, which has the smallest cash amount; thus, the result is a



**Table 3: The seven principle rankings for four C/V alternatives depending on principle WTP values in which ranking scores swap pairwise**

	Consumers' Willingness to Pay (principle WTP)								
	Cash	Voucher	€0.00	€2.00	€4.50	€6.00	€9.00	€10.50	€12.00
Alternative 1	4.50	12.20	4.	4.	4.	<b>3.</b>	3.	<b>2.</b>	<b>1.</b>
Alternative 2	6.00	10.50	3.	3.	<b>2.</b>	2.	<b>1.</b>	1.	<b>2.</b>
Alternative 3	8.50	6.50	2.	<b>1.</b>	1.	1.	<b>2.</b>	<b>3.</b>	3.
Alternative 4	9.00	1.50	1.	<b>2.</b>	<b>3.</b>	<b>4.</b>	4.	4.	4.

4 = highest preference, 1 = lowest. Swapping scores are bold.

ranking of 4. – 3. – 2. – 1. However, from a WTP of €2.00 upward, alternative 3 becomes superior, and the ranking changes to 4. – 3. – 1. – 2.

However, with four scores, there exist a total of  $4! = 24$  possible ranking orders. More information could be elicited from consumers when they are not constrained to choose one out of only seven of the principle alternatives above. Additionally, more WTP values could fill the gaps between the existing alternatives. As a solution for the WTP calculation of other rankings (e.g., 2. – 3. – 4. – 1.), the part-worths of the cash and voucher amounts are determined following the methods of CA. The scores of each alternative are regressed on their specific cash and voucher amounts ( $score_i = b_1 \cdot cash_i + b_2 \cdot voucher_i; i = 1, \dots, 4$ ).

The differences between the cash and voucher amounts from one alternative to another are far from being equidistant. The steps follow an ordered monotonic pattern; however, it is not linear. From alternative  $x = 1$  to 4, the cash amounts can be calculated by cubic polynomials ( $cash = -.5x^3 + 3.5x^2 - 5.5x + 7$ ;  $voucher = .217x^3 - 2.45x^2 + 4.133x + 10.3$ ). Consequently, linear regression is not a suitable method. Thus, a MONANOVA procedure (ordinal CA; De Leeuw, Young, and Takane 1976) for ordered dependent variables is applied instead. The regression procedure results in a coefficient  $b_1$  for cash and coefficient  $b_2$  for the voucher. Because each of the two coefficients is an expression of the contribution of cash and voucher, the ratio between the voucher and the cash coefficient ( $b_2/b_1$ ) represents the voucher's cash equivalent (CE) as a percentage. After assigning the respective CEs to the seven known alternatives above, the unknown WTP values can be calculated by an interpolation procedure between the CEs. Because no differentiability of the

interpolation function is required, no complicated approximation with polynomials or splines is applied; rather, the much simpler linear interpolation  $WTP_{unknown} = WTP_{lower} + (CE_{Ranking \text{ with unknown WTP}} - CE_{lower}) / (CE_{higher} - CE_{lower}) \times (WTP_{higher} - WTP_{lower})$  is employed.

Based on the four combinations in Table 2, the MONANOVA regressions resulted in a CE of .4779 for the ranking 3. – 2. – 1. – 4., with a principle WTP of €6.00 (Table 3) and a CE of .6244 for the next higher ranking 3. – 1. – 2. – 4., with a principle WTP of €9.00. The ranking 2. – 3. – 4. – 1. resulted in a CE of .5221; thus, its WTP is equal to  $€6.00 + (.5221 - .4779) / (.6244 - .4779) \times (€9.00 - €6.00) = €6.91$ .

#### 4 The Ecological Validity of the C/V Method

Although the concept of ecological validity was coined by Brunswik and Kamiya (1953) in a "slightly different" context (Neisser 1976: 48), the term has been increasingly used since the 1970s to describe the generalizability of experimental results to reality. In our case, the WTP results are ecologically valid if the respondents would actually buy products that are less expensive, as their stated WTP indicated. It is a widespread belief that hypothetical biases prevent survey results from being ecologically valid. However, solutions to hypothetical tasks may nevertheless reflect reality. Bornstein (1999) analyzed results from simulated juries in court settings and found that feasible generalizations regarding the behavior of real jurors could be made from simulation studies. Bateson and Hui (1992) reported the ecological validity of photographic slides and videotapes in the service setting. Rogers, Kadar, and Costall (2005) reported the ecological validity of



surveys with car simulators. It is not even necessary for a simulation to be a high-fidelity simulation. Low-fidelity environments, such as paper and pencil tasks, also result in realistic behavior (Motowidlo, Dunette, and Carter 1990). The rationale behind this approach is that “when people are asked what they would do if they were in a hypothetical work situation, they may try to recall how they behaved in similar situations in the past” (Motowidlo, Dunette, and Carter 1990). This rationale holds even for situations that participants have never experienced before and results from their ability to extrapolate from other situations.

Based on the above considerations, the measures in a hypothetical environment do not necessarily need to be ecologically invalid. The setting of the questions could be designed such that respondents’ answers converge with their real behaviors. Although a hypothetical bias is unlikely to be completely eliminated in a non-real situation, it could eventually be remedied. The answer to an open-ended question is given with a mere pointing of the finger at a probable amount of WTP with little contemplation. In the C/V procedure, however, some thought is required. First, respondents must understand the task. Such a task need not be elaborate because individuals are likely to have previously experienced situations in which they received vouchers as a gift. The recognition of the trade-off follows. Most importantly, respondents will learn about the extremes of the options: either cash only or less cash with a large voucher. If participants are willing to solve this trade-off, then they will be obligated to use their minds. This feature distinguishes the C/V procedure from a mere bidding scenario and may result in greater ecological validity.

For cases in which the ecological validity of a method is in question, Raser (1969) suggested comparing the results of hypothetical situations with the outcomes of reconstructed known situations. This approach was used for this article through the comparison of five different WTP elicitation methods, including two C/V procedures, one open-ended question, a Vickrey auction, and a BDM procedure. The last two methods are well-known methods that are assumed to be incentive compatible and thus are able to predict WTP values that are fairly close to “real” WTP values.

## 5 Empirical Study 1

This study analyzes the effect of different C/V alternatives as well as the internal and external validity of the C/V procedure.

### 5.1 Propositions and expectations

First, the effect of different sets of C/V combinations should be analyzed. When comparing two sets – for example, one with principle WTP values of €0, €2, €3, €4, €5, €9, and €12 and another with values of €0, €3, €6, €8, €9, €10, and €12 – will the average WTP or the distribution of WTP be the same? Furthermore, what is the effect of the number of alternatives? A MONANOVA with two independent variables requires at least four observations, that is, C/V combinations. Six alternatives appear to be the maximum that are manageable with 720 possible rankings (6!). With seven alternatives, there would be 5,040 different rankings, which would be difficult to handle. Consequently, in a within-subject design, we will compare results from a set of four alternatives and a different set of six alternatives, and we expect the following proposition to hold:

P<sub>1</sub>: In a within-subject design, C/V procedures with four alternatives and with six alternatives will result in similar WTP values.

The WTP results following open-ended elicitation methods are known to be overestimated because of the hypothetical bias. As a necessary condition, the results of the two C/V surveys should thus be significantly lower than the WTP results that are derived from an open-ended question:

P<sub>2</sub>: The WTP values that result from a C/V method will be lower than the values that result from answers to an open-ended question.

Studies applying Vickrey auctions and BDM procedures provide support for the assumption of ecological validity; thus, these methods are assumed to be capable of predicting real WTP. To test whether the new C/V method can also be assumed to be ecologically valid, a Vickrey auction and a BDM survey will be conducted with the same commodity as basis for comparison.

P<sub>3</sub>: The predicted WTP values from C/V procedures will be similar to the results from a Vickrey auction and a BDM procedure.

**Table 4: Principle WTP values in which two out of six C/V alternatives are equivalent**

	Alternative 2		Alternative 3		Alternative 4		Alternative 5		Alternative 6	
	Cash	Voucher	Cash	Voucher	Cash	Voucher	Cash	Voucher	Cash	Voucher
	4.90	11.00	5.50	9.40	6.25	7.00	7.00	4.00	7.50	0.00
	<b>Cash Voucher ↓Sum →</b>		15.90	14.90	13.25		11.00		7.50	
Alternative 1	3.90	12.25	16.15	<b>12.00</b>	<b>11.00</b>	<b>9.35</b>	<b>7.10</b>	<b>3.60</b>		
Alternative 2	4.90	11.00	15.90	-	<b>10.00</b>	<b>8.35</b>	<b>6.10</b>	<b>2.60</b>		
Alternative 3	5.50	9.40	14.90		-	<b>7.75</b>	<b>5.50</b>	<b>2.00</b>		
Alternative 4	6.25	7.00	13.25			-	<b>4.75</b>	<b>1.25</b>		
Alternative 5	7.00	4.00	11.00				-	<b>0.50</b>		
Alternative 6	7.50	0.00	7.50						-	

*Principle WTP values are in bold.*

## 5.2 Design, sample, and procedure

Five surveys have been conducted to identify differences in the results of various elicitation procedures. These surveys included the open-ended question, two measurements that use the above-described ranking procedure (one with four C/V alternatives and one with six), a Vickrey auction, and a BDM procedure. The commodity for which the WTP had to be elicited consisted of a pack of four cans of chocolate sandwich cookies (total weight 460 grams = 16.23 oz.). This offer was the same in all five surveys. The studies were performed in Switzerland, and this product was chosen because it was not available in local stores and was thus expected to be unknown to the survey participants. Thus, there was no reference price as an anchor, and participants had to make their own decisions regarding their WTP. Comparable retail prices in other European countries were in the range of €8 to €10 (CHF 9.60 to CHF 12.00). The survey participants were undergraduates from a northeastern Swiss university.

**Open-ended question.** The situation was hypothetical. The participants were given the following instruction: “Please state the price that you would be willing to pay to acquire this set of four cans with chocolate sandwich cookies.”

**C/V alternatives.** Similar to the open-ended procedure, the situation was hypothetical, and the participants were not obliged to purchase the cookies. However, preferences needed to be stated for the alternative combinations of cash and voucher, the latter of which was earmarked exclusively for the

chocolate sandwich cookies. The hypothetical situation could be a Christmas or anniversary gift. The participants were informed that the vouchers were not transferable and that they could be redeemed at their face value when purchasing the cookies. If the voucher’s value was not sufficient, then the remaining amount needed to be paid in cash, and if the voucher’s value was more than the sales price of the cookies, then the remaining amount of the voucher was forfeited. The above-described procedure was applied in one version with four C/V alternatives, and to compare the results, the procedure in a version with six alternatives. Because only a single questionnaire was used, the order of the two tasks was not randomized across participants.

**Four C/V alternatives.** The four alternatives described in Table 2 were chosen. These alternatives offer a range of WTP values from 0.00 to 12.00. The participants were asked to rank the alternatives according to their preferences, with 1 as the first choice and 4 as the last choice.

**Six C/V alternatives.** Another set of alternatives with six combinations of cash and voucher was chosen. The cash and voucher amounts as well as the respective equivalent principle WTP values are summarized in Table 4. These values were calculated in the same manner as that used for the procedure with four alternatives, by comparing two of them and adding the sum of cash and voucher from the couple with the higher cash portion and deducting the cash amount from the other (e.g., alternative 1 and 6:  $7.50 + 0.00 - 3.90 = 3.60$ ).



**Table 5: Ranking scores swapped in a pairwise fashion**

	Consumers' Willingness to Pay																	
	Cash	Voucher	0.00	0.50	1.25	2.00	2.60	3.60	4.75	5.50	6.10	7.10	7.75	8.35	9.35	10.00	11.00	12.00
Alternative 1	3.90	12.25	6.	6.	6.	6.	6.	<b>5.</b>	5.	5.	5.	<b>4.</b>	4.	4.	<b>3.</b>	3.	<b>2.</b>	<b>1.</b>
Alternative 2	4.90	11.00	5.	5.	5.	5.	<b>4.</b>	4.	4.	4.	<b>3.</b>	3.	3.	<b>2.</b>	2.	<b>1.</b>	1.	<b>2.</b>
Alternative 3	5.50	9.40	4.	4.	4.	<b>3.</b>	3.	3.	3.	<b>2.</b>	2.	2.	<b>1.</b>	1.	1.	<b>2.</b>	<b>3.</b>	3.
Alternative 4	6.25	7.00	3.	3.	<b>2.</b>	2.	2.	2.	<b>1.</b>	1.	1.	1.	<b>2.</b>	<b>3.</b>	<b>4.</b>	4.	4.	4.
Alternative 5	7.00	4.00	2.	<b>1.</b>	1.	1.	1.	1.	<b>2.</b>	<b>3.</b>	<b>4.</b>	<b>5.</b>	5.	5.	5.	5.	5.	5.
Alternative 6	7.50	0.00	1.	<b>2.</b>	<b>3.</b>	<b>4.</b>	<b>5.</b>	<b>6.</b>	6.	6.	6.	6.	6.	6.	6.	6.	6.	6.

6 = highest preference, 1 = lowest. Swapping scores are in bold

With six alternatives, 16 such equivalent points exist, with their respective principle rankings. These rankings are depicted in Table 5.

**Vickrey auction.** When a Vickrey auction was undertaken, the situation was no longer hypothetical. All of the undergraduates were familiar with this technique because of a preceding pricing lecture. The participants were instructed to bid for the cookies secretly by writing their WTP on paper and sealing it in an envelope. They were aware that the participant with the highest bid was obliged to buy the product, paying the price of the second-highest bid in cash immediately after the auction.

**BDM procedure.** Again, the participants were aware that it was a real situation in which all bidders had to purchase the cookies if their WTP values were above or equal to a minimum price, which was to be unveiled after the bidding process. The price that they needed to pay would be this minimum price, regardless of what the participants had bid. Those bidders whose WTP values were below that price were not eligible to acquire the cookies.

### 5.3 Results

**Open-ended question.** In total, 111 undergraduate students participated (36.1% female,  $M_{Age} = 23.4$ ,  $SD_{Age} = 2.4$ ). The average WTP was 7.71, with a standard deviation of 3.39 (for better readability, CHF as the sign for Swiss francs is omitted). The lowest WTP was 0.00 ( $n = 3$ ), and the highest was 20.00 ( $n = 1$ ). Four modes existed, one at 5.00 ( $n = 12$ ), one at 6.00 ( $n = 10$ ), one at 8.00 ( $n = 18$ ), and one at 10.00 ( $n = 14$ ).

**Four C/V alternatives.** Sixty-one undergraduates participated in this survey (39.3% female,  $M_{Age} =$

23.1,  $SD_{Age} = 2.4$ ). Their rankings were assigned to the WTP values according to the procedure that is described above. The average was 5.00, with a standard deviation of 4.09. The maximum was 12.00 ( $n = 6$ ), and the minimum was 0.00 ( $n = 14$ ) as a strong mode.

**Six C/V alternatives.** The same 61 undergraduates who participated in the procedure with four alternatives also participated in this survey. Their rankings were assigned to the WTP. The average WTP was 4.99, with a standard deviation of 4.23. The maximum was 12.00 ( $n = 4$ ), and the minimum was 0.00 ( $n = 13$ ) as a mode.

Both C/V surveys were conducted in a within-subject design (i.e., the respondents ranked the alternatives for the versions with four alternatives and with six alternatives). In such a one-group pre-test-posttest design, there was a threat of biasing testing effects (by learning, by order, or by both; see Campbell and Stanley 1963). To estimate such a risk, the second survey of the subsequent study 2 was conducted with a between-subject design. In this case, no differences between both respondent groups were observed.

**Vickrey auction.** Fifty undergraduates participated in the auction (31.9% female,  $M_{Age} = 23.8$ ,  $SD_{Age} = 2.3$ ). The highest bid was 11.50. The winner acquired the cookies and paid 10.50, which was the second-highest bid. The average bid was 5.03, with a standard deviation of 2.89. The minimum price was 0.00 ( $n = 3$ ). All of the participants answered correctly when they were asked how much the winner had to pay (i.e., the price of the second-highest bid). Additionally, the participants were asked whether they had placed a low bid if they did not

**Table 6: Average WTP results and significances of pairwise T-tests**

	Open-ended	Vickrey auction	BDM	4 C/V alternatives	6 C/V alternatives
Average WTP →	7.71	5.03	4.61	5.00	4.99
Open-ended question	-	$p < .001$	$p < .001$	$p < .001$	$p < .001$
Vickrey		-	$p = .414$	$p = .953$	$p = .946$
BDM			-	$p = .534$	$p = .551$
4 C/V alternatives				-	$p = .973$

want to become the winner (1 = do not agree at all, 5 = completely agree) and asked to indicate their overall interest in the cookies (1 = no interest, 5 = strong interest). The answers were significantly correlated with the WTP values ( $r = -.508, p < .001$ , and  $.481, p < .001$ , respectively), which supports the evidence of sufficient face validity.

**BDM procedure.** Fifty-six undergraduates participated in this procedure (35.7% female,  $M_{Age} = 23.0$ ,  $SD_{Age} = 2.5$ ). Seven bidders offered more than the minimum price of 7.50 (these bids were 7.55, 7.90, 8.00, 8.90, 9.00, 9.90, and 10.00). The cookies were given to these participants, and 7.50 was collected from each person. The distribution had a minimum of 0.00 ( $n = 3$ ) and a maximum of 10.00 ( $n = 1$ ). The average WTP was 4.61, with a standard deviation of 2.44. Regarding the two questions concerning face validity, the answers were again significantly correlated with the WTP values ( $r = -.627, p < .001$ , and  $r = .331, p < .050$ , respectively).

An ANOVA for all five procedures showed that the method had a significant effect on the WTP values ( $F(4,334) = 12.177, p < .001$ ). However, when the open-ended method was removed, no such effect emerged ( $F(3,224) = .176, p = .913$ ). Subsequently, the results of the two C/V methods were compared. The procedure with four alternatives resulted in an average WTP of 5.00, and the version with six alternatives resulted in an average WTP of 4.99. The assumption that these versions were equal could not be rejected ( $p = .973$ ), and  $P_1$  is thus supported. Because this was a paired sample  $T$ -test with a standard error of only .213, the confidence interval (95%) of the differences between the two versions ranged from  $-.42$  to  $.43$ , which was quite low. This result supports the assumption that the respondents were able to solve the subsequent tasks with four and six C/V alternatives consistently, which gives the procedure internal and external validity.

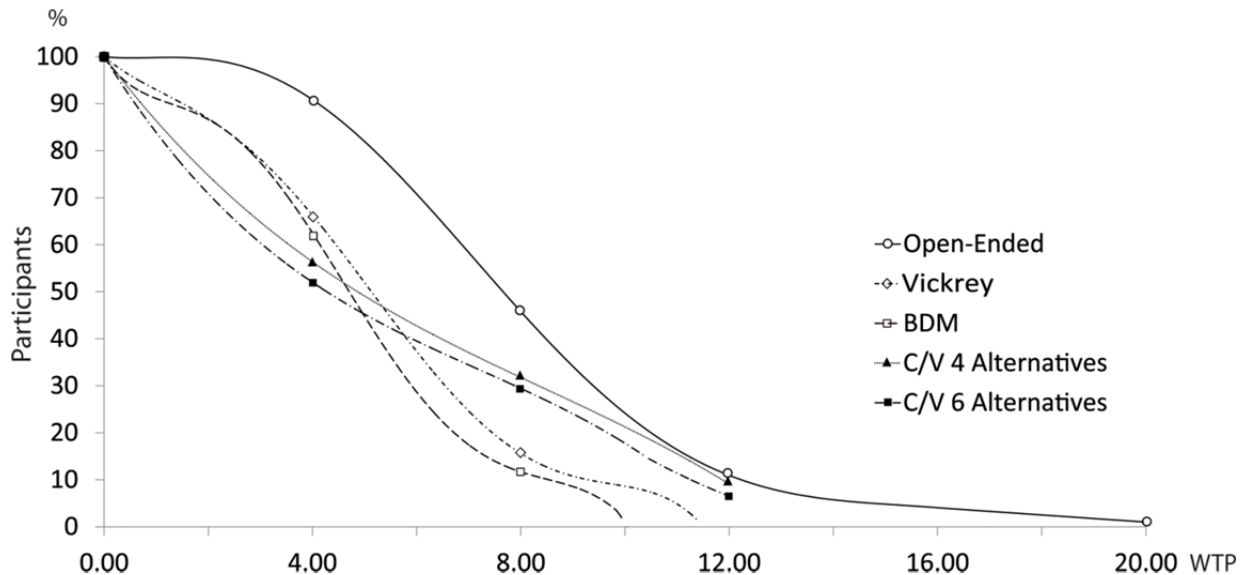
The BDM result is slightly lower than the result from the Vickrey auction (4.61 vs. 5.03). However, an assumption that the results were equal could not be rejected ( $p = .414$ ). Most importantly, the assumption that the results from the cash and voucher procedures did not differ from those of the Vickrey auction and the BDM elicitation could not be rejected. However, the result of the open-ended procedure (7.71) was significantly higher than the results of the other four procedures. All significances for the equality hypotheses can be found in Table 6. As a result, the C/V procedure was found to be a suitable alternative to the widely accepted Vickrey auction and the BDM procedure. Thus,  $P_2$  and  $P_3$  are also supported.

Given the C/V data, there could be concerns regarding the normality assumption of the distributions, which is necessary for the  $T$ -tests presented above. The density functions appear rather U-shaped, with two modes at the ends of the scale from 0.00 to 12.00. Although a Kolmogorov-Smirnov test did not reject the hypotheses of normality for all five distributions ( $ps > .06$ ), an additional nonparametric analysis was conducted. A Kruskal-Wallis test rejected the hypothesis that all five samples came from the same population ( $\chi^2(4) = 42.893, p < .001$ ). However, when the open-ended procedure was removed, the assumption that the two C/V procedures, the Vickrey auction, and the BDM procedure came from the same population could not be rejected ( $\chi^2(4) = .522, p < .914$ ). This finding again supports the above propositions.

#### 5.4 Discussion

The results showed standard deviations below 3.00 for the BDM and Vickrey procedures. This result may have been observed because the survey participants tended to bid in the middle range, which served as an anchor. During auctions, even if a

**Figure 1: The demand functions of all five WTP elicitation procedures**



consumer is not willing to acquire a good at all, there is no incentive for a total rejection of an offer because bids in the midrange are sufficiently safe to prevent one from winning. Consequently, there is no need to bid zero. This behavior is a disadvantage, at least in the Vickrey auction and perhaps also in the BDM procedure. Only three participants offered 0.00 each in the Vickrey and BDM procedures, whereas there were 14 such bids in the version with four C/V alternatives and 13 in the version with six alternatives. Similarly, only three respondents bid at least 10.00 in the Vickrey auction, and only one individual offered 10.00 in the BDM session. In the C/V elicitations (the version with six alternatives and four alternatives), there were six and four participants, respectively, who chose a ranking equivalent to 12.00. As a result, the WTP distributions in the C/V procedures have higher standard deviations (approximately 4.00) and are flatter than the traditional procedures. Thus, the demand functions of the Vickrey auction and the BDM procedure have a steeper middle part because of the accumulation of bids with midrange prices (Figure 1). The results show rather logistic functions, with higher price elasticity in the middle part; that is, small changes in the mid-price range produce high demand leaps. In contrast, the C/V curves have the opposite shape, in which the mid-price range is rather inelastic. This type of a kinked demand curve is also called the Gutenberg demand function, and it was initially introduced by [Sweezy \(1939\)](#) to describe the prevail-

ing behavior in oligopolies or imperfect markets ([Barrot, Albers, Skiera, and Schäfers 2010](#); [Simon 1969](#)).

To the marketing practitioner, the average WTP does not adequately explain consumer behavior. An average WTP concerns the average price that an average customer would pay. However, in contexts of differentiation and price discrimination, more information about market segments is required. Therefore, price-demand curves are critical to marketing planners attempting to determine sales prices that maximize expected profits. Thus, the shape of this function is crucial. Figure 1 shows that the C/V curves predict a market share of 30% and higher at a sales price of 8.00. In contrast, the Vickrey and BDM methods predict a market share of only 15%, which may be deemed to be insufficient, thus triggering marketers to reduce prices. In the case of viable demand curves for the C/V method, this reduction would result in an unnecessary waste of profit.

The standard deviations in the Vickrey auction and the BDM method were rather small (below 3.00) because of the many WTP values in the mid-price range. From this finding, one could conclude that an average consumer will buy a product anyway, as soon as the price falls below a midrange average. However, this conclusion may not reflect reality. This examination does not seek information regarding an average consumer but rather seeks data on the average WTP values of many consumers. It

**Table 7: Two sets of C/V combinations and their respective principle WTP values**

1 <sup>st</sup> Version						2 <sup>nd</sup> Version					
Cash →		225.00	300.00	425.00	450.00	Cash →		150.00	350.00	400.00	450.00
Voucher →		610.00	525.00	325.00	75.00	Voucher →		650.00	400.00	150.00	0.00
Cash ↓	Voucher ↓					Cash ↓	Voucher ↓				
225.00	610.00	-	<b>600.00</b>	<b>525.00</b>	<b>300.00</b>	150.00	650.00	-	<b>600.00</b>	<b>400.00</b>	<b>300.00</b>
300.00	525.00	-	<b>450.00</b>	<b>225.00</b>		350.00	400.00	-	<b>200.00</b>	<b>100.00</b>	
425.00	325.00			<b>100.00</b>		400.00	150.00			<b>50.00</b>	
450.00	75.00					450.00	0.00				

Principle WTP values are in bold.

appears to be likely that a reasonable number of customers will buy a product even at high prices (such as first movers and early adopters), and there will be a high number of consumers who will not acquire the good at all. Thus far, there appears to be support for the assumption that the C/V method results in more realistic demand functions than that yielded by the Vickrey auction and the BDM procedure.

## 6 Empirical Study 2

Study 1 examined one C/V set with four alternatives and one alternative set with six alternatives using a within-subject design. Study 2 analyzes (1) how different C/V sets with four alternatives affect WTP values and (2) whether a between-subject design yields results that differ from the values of a within-subject design. Rather than a fairly unknown commodity, such as the cookies used in the first empirical study, the category of tablet PCs was chosen for study 2.

**(1) Two C/V sets of four alternatives.** Tablet PCs are regularly sold at prices from approximately €100 to €600. As a consequence, the C/V alternatives had to be selected within that range. Two different sets of four combinations were chosen (Table 7). Price ranges can easily be altered by multiplying the C/V portions of a known set of alternatives with a scalar. As an example, the figures from the first version have merely been calculated by multiplying the numbers of the alternative with four combinations (Table 2) by 50. No further regression procedure was necessary. The CEs remain the same, and the WTP values only need to be multiplied by the same scalar. The second version was calculated as a completely new set. Although the principle WTP values in the first version were relatively higher than

those in the second version, we did not expect different outcomes of the average WTP:

P<sub>4</sub>: The choice of relatively lower or higher principle WTP values in two sets of four C/V alternatives covering the same price range does not result in different average WTP values.

**(2) Within- and between-subject measurement.** In study 1, the participants stated their preferences for the four and six alternatives consecutively. Hence, the second set of statements could be biased as a result of a learning effect (Campbell and Stanley 1963). To be able to compare such a within-subject design with a between-subject design, both the first and second versions of the C/V alternatives were applied in two separate surveys.

### 6.1 Design, stimuli, and procedure

The survey participants were instructed with the following text:

“Imagine someone offers you a gift, such as for Christmas or your birthday. The gift consists of one portion of cash money, which you can spend at your own discretion. The other part of the gift is a voucher that is earmarked for the purchase of a tablet PC. You may choose any product from any price range (e.g., Apple, Samsung, Asus, Nexus, Amazon, Acer). Additionally, the voucher is not constrained to be redeemable only at specific places. You may choose any dealer you like. However, you may not transfer the voucher to someone else. You may only redeem the voucher in its entirety. Assume that you receive a voucher worth €200 and that the product costs €300; then you would pay only the remaining €100 in cash. If the voucher’s value is €400 and the tablet that you choose costs only €300, then you do not have to pay anything. However, the remainder of the voucher’s value is forfeited.



Hereafter, we present four alternatives of such C/V combinations. Please rank these alternatives according to your preferences.”

In the first survey, the first and second versions were chosen by a random rotation procedure, in which the sequence of the versions was changed by coincidence (within subjects). In the second survey, only one of the two versions was selected randomly (between subjects). The surveys were conducted online, inviting participants through their e-mail addresses, which were supplied by a German provider with a panel consisting of approximately 65,000 regular e-commerce customers.

## 6.2 Results

In the random rotation survey (within subjects), 125 respondents participated (31.2% female,  $M_{Age} = 45.8$ ,  $SD_{Age} = 13.5$ ). The average WTP in the group for the first version was €319.11 ( $SD = €221.80$ ), and it was €307.53 ( $SD = €225.05$ ) in the second version. In the random selection survey (50.9% female,  $M_{Age} = 41.5$ ,  $SD_{Age} = 12.5$ ), 102 participants rated the alternatives in the first version ( $M = €291.06$ ,  $SD = €249.09$ ), and 81 rated the second version of C/V sets ( $M = €262.10$ ,  $SD = €215.21$ ). A paired *T*-test for both versions in the within-subject survey did not show a significant difference ( $T(124) = .898$ ,  $p = .371$ ). Thus, the within-subject judgments can be viewed as consistent over the two versions ( $SE$  of the difference = 12.90, 95% confidence interval:  $-13.95 \leq 37.11$ ). An independent-samples *T*-test in the between-subject survey resulted in  $t(181) = .843$ ,  $p = .400$ . In addition, the assumption of equality could not be rejected here. As a result, there is evidence that the C/V procedure possesses internal and external validity. The reason may be the complexity of the trade-off task: it is still sufficiently plain and simple to solve it intuitively; however, the mechanics behind it are not sufficiently obvious to facilitate cognitive understanding of the solution process.

The second (between-subject) survey was conducted approximately four months after the within-subject study. The average WTP values emerged to be approximately €30 lower than before. This 5% reduction was consistent with the regular price decline in the IT business. Study 2 is another example of the flexibility of the C/V method, which needs neither real nor single identifiable products. Here, the WTP

values for an entire product range could be elicited by a lean and concise procedure.

## 7 Empirical Study 3

Study 3 analyzed the elicitation of WTP values for an unavailable good, which presents an advantage over incentive-compatible methods.

### 7.1 Propositions and expectations

Two more versions of C/V alternatives were chosen (Table 8). One version covered a range from €0 to €1,500 (third version), and another covered amounts from €0 to €2,000 (fourth version). We expected a higher average WTP in the fourth version as a result of the anchoring effect (Tversky and Kahneman 1974; Wansink, Kent, and Hoch 1998). However, we expected that this result would not be caused by the use of a demand function that was similar to the function describing the results of the open-ended procedure in study 1 (Figure 1). Rather, only the selected WTP values above €1,500 in the fourth version were expected to contribute to the higher average. The average WTP below €1,500 should not be different, regardless of the version. In both cases, the shape of the demand curves from €0 to €1,500 should be similar. This property makes the C/V procedure immune to anchoring effects, at least in the low- and mid-price ranges. The respective propositions are as follows:

- P<sub>5</sub>: Because of the anchor effect, the average WTP from a set of alternatives covering the price range from €0 to €2,000 will be higher than the average WTP from a set covering the range from only €0 to €1,500.
- P<sub>6</sub>: The higher average WTP exclusively results from the WTP values above €1,500 in the version with the higher range. Average WTP values do not differ below €1,500; thus, the demand curves in the range from €0 to €1,500 are similar for both versions.

Study 3 combined the analysis of the anchor effect with the evaluation of different elicitation procedures. Several stated preference methods exist (Hanley, Mourato, and Wright 2001). Among these methods are the rankings that have been used in this article thus far. In addition, there are choice experiments in which participants choose between two alternatives. In CA, the choice-based procedure has increasingly gained acceptance. With a simple



**Table 8: Third and fourth versions of C/V combinations**

3 <sup>rd</sup> Version						4 <sup>th</sup> Version					
Cash →	100.00	500.00	700.00	900.00		Cash →	200.00	700.00	850.00	900.00	
Voucher →	1700.00	1100.00	650.00	0.00		Voucher →	2100.00	1500.00	650.00	0.00	
Cash ↓	Voucher ↓					Cash ↓	Voucher ↓				
100.00	1700.00	-	<b>1500.00</b>	<b>1250.00</b>	<b>800.00</b>	200.00	2100.00	-	<b>2000.00</b>	<b>1300.00</b>	<b>700.00</b>
500.00	1100.00		-	<b>850.00</b>	<b>400.00</b>	700.00	1500.00		-	<b>800.00</b>	<b>200.00</b>
700.00	650.00			-	<b>200.00</b>	850.00	650.00			-	<b>50.00</b>
900.00	0.00				-	900.00	0.00				-

Principle WTP values are in bold.

decomposition of the rankings for  $(n - 1) \times n/2$  pairwise choices, this procedure could also be employed for the C/V version. It would thus be of interest to analyze the effects on the stated WTP values. In the version with four alternatives, this analysis would require that participants had to compare no more than six choices, which could make it an easy task. By virtue of the convenient nature of the C/V method, we do not expect any difference between the ranking and choice procedures.

P<sub>7</sub>: When applying two different stated preference methods in a C/V experiment, no differences will result between a ranking and a pairwise choice procedure.

### 7.2 Design, sample, and procedure

WTP values were elicited for a hypothetical product called “Pivotablet,” which can be folded from a 4.5” smartphone to a 10” tablet. The range of retail prices is expected to be near €1,000.

**Figure 2: The hypothetical “Pivotablet” with hinged 4.5”, 7”, and 10” screen**



The survey itself had a design with 2 version  $\times$  2 stated preference methods. The C/V method was explained to the participants, and the “Pivotablet” was introduced as the underlying product. A photograph (Figure 2) was shown, and a short description explained the advantages of the new innovation.

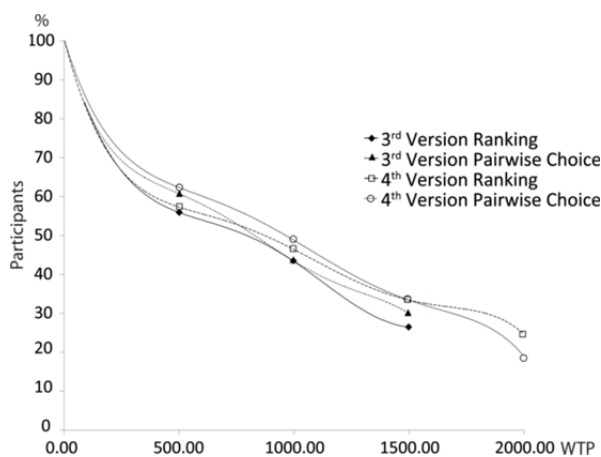
The respondents first ranked the randomized alternatives from the third and fourth versions (Table 8). Thereafter, they needed to choose in a pairwise fashion which of the two alternatives they preferred. To this end, the rankings from each version were presented as six pairs (alternative 1 vs. 2, 1 vs. 3, 1 vs. 4, 2 vs. 3, 2 vs. 4, 3 vs. 4) and were presented randomly. After the pairwise comparison, the choices were recomposed to the rankings. The survey was conducted online again, choosing e-mail addresses randomly from the same German panel.

### 7.3 Results

In total, 249 respondents participated in this survey (38.5% female,  $M_{Age} = 46.2$ ,  $SD_{Age} = 11.8$ ). The rankings in the third and fourth versions were successfully completed by all of the participants, perhaps because of the rather intuitive procedure of the online survey, in which the four alternatives could be placed as cards that could be dragged and dropped in the preferred sequence. When making their pairwise choices, some participants tended to intransitivity. An example is  $1 > 2$ ,  $2 > 3$ ,  $3 > 1$ ,  $1 > 4$ ,  $2 > 4$ , and  $3 > 4$ , with a reversal between 1 and 3, thus leading to an intransitive preference of  $1 > 2 > 3 > 1 > 4$ . The respective preference points are then  $3 - 3 - 3 - 1$  (note that this is the rating but not the ranking). However, these ratings can also be used in the MONANOVA. In the third version, the results show a WTP of €1178.90. For the paired choice method, 22 WTPs could be calculated similarly in the third version and 28 in the fourth version. How-

ever, in 42 cases in the third version and in 38 cases in the fourth version, the participants did not make all six pairwise choices (i.e., at least one of the six binary choices was missing). In some cases, a calculation of the priorities would be possible. Assumed in the example above, the intransitivity of  $3 > 1 > 2 > 3 > 4$ . However, this assumption seemed somewhat arbitrary, and according to Hair, Black, and Anderson (2010), those responses in which all pairwise choices were available were corrected (although there were intransitivities). Where choices were missing, however, the whole pairwise choice case was removed.

**Figure 3: The demand curves for two sets of C/V alternatives and two elicitation methods**



The third version with ranked alternatives ( $n = 249$ ) resulted in an average WTP of €749.52 ( $SD = 610.04$ ), and the third version with pairwise choices ( $n = 207$ ) resulted in an average WTP of €790.39 ( $SD = 606.91$ ). The fourth version with ranked alternatives ( $n = 249$ ) resulted in an average WTP of €931.18 ( $SD = 804.81$ ), and the fourth version with pairwise choices ( $n = 211$ ) resulted in an average WTP of €959.86 ( $SD = 768.66$ ). The equality of WTP in the third version cannot be rejected ( $p = .476$ ), nor can the fourth version ( $p = .696$ ), which supports  $P_7$ . However, the anchoring effect of the fourth version with a higher maximum WTP of €2,000 resulted in a higher average WTP ( $p < .010$  for the ranking method and  $p < .050$  for the pairwise choices, respectively). As expected,  $P_5$  is supported. However, the removal of stated WTP values at or above €1,500 resulted in WTP values for which the assumption of equality could not be rejected ( $p$

$= .801$ ) between the third and fourth versions using the ranking procedure;  $p = .162$  in the choice method). For all four methods, an ANOVA could no longer find an effect for values below €1,500 ( $F(3,641) = .896$ ,  $p = .443$ ). The demand curves confirm these findings (Figure 3). Up to the lower maximum of €1,500 in the third version, their shape is similar. Only the stated WTP values above €1,500 contribute to the higher average in the fourth version. Thus,  $P_6$  cannot be rejected.

#### 7.4 Discussion

Study 3 again shows evidence of the relative robustness of the procedure. Particularly in the case of new products, a suitable price range cannot always be determined. However, the possibility that a price range may have been selected up to an overestimated maximum does not render all of the data unusable. Low- and mid-price WTP values remain validly represented in the demand functions. Here as well, the distributions did not appear to be normal, and nonparametric tests were applied. The results did not offer different evidence compared with the parametric analyses. The implementation of either a ranking or pairwise-choice method did not result in any differences. The surveys had been designed such that the participants were not obliged to answer the questions. Therefore, missing values could appear. Surprisingly, no such missing values occurred in the ranking procedure; however, approximately 25% of the values were missing when the paired choices needed to be made. When we asked the participants which method they preferred, 111 out of 249 voted for the ranking (44.6%), and 138 voted for the choice method. Given this insignificant difference ( $\chi^2(1) = 2.928$ ,  $p = .087$ ) and given the lack of average WTP difference resulting from both procedures, there appears to be no reason to implement a pairwise-choice task.

## 8 General Discussion

### 8.1 Results

This paper proposes a new method for measuring WTP. The new method is based on disposable cash (C) and specific-purpose vouchers (V) that are earmarked for particular goods such that the hypothetical bias is minimized. The so-called C/V procedure is a hypothetical task. No real environment needs to be established, as with auctions. Consequently, no



real money is required, and the products under investigation can have high sales prices. In addition, the goods need not exist or be defined in detail; they need only be determinable rather than determinate. The results of the study showed evidence of criterion validity compared with the Vickrey auction and BDM procedures (study 1). As with these other methods, the C/V procedure also demonstrated ecological validity.

Furthermore, it could be demonstrated that there is both internal and external validity, showing that the respondents were able to solve such tasks consistently. As shown above, WTP values for an entire product range with various price ranges can be elicited (study 2).

Consumer spending behavior can be assessed by C/V surveys with divisible goods, such as food stamps and event vouchers. Additionally, unavailable products or services can be the underlying good of the vouchers; thus, WTP values for brand new commodities that are still in the design process can be assessed with the demonstrated C/V method (study 3).

## 8.2 Implications

No differences were revealed with regard to the results that were achieved between C/V procedures that offered four or six alternatives of C/V combinations. With fewer alternatives, the principle WTP values, rankings, and regression results are easier to handle. Thus, it appears advisable to choose the construction with four sets.

Finding a suitable set of alternatives of C/V combinations for specific demand functions requires some heuristic effort. However, this issue is well known with respect to CA. Although these alternatives – particularly price settings – affect CA results, we were able to show that different sets of combinations have no effect on the average WTP but that they affect only the shape of the demand curve. A researcher can take advantage of this finding, depending on the assumptions that underlie the respective demand curve.

We chose to calculate the regressions with a MONANOVA (i.e., ordinal regression), although it has become standard to treat ordinal data as metric and to employ a linear regression procedure. This tendency may hold for Likert scales, which are believed to have an equidistant, or pseudo-interval characteristic (Green and Srinivasan 1978). In addition, Darmon and Rouziès (1994) claimed that there are

no significant differences between methods. However, examining the alternatives for C/V in the above examples, we find that equidistance is rather unlikely. We also conducted sensitivity analyses and found that all principle rankings were quite similar when comparing OLS and MONANOVA. However, with regard to more complicated rankings, there were more significant differences.

## 8.3 Limitations

The system of principle WTP values and rankings is based on decisions of rational decision makers with the axioms of normative decision theory. It may be argued that decisions in the real world often contradict these axioms (Kahnemann and Tversky 1984) and that descriptive decision theory instead informs us that humans generally do not make rational decisions. In a CA, different incompatible scales may compete (e.g., the color, power, or price of a car). No attribute can be expressed by another. In this case, trade-offs must be resolved by respondents' utility, which could be subject to inconsistencies. In the C/V procedure, however, the bases of decision making are calculated deterministically because one attribute can be expressed in terms of another. The trade-offs concern monetary decisions, and they can be calculated in exact terms. Consequently, decision makers can easily follow the logical principle of the avoidance of willful waste. Assume that someone has absolutely no interest in buying a tablet PC, and two alternatives of C/V combinations are offered to her or him: €100 cash plus a €100 voucher for a tablet PC versus €150 cash but no voucher. Because he or she is absolutely unwilling to acquire a tablet PC, it would be irrational to choose the first alternative and to waste €50 in cash. The possibility that this result may occur because of a lack of awareness or interest in the ranking task classifies these answers as errors. Statistical theory is able to handle these errors as noise, but the basic principle of rationality holds.

The survey participants in this study were informed that the vouchers would not be transferable and, in particular, that they would not be for sale. A tactical answer could involve choosing the alternative with the highest sum of cash and voucher and selling the commodity after the redemption. However, transaction costs could arise, and the risk may remain that the full amount is not obtained, for example, in an online auction. The high number of rankings that result in a WTP of zero is not indicative that many



tactical answers exist. Further research should analyze such potential behavior in greater detail.

Another shortcoming of the C/V procedure could be the difficulty in calculating regressions with good fit indices for some of the alternatives of C/V combinations. However, this problem is not unique to this method. In fact, these problems are well known with any regression-like estimation procedures (Green and Srinivasan 1990). Hair, Black, and Anderson (2010) recommended three measures as remedies: (1) ignore these problems, (2) remove them, or (3) correct them. With regard to correction, the parameter estimations can be constrained following the monotonicity that underlies the characteristics of the attributes (Srinivasan, Jain, and Malhotra 1983; Allenby, Arora, and Ginter 1995). We followed this recommendation through an alignment of the alternatives. The cash amounts from alternative to alternative are reduced stepwise while simultaneously increasing the voucher amounts and the sums of both in a stepwise manner (Table 2). Through this procedure, when two scores are swapped, for example, the higher score steps forward, and the lower score steps backward; therefore, we know that this ranking must have higher cash affinity and consequently that the WTP must be higher. Thus, a ranking of 4. – 3. – 1. – 2. has a higher WTP than 4. – 3. – 2. – 1. This constraint makes it simple to identify intransitivities and, if necessary, to correct them. Only five out of 61 participants in study 1 had chosen such intransitive rankings in the version with four alternatives. Ignoring, removing, or correcting them had no significant effect on WTP values and demand curves. The percentage in the version with six alternatives was higher; however, the results did not differ regardless of the action that was taken.

The context of the WTP values regarding two C/V combinations can be expressed by a rather trivial equation by adding and subtracting three numbers. However, this finding is not obvious, particularly to survey participants who are in the process of resolving such a stated preference task for the first time. When this method is adopted more frequently, the rationales behind it could theoretically be discovered by participants, who could learn how to solve the task in a manner that is similar to an open-ended question. Further research should concentrate on analyzing whether this problem could actually occur and how it could be avoided.

Researchers and practitioners may apply the method in several variations. In the studies above, the

respondents were free to choose the rankings, for example, out of 24 in the version with four alternatives. However, the number of rankings could be limited to a constrained set. One could offer exclusively principle WTP values and those derived from the regressions with excellent fit indices ( $R^2$  and Kendall's  $\tau$ ). This method has the advantage of avoiding any inconsistent rankings. Thus, the versions with seven alternatives could also be chosen because only a reduced set would need to be handled. Further research could analyze the effects on average WTP values and demand curves.

The validity and particularly the robustness of the procedure that is described may be rather surprising and – if resulting from principles from cognitive psychology – these underlying rules have not yet been discussed in detail. Further research may reveal such fundamentals and may help to release the method from its somewhat heuristic nature.

## References

- Ajzen, Icek, Thomas C. Brown, and Franklin Carvajal (2004): Explaining the Discrepancy between Intentions and Actions: The Case of Hypothetical Bias in Contingent Valuation, *Personality and Social Psychology Bulletin*, 30 (9): 1108-1121.
- Allenby, Greg M., Neeraj Arora, and James L. Ginter (1995): Incorporating Prior Knowledge into the Analysis of Conjoint Studies, *Journal of Marketing Research*, 32 (2): 152-162.
- Barrot, Christian, Sönke Albers, Bernd Skiera, and Björn Schäfers (2010): Vickrey vs. eBay: Why Second-Price Sealed-Bid Auctions Lead to More Realistic Price-Demand Functions, *International Journal of Electronic Commerce*, 14 (4): 7-38.
- Bateson, John E. G. and Michael K. Hui (1992): The Ecological Validity of Photographic Slides and Videotapes in Simulating the Service Setting, *Journal of Consumer Research*, 19 (2): 271-281.
- Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak (1964): Measuring Utility by a Single-Response Sequential Method, *Systems Research and Behavioral Science*, 9 (3): 226-32.
- Blatter, David, Klaus Miller, Reto Hofstetter, and Harley Krohmer, H. (2009): Improving the Direct Estimation of Demand by Adjusting for Incorrect Price-Statements, *Proceedings of the 38<sup>th</sup> EMAC Conference*, 26-29 May 2009, Nantes, France.
- Bornstein, Brian H. (1999): The Ecological Validity of Jury Simulations: Is the Jury Still out?, *Law and Human Behavior*, 23 (1): 75-91.
- Brunswik, Egon and Joe Kamiya (1953): Ecological Cue-Validity of 'Proximity' and of Other Gestalt Factors, *American Journal of Psychology*, 66 (1): 20-32.
- Campbell, Donald. T. and Julian Stanley (1963): *Experimental and Quasi-Experimental Designs for Research*, Houghton Mifflin: Boston et al., MA.





- Darmon, René Y. and Rouziès, Dominique (1994): Reliability and Internal Validity of Conjoint Estimated Utility Functions under Error-Free versus Error-Full Conditions, *International Journal of Research in Marketing*, 11 (5): 465-476.
- Dasgupta, Partha, Peter Hammond, and Eric Maskin (1979): The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility, *The Review of Economic Studies* 46 (2): 185-216.
- De Leeuw, Jan, Forrest W. Young, and Yoshio Takane (1976): Additive Structure in Qualitative Data: An Alternating Least Squares Method with Optimal Scaling Features, *Psychometrika*, 41 (4): 471-503.
- Ding, Min (2007): An Incentive-Aligned Mechanism for Conjoint Analysis, *Journal of Marketing Research*, 44 (2): 214-223.
- Ding, Min, Rajdeep Grewal, and John Liechty (2005): Incentive-Aligned Conjoint Analysis, *Journal of Marketing Research*, 42 (1): 67-82.
- Dong, Songting, Min Ding, and Joel Huber (2010): A Simple Mechanism to Incentive-Align Conjoint Experiments, *International Journal of Research in Marketing*, 27 (1): 25-32.
- Dost, Florian and Robert Wilken (2012): Measuring Willingness to Pay as a Range, Revisited: When Should We Care?, *International Journal of Research in Marketing*, 29 (2): 148-166.
- Green, Paul E., Kristiaan Helsen, and Bruce Shandler (1988): Conjoint Internal Validity under Alternative Profile Presentations, *Journal of Consumer Research*, 15 (3): 392-397.
- Green, Paul E. and Vithala R. Rao (1971): Conjoint Measurement for Quantifying Judgmental Data, *Journal of Marketing Research*, 8 (3): 355-363.
- Green, Paul E. and V. Srinivasan (1978): Conjoint Analysis in Consumer Research: Issues and Outlook, *Journal of Consumer Research*, 5 (2): 103-123.
- Green, Paul E. and V. Srinivasan (1990): Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice, *Journal of Marketing*, 54 (4): 3-19.
- Hair, Joseph F. Jr., William C. Black, Barry J. Babin, and Rolph E. Anderson (2010): *Multivariate Data Analysis: A Global Perspective*, 7<sup>th</sup> ed., Pearson: Upper Saddle River, NJ.
- Hamlin, Robert P. (2005): The Rise and Fall of the Latin Square in Marketing: A Cautionary Tale, *European Journal of Marketing*, 39 (3/4): 328-350.
- Hanemann, W. Michael (1991): Willingness to Pay and Willingness to Accept: How Much Can They Differ?, *The American Economic Review*, 81 (3): 635-647.
- Hanley, Nick, Susana Mourato, and Robert E. Wright (2001): Choice Modelling Approaches: A Superior Alternative for Environmental Valuation?, *Journal of Economic Surveys*, 15 (3): 435-462.
- Heckman, James (1990): Varieties of Selection Bias, *American Economic Review*, 80 (2): 313-318.
- Herrmann, Andreas, Mark Heitmann, Robert Morgan, Stephan C. Henneberg, and Jan Landwehr (2009): Consumer Decision Making and Variety of Offerings: The Effect of Attribute Alignability, *Psychology & Marketing*, 26 (4): 333-358.
- Kahneman, Daniel and Amos Tversky (1984): Choices, Values and Frames, *American Psychologist* 39 (4): 341-350.
- List, John A. and Craig A. Gallet (2001): What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values? Evidence from a Meta-Analysis, *Environmental and Resource Economics*, 20 (3): 241-254.
- Little, Joseph and Robert Berrens, (2004): Explaining Disparities between Actual and Hypothetical Stated Values: Further Investigation Using Meta-Analysis, *Economics Bulletin*, 3 (6): 1-13.
- Milgrom, Paul R. and Robert J. Weber (1982): A Theory of Auctions and Competitive Bidding, *Econometrica*, 50 (5): 1089-1122.
- Miller, Klaus M., Reto Hofstetter, Harley Krohmer, and Z. John Zhang (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.
- Moffitt, Robert (1989): Estimating the Value of an In-Kind Transfer: The Case of Food Stamps, *Econometrica*, 57 (2): 385-409.
- Motowidlo, Stephan J., Marvin D. Dunette, and Garry W. Carter (1990): An Alternative Selection Procedure: The Low-Fidelity Simulation, *Journal of Applied Psychology*, 75 (6): 640-647.
- Müller, Holger, Steffen Voigt, and Bernd Erichson (2009): Befragungs-basierte Methoden zur Ermittlung von Preisresponsefunktionen: Preisbereitschaft oder Kaufbereitschaft, Working Paper, RePEc.
- Murphy, James J., P. Geoffrey Allen, Thomas H. Stevens, and Darryl Weatherhead (2005): A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation, *Environmental and Resource Economics*, 30 (3): 313-325.
- Neisser, Ulric (1976): *Cognition and Reality: Principles and Implications of Cognitive Psychology*, Freeman: New York, NY.
- Park, Young-Hoon, Min Ding, and Vithala R. Rao (2008): Eliciting Preference for Complex Products: A Web-Based Upgrading Method, *Journal of Marketing Research*, 45 (5): 562-574.
- Raser, John R. (1969): *Simulation and Society: An Exploration of Scientific Gaming*, Allyn and Bacon: Boston, MA.
- Rogers, Steven D., Endre E. Kadar, and Alan Costall (2005): Gaze Patterns in the Visual Control of Straight-Road Driving and Braking as a Function of Speed and Expertise, *Ecological Psychology*, 17 (1), 19-38.
- Simon, Julian L. (1969): A Further Test of the Kinky Oligopoly Demand Curve, *American Economic Review*, 59 (5): 971-975.
- Srinivasan, V., Arun K. Jain, and Naresh K. Malhotra (1983): Improving Predictive Power of Conjoint Analysis by Constrained Parameter Estimation, *Journal of Marketing Research*, 20 (4): 433-438.
- Sweezy, Paul M. (1939): Demand under Conditions of Oligopoly, *The Journal of Political Economy*, 47 (4): 568-573.
- Teichert, Thorsten and Edlira Shehu (2010): Investigating Research Streams of Conjoint Analysis: A Bibliometric Study, *BuR – Business Research*, 3 (1): 49-68.
- Tversky, Amos and Daniel Kahneman (1974): Judgment under Uncertainty: Heuristics and Biases, *Science*, 185: 1124-1131.





Varian, Hal R. (1992): *Microeconomic Analysis*, 3<sup>rd</sup> ed., Norton: New York et al., NY.

Vickrey, William (1961): Counterspeculation, Auctions, and Competitive Sealed Tenders, *Journal of Finance*, 16 (1): 8-37.

Völckner, Franziska (2006): An Empirical Comparison of Methods for Measuring Consumers' Willingness to Pay, *Marketing Letters*, 17 (2): 137-149.

Wang, Tuo, R. Venkatesh, and Rabikar Chatterjee (2007): Reservation Price as a Range: An Incentive-Compatible Measurement Approach, *Journal of Marketing Research*, 44 (2): 200-213.

Wansink, Brian, Robert J. Kent, and Stephen J. Hoch (1998): An Anchoring and Adjustment Model of Purchase Quantity Decisions, *Journal of Marketing Research*, 35 (1): 71-81.

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

Wittink, Dick R., Lakshman Krishnamurthi, and David J. Reibstein (1989): The Effect of Differences in the Number of Attribute Levels on Conjoint Results, *Marketing Letters*, 1 (2): 113-123.

## Biographies

**Theo Lieven** is a research associate at the Center for Customer Insight at the University of St. Gallen. He received his Doctoral Degree in 2009 at the same university.

**Silke Lennerts** is Assistant Professor and Managing Director at the Center for Customer Insight at the University of St. Gallen. Her dissertation has been awarded the Amicitia prize of the University of St. Gallen in 2005.