**A De-biased Direct Question Approach to Measuring**

**Consumers’ Willingness to Pay**

**WEB APPENDIX**

**Max 2020**

Reto Hofstetter

Professor of Marketing

Faculty of Economics and Management, University of Lucerne

Frohburgstrasse 3, P.O. Box 4466, 6002 Lucerne, Switzerland

Phone: (41) 41 229 58 80

Email: reto.hofstetter@unilu.ch

Klaus M. Miller

Assistant Professor of Quantitative Marketing

Faculty of Economics and Business, Goethe University Frankfurt

Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main

Phone: (49) 69 798 34601

Email: klaus.miller@wiwi.uni-frankfurt.de

Harley Krohmer

Professor of Marketing

University of Berne, Engehaldenstrasse 4, 3012 Berne, Switzerland

Phone: (41) 31 631 8031

Fax: (41) 31 631 8032

Email: krohmer@imu.unibe.ch

Z. John Zhang

Tsai Wan-Tsai Professor, Professor of Marketing

Director, Penn Wharton China Center

The Wharton School

University of Pennsylvania, 3730 Walnut St, Philadelphia, PA 19104, USA

Phone: (01) 215 898 1989

Fax: (01) 215 898 2534

Email: zjzhang@wharton.upenn.edu

# WEB APPENDIX A

To better understand the adoption of demand measurement approaches in business practice, we conducted a brief survey among pricing managers in Switzerland. The basic population for our survey comprised all firms which are registered in the Swiss commercial registry. Based on a stratified sampling approach, which equally accounted for all sectors and firm sizes, we contacted 486 firms. After two weeks, we reminded all firms that had not yet participated to do so via the telephone. After this process, we ended up with 82 completed surveys, which corresponds to a completion rate of 17%. 65% of the completed surveys were answered by members of the senior management, 22% by department heads of marketing, and 13% by other functions within the firm. With regard to economic sectors, 44% of the firms covered in our survey came from industry, 39% from services, and 16% were in retailing. Firm revenue was almost evenly distributed and ranged from less than CHF 1 Mio. to more than CHF 500 Mio. Similarly, market shares ranged from less than 1% to more than 50% in our sample. Finally, the number of competitors was evenly distributed and ranged from monopoly, state-owned firms to firms in highly-competitive markets with more than 50 major competitors. We administered a survey consisting of 53 questions regarding pricing-related topics to the firms. Here, we report the results of the question that referred to the adoption of demand measurement approaches in business practice (see Figure A.1).

FIGURE A.1  
ADOPTION OF DEMAND MEASUREMENT APPROACHES IN BUSINESS PRACTICE



# WEB APPENDIX B

We assume with mean(. What we say is

1. f( ) need not be symmetrical
2. we can still have bias if f( ) is not symmetrical and mean ( ,

To show this:

1. Distribution function: where x1 is the range below zero with density y1 and x2 the range above zero with density y2.
2. Mean = 0

Mean =

=

=

Set mean = 0

1. and 3. can exist

3.

Proof: Let

0.7

Specifically:

If

Basically, this shows that with mean () = 0, you can still have lots of biases.

# WEB APPENDIX C

as

***WEB APPENDIX D******STIMULUS STUDY 1: UNIVERSITY GYM BAG***

Below we include the stimulus used in Study 1. We blinded the university logo for review and will unblind them if our paper is accepted for publication.

|  |  |
| --- | --- |
| https://unibe.eu.qualtrics.com/CP/Graphic.php?IM=IM_bQLPdnOjKIKJQYR | * Functional university gym bag. * Ideal for party, shopping, university, and sports. * 100% organic cotton. * High quality screen print. * Printed in Switzerland by hand. * Color: royal blue. * Size: 37 x 46 cm. * Free shipping. |

# WEB APPENDIX E TRANSPARENCY AND ACCEPTABILITY RATING QUESTIONS

Note: The specific experimental groups are denoted as OE (Open-ended Question Format), DC (Dichotmous-choice Question Format), and BDM (Becker, deGroot, Marschak Mechanism).

1. If the gym bag (sweatshirt) was actually offered to you for purchase in an online shop, how certain are you that you would purchase the product at **your** (OE/BDM) / **the** (DC)stated price? (1 = “very uncertain,” 7 = “very certain”)
2. Is it clear to you why it is in your best interest to **state** (OE/BDM) / **accept** (DC) exactly the price you are willing to pay for the gym bag (sweatshirt)? (1 = “not at all,” 7 = “very much so”)
3. OE/DC: Was it confusing to you to state your maximum willingness to pay for the gym bag (sweatshirt)? (1 = “not at all,” 7 = “very much so”); BDM: Was the buying process confusing to you? (1 = “not at all,” 7 = “very much so”)
4. BDM only: Did you understand the buying process? (1 = “not at all,” 7 = “very much so”)
5. OE/DC/BDM: This task was very easy to understand and complete. (1 = “not at all,” 7 = “very much so”)
6. OE/DC/BDM: Did you perceive this task as fair? (1 = “not at all,” 7 = “very much so”)
7. OE/DC/BDM: I will be happy to do this task again in the future. (1 = “not at all,” 7 = “very much so”)

# WEB APPENDIX F

FIGURE F.1  
AGGREGATE DEMAND CURVES FOR UNIVERSITY GYM BAG   
BASED ON DIFFERENT WTP MEASUREMENT AND DE-BIASING APPROACHES

# 

Note: The curves show the aggregate demand based on aggregating individuals’ responses to the OE, DC, or BDM questions and based on the different de-biasing approaches. OE: Open-ended question format; DC: Dichotomous-choice question format; BDM: Becker, DeGroot, and Marschak Mechanism; Quantity is the aggregate demand that is normalized to a range between 0 and 1 due to differing sample sizes. Prices are in CHF (Swiss Francs), which at the time of the study equal approximately to prices in USD.

# WEB APPENDIX G UNIVERSITY SWEATSHIRT STUDY (STUDY 2)

We replicated our de-biasing procedure demonstrated in the university gym bag study (Study 1) in a second study using a university sweatshirt as a stimulus.

# METHOD AND DATA COLLECTION

***Participants***. The data for the university sweatshirt study was also collected through an online experiment. To recruit participants, we sent out 12,448 invitation emails to the entire student body (undergraduate, graduate, and Ph.D. candidates) of a large Swiss university. We motivated participation by offering all survey participants a chance to win an Apple iPhone 7 Plus in a raffle[[1]](#footnote-2). The participants were further informed that their chance to win the raffle was independent of their experimental responses[[2]](#footnote-3). A total of 772 participants chose to take part in the sweatshirt-study, which represents a response rate of 6.18%. We pre-specified that data collection would end after seven days (i.e., the decision to stop collecting data was independent from the experimental results; we did not analyze the data until after data collection for the given study had been completed). Within the seven-day period, we collected as much data as we could.

***Stimulus.*** The stimulus we used in our second study was a hooded sweatshirt imprinted with a logo of the university that was not available in the market at the time of the study (see Figure G.1 for a depiction of the stimulus). The sweatshirt was designed and fabricated exclusively for the purpose of the study. We expected the university sweatshirt to be both interesting and affordable for most of the students, who actually represent the target segment of the product. As apparel and accessories are also often sold online, the online channel represents the known and accustomed distribution channel for these categories.[[3]](#footnote-4). Further, because the students had no reference for the exact market price for this distinctive university sweatshirt, their HWTP or AWTP statements would not be capped[[4]](#footnote-5).

As our stimulus were new to the market, no repeat purchases were observed and participants were allowed to state their WTP only once. Further, the university sweatshirt was not displayed in a competitive setting. This meant that students were unable to select the stimulus from a group of competing products as they would be able to in a real online store. Finally, because we were using an online experiment, shipping the product had to be easy and cheap, which was the case with the sweatshirt.

FIGURE G.1

STIMULUS STUDY 2: UNIVERSITY SWEATSHIRT

Below we include the stimulus used in Study 2. (We blinded the university logo for review and will unblind them in case our paper gets accepted for publication.)

|  |  |
| --- | --- |
| https://unibe.eu.qualtrics.com/CP/Graphic.php?IM=IM_0OKhuDEX5rwewRL | * Design and ethics combined in a high-quality  university sweatshirt (unisex). * Fair production according to high environmental and social standards. * Certification according to Oekotex Standard 100. * 100% organic cotton. * Durable screen printing. * Printed in Switzerland by hand. * Available in all sizes from S to XL. * Free shipping. |

***Experimental Design.*** We developed three different independent experimental groups per study and used a between-subjects design. Each participant was randomly assigned to one treatment group.

In the open-ended (OE) question format group, each participant directly stated her hypothetical WTP for the university sweatshirt.

In the dichotomous-choice (DC) question format group, we used a total number of 21 price levels, ranging from CHF 10 up to CHF 110 incremented in steps of CHF 5 for the sweatshirt study. We set the market price because the most expensive sweatshirt similar to our products as our upper limit was CHF 110. Each respondent received a price level that was chosen randomly out of the 21 available levels. The random distribution of the price levels was even, meaning that all the price levels had an equal chance of one in 21 to appear in the respondent’s DC question.

For the BDM group, our control group for validating our de-biasing procedures, we determined our benchmarking AWTP data by using an incentive-aligned mechanism, the BDM mechanism proposed by Becker, DeGroot, and Marschak (1964). We chose the BDM mechanism as WTP from BDM has been found to not significantly differ from WTP based on real purchase data (Miller et al. 2011). We implemented the BDM mechanism in a way similar to what Wertenbroch and Skiera (2002) did. In our particular application of the BDM mechanism, we told participants that they would have a chance to purchase the university sweatshirt without having to invest more money than they would be willing to pay for the product, that the price for the university sweatshirt was not yet set, and that it would be determined randomly from a predefined uniform distribution unknown to the participants. Participants were further told that they were obligated to buy the university sweatshirt at the randomly determined price if the price was less than or equal to their stated WTP. However, if the randomly determined price was higher, a respondent would not have a chance to buy the product. This mechanism ensures that participants have no incentive to state a price that is higher or lower than their true WTP.

To carry out the buying obligation, we recorded the name and address of each participant in the BDM group. After the completion of the study, we determined each individual participant’s buying obligation by drawing from a discrete uniform distribution which corresponded to the price-levels used under the DC question format. The distribution thus ranged from CHF 10 to CHF 110 incremented in steps of CHF 5 and included a total of 21 price levels. We determined per participant whether the randomly drawn price was smaller or equal to her stated WTP. Thus, none of the participants had to purchase the sweatshirt at a price that was larger than their stated WTP. Out of all participants in the BDM group, 12.70% were obliged to buy the product at an average price of 26.77 (SD = 16.26, min = 10, max = 75). Only 5 participants (16.13%) of those who were obliged to buy paid a price higher than the average BDM WTP of 37.12. After the completion of the study, all participants who were obligated to buy were sent the sweatshirt with an invoice, via the post. The invoice was due within 14 days and payable with cash or credit card. (This payment process was officially approved by the appropriate university authorities). In the sweatshirt study, out of all 244 respondents in the BDM group, 31 (12.7%) were required to purchase the sweatshirt[[5]](#footnote-6). Only one respondent refused to comply with her purchase obligation and returned the product[[6]](#footnote-7).

We obtained 262 responses in the OE group, 266 responses in the DC group, and 244 responses in the BDM group. Our realized sample size exceeds current expectations in experimental studies of larger than 50 respondents per cell (Simmons 2014).

The three experimental groups did not differ significantly in terms of socio-demographics or socioeconomic status. We performed a multivariate analysis of variance for the sweatshirt-study (Pillai-Spur: F = .984, p = .480) for age (F = .925, p = .428), sex (F = .879, p = .452), education (F = 1.696, p = .167), occupation (F = 1.712, p = .164), income (F = .071, p = .975), budget for clothing and accessories (F = .735, p = .532), and purchase interest (F = .695, p = .555).

***Experimental Procedure***. We divided our online experiment into three parts. The first part described the product (i.e., the university sweatshirt) in the OE, DC, and BDM groups. The second part consisted of the WTP task in the different experimental treatment groups. In the third part of the online experiment, we conducted a brief survey on the respondents’ socio-demographics and economics, and we made sure the participants understood the WTP elicitation method to which they were exposed (see Web Appendix D of the main paper for further details).

***WTP Estimation Procedure*.** Figure G.2 plots the observed demand in each treatment group. For the OE and BDM groups, we obtained respondent’s hypothetical WTP and actual WTP directly from the survey data and plot demand as the probability that a respondent`s WTP is equal to or greater than a certain price using the demand function of the form . For the DC group we plotted the choice share for each price level. We determine the face validity of WTP measures by correlating elicited WTP with the respondent`s purchase interest. We measured purchase interest itself by using a seven-point Likert scale ranging from one (low interest in the product) to seven (high interest in the product). Face validity is high for all methods because correlations are positive and significant (OE: r = .379, p = .000, BDM: r = .314, p = .000). We did not test the DC question format because HWTP data is not available on an individual level here.

# RESULTS

The university gym bag study (Study 1) demonstrated the great promise of our de-biasing procedure for the direct single question approach. In this study, we will apply the same procedure to a different data set for a university sweatshirt we have collected to investigate the robustness of our de-biasing approach.

As shown in G.2, the demand functions show a similar pattern compared to the gym bag data. At first glance, the OE demand is consistently inflated, while the DC demand appears to be rotated counter-clockwise. A closer inspection also confirms, as shown in Table F.1, that the OE data series is significantly different from the BDM both in mean [t(543.92)OE vs. BDM = 4.166, p < .01] and distribution (DKS-Test, OE vs. BDM = .181, p < .01; DLR-Test, OE vs. BDM = 38.65, p < .01). In addition, the 95%-confidence intervals for the OE mean and BDM mean do not overlap. As expected, the OE mean is significantly inflated relative to the BDM. However, the DC data series is closer to the BDM and passes all tests except the LR-test (DLR-Test, DC vs. BDM = 17.681, p < .01).

Table D.2 also confirms that our partial, BASIC and EPSILON, as well as FULL de-biasing procedures all have improved on the OE data series. All three de-biased data series pass all the tests except the KS-test, similar to the DC data series [t(502.67)BASIC vs. BDM = 1.86, p > .05, t(476.76)EPSILON vs. BDM = 1.22, p > .22, t(472.05)FULL (cov = -3.20) vs. BDM = .49, p > .05, t(475.06)FULL (cov = -5.64) vs. BDM = -.96, p > .05], (DKS-Test, BASIC vs. BDM = .17, p < .01, DKS-Test, EPSILON vs. BDM = .14, p < .05, DKS-Test, FULL (cov = -3.20) vs. BDM = .14, p < .05, DKS-Test, FULL (cov = -5.64) vs. BDM = .20, p < .01). However, the estimated means from our de-biased data series are significantly closer to the BDM mean and two of them beat the DC mean.

FIGURE G.2  
AGGREGATE DEMAND FOR UNIVERSITY SWEATSHIRT   
BASED ON DIFFERENT WTP MEASUREMENT APPROACHES



Note: The curves show the aggregate demand based on aggregating individuals’ responses to the OE, DC, or BDM questions. OE: Open-ended question format; DC: Dichotomous-choice question format; BDM: Becker, DeGroot, and Marschak Mechanism; Quantity is the aggregate demand that is normalized to a range between 0 and 1 due to differing sample sizes. Prices are in CHF (Swiss Francs), which at the time of the study equal approximately to prices in USD.

To see how these improvements might affect managerial decisions, we conducted a similar economic analysis as in Study 1. Here, we used marginal costs for the university sweatshirt as obtained from the manufacturer of c = CHF 15.00. [Table](#A646750634754315AE42658FAB8971B9) F.2 summarizes the results from this analysis. Similar to the previous study, the OE data series on the university sweatshirt generated statistically different optimal quantity and the DC data series different optimal price estimates. More importantly, both data series led to wildly overestimated profits, 48.83% and 23.03% respectively. The first BASIC de-biasing procedure subtracting only improved significantly on the original OE data, but did not beat the DC data series. The second one, EPSILON, where individual-specific variation is accounted for, improved on DC significantly and the profit estimate differed from BDM by only 6.16%. As [Table](#A646750634754315AE42658FAB8971B9) E.3 further shows, our FULL de-biasing procedure performs, once again, remarkably well. The estimates of the optimal price and quantity are not statistically different from BDM and the estimate of optimal profits is within 5.26%, -.01% of the BDM. Please see Figure G.3 for a visualization of the different demand curves resulting from these de-biasing approaches.

TABLE G.1  
STATISTICAL ANALYSIS OF COLLECTED DATA AND   
DE-BIASED DATA OF UNIVERSITY SWEATSHIRT

|  |  |
| --- | --- |
| Collected Data | |
| *Data Source* | *Mean [Confidence Interval]* |
| OE | 45.758a, b, c [43.277, 48.250] |
| DC | 40.324b [34.878, 45.550] |
| BDM | 37.120 [34.677, 39.564] |
| De-Biased Data | |
| *Data Source* | *Mean [Confidence Interval]* |
| BASIC[cov = 0, epsilon = 0] | 40.40b [37.92, 42.87] |
| EPSILON[cov = 0, epsilon = SD(OE)] | 39.63b [36.40, 42.86] |
| FULL[cov = -3.20, epsilon = SD(OE)] | 38.81b [35.69, 41.93] |
| FULL[cov = -5.64, epsilon = SD(OE)] | 35.14b [31.89, 38.40] |
| BDM | 37.12 [34.68, 39.56] |

Note: OE: Open-ended question format; DC: Dichotomous-choice question format; BDM: Becker, DeGroot, Marschak Mechanism; BASIC, EPSILON, FULL: Refer to the steps of our de-biasing procedure; Values are shown with their 95% confidence interval in brackets; Superscript a indicates significant difference in terms of mean (t-test) relative to the benchmark; Superscript b indicates significant difference in terms of distribution (KS-test). For the DC data we used a Likelihood ratio test for the distribution comparison and compared confidence intervals (calculated based on Krinsky and Robb’s (1986) procedure) for the mean comparison; Superscript c indicates non-overlapping confidence intervals relative to the BDM benchmark.

TABLE G.2  
ECONOMIC ANALYSIS RESULTS FOR UNIVERSITY SWEATSHIRT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Data Source* | *Optimal Price* | *Optimal Quantity* | *Optimal Profit* | *Profit Percentage  Difference to BDM* |
| OE | 39.9 [26.330, 44.430] | .664c,d [.574, .911] | 16,536c,d [14540, 17,947] | 48.83%c,d |
| DC | 49.112d [43.354, 55.342] | .402 [.335, .4765] | 13,670d  [11,173, 16,2701] | 23.03%c,d |
| BASIC[cov = 0, epsilon = 0] | 43.566 [39.100, 48.780] | .484 [.364, .579] | 13,846d [12,041, 15,460] | 24.62%d |
| Epsilon[cov = 0, epsilon = SD(OE)] | 43.614 [34.990, 54.630] | .412 [.224, .512] | 11,795 [9,947, 13,020] | 6.16% |
| FULL[cov = -3.20, epsilon = SD(OE)] | 44.748 [36.360, 54.360] | .393 [.249, .490] | 11,695 [9,834, 13,215] | 5.26% |
| FULL[cov = -5.64, epsilon = SD(OE)] | 43.266 [37.580, 46.550] | .389 [.328, .477] | 11,004 [9,148, 12,603] | -.01% |
| BDM | 40 [35.520, 45.510] | .444 [.331, .530] | 11,111 [9,605, 12,573] | N.A. |

Note: OE: Open-ended question format; DC: Dichotomous-choice question format; BASIC, EPSILON, FULL: Refer to the steps of our de-biasing procedure; BDM: Becker, DeGroot, Marschak Mechanism; Quantity scaled from [0, 1]. N.A. = not applicable; Values are shown with their 95% confidence interval in brackets; We checked for a non-overlapping of confidence intervals as a test for significant differences; Superscript c indicates non-overlapping confidence intervals; Superscript d indicates significant difference at p = .05 (bootstrapping the difference between the measures); The shaded cells indicate that the confidence interval of the specific measure overlaps with the confidence interval of the corresponding benchmark measure obtained from our BDM data. Thus, shaded areas imply no statistical difference between the estimated measure and the benchmark.

FIGURE G.3  
AGGREGATE DEMAND FOR UNIVERSITY SWEATSHIRT   
BASED ON DIFFERENT WTP MEASUREMENT AND DE-BIASING APPROACHES

******

Note: The curves show the aggregate demand based on aggregating individuals’ responses to the OE, DC, or BDM questions and based on the different de-biasing approaches. OE: Open-ended question format; DC: Dichotomous-choice question format; BDM: Becker, DeGroot, and Marschak Mechanism; Quantity is the aggregate demand that is normalized to a range between 0 and 1 due to differing sample sizes. Prices are in CHF (Swiss Francs), which at the time of the study equal approximately to prices in USD.

***WEB APPENDIX H  
ROBUSTNESS OF RESULTS TO THE DC PRICE RANGE***

***Setup of Simulation.*** We explored thesensitivity of our results to the chosen DC price range in a simulation study. We followed the setup of our experimental design, but generated WTP values for OE, DC, and BDM from a known distribution. This allowed us to know a consumer’s actual WTP and to keep the number of individuals in the DC group constant, despite narrowing the range of DC price levels (i.e., by removing price levels)[[7]](#footnote-8).

We randomly drew WTP values from a normal distribution with a mean of CHF 50 and standard deviation of CHF 10 within a range between CHF 15 and CHF 85. We randomly drew 250 actual WTP values in each of the three experimental groups, OE, DC, and BDM. We biased the randomly drawn OE and DC data according to our bias models. For OE, we set the product category-specific inflator (i.e., half the mean of the OE data from the university sweatshirt study; see Web Appendix F). We sampled randomly at the individual-level according to the function described in Web Appendix B (ranging from -1 to +2). To generate the simulated DC data, we randomly assigned a price level out of all price levels and simulated a consumer’s response to the DC question depending on her generated actual WTP.

We repeated the above steps for 10 different sets of price levels. We started with 21 price levels ranging from CHF 0 to CHF 100 incremented in steps of CHF 5 (i.e., 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100). In a second step, we removed the CHF 0 and CHF 100 levels. In a third step, we further removed the CHF 0 and CHF 100, and the CHF 5 and CHF 95 price levels, and so on until we narrowed down the DC price range to only three remaining levels, CHF 45, CHF 50, and CHF 55. For each resulting set of price levels, we randomly drew 1,000 WTP samples for each simulated group OE, DC, and BDM. Importantly, the price levels were randomly allocated to DC WTP values in each set. This allowed keeping the number of respondents constant despite having fewer price levels (i.e., when there are fewer price levels they are simply randomly distributed to more individuals). We then applied the bias correction approaches BASIC, EPSILON and FULL, and calculated mean WTP, the optimal price, optimal quantity and profits including a 95%-confidence interval based on the 1,000 samples of the de-biased data. All results are depicted in Figures H.1, H.2.

***Results of Simulation.*** We distinguished two different cases for the range of the DC price levels: First, the DC price range could be wider than the price range of the actual WTP (i.e., to the left of the vertical line in Figures H.1, H.2), or narrower than the price range of the actual WTP (i.e., to the right of the vertical line in Figures H.1, H.2). If the DC range chosen is wider than the actual price range, both non-parametric and parametric versions of all three de-biasing approaches are able to recover the actual WTP values. If the DC range is narrower than the actual price range, we see that the de-biasing approaches will still be able to generate the actual values, as long as the range is not too narrow (i.e., the range is reduced down to seven price levels, which means that the DC range is 35% narrower than the actual price range). We see that if the DC range is not too narrow, the confidence intervals still overlap (> 7 price levels). If there are fewer price levels, the results start to change. For instance, the mean values significantly reduce. This can be explained by the non-parametric approach we used to calculate mean WTP for DC (Figure H.1). In this case, narrowing the DC range will result in missing observations for large values and a missing upper bound for WTP, which will be ignored and not count towards the non-parametric mean. The lower values still count as the lower bound for WTP is always zero. In turn, this leads to a reduction of the mean at narrower DC ranges.

Alternatively, the DC mean can also be calculated using a parametric approach as follows: First, a logistic regression is estimated based on the price levels and yes/no responses. The estimated coefficients for the intercept and slope then define a parametric logistic function. This function covers the full range of possible prices, allowing an unbiased calculation of the DC mean (using ); see also Hanemann 1989). This is confirmed in figures H.2 where the parametric de-biasing approaches generated accurate mean WTPs, optimal price, optimal quantity, and optimal profit values over all ranges of DC price levels.

FIGURE H.1  
SENSITIVITY OF NON-PARAMETRIC DE-BIASING APPROACHES TO   
A REDUCTION OF THE NUMBER OF DC PRICE LEVELS



Note: We use a dodged presentation to facilitate interpretation of the graphs. Pluses indicate the true value of Mean WTP, Optimal Price, Quantity, and Profit. The circles, triangles and squares indicate the values recaptured by the respective non-parametric de-biasing approach, depending on the number of DC price levels used when collecting DC data (x-axis). The vertical line indicates the actual price range of the simulated WTP data. On the left of the vertical line, the additional price levels are outside the range of actual WTP values. On the right of the vertical line, the price levels are within the range of actual WTP values.

FIGURE H.2  
SENSITIVITY OF PARAMETRIC DE-BIASING APPROACHES TO   
A REDUCTION OF THE NUMBER OF DC PRICE LEVELS



Note: We use a dodged presentation to facilitate interpretation of the graphs. Pluses indicate the true value of Mean WTP, Optimal Price, Quantity, and Profit. The circles, triangles and squares indicate the values recaptured by the respective parametric de-biasing approach, depending on the number of DC price levels used when collecting DC data (x-axis). The vertical line indicates the actual price range of the simulated WTP data. On the left of the vertical line, the additional price levels are outside the range of actual WTP values. On the right of the vertical line, the price levels are within the range of actual WTP values.

# WEB APPENDIX I MANAGERIAL APPLICATION OF DE-BIASING APPROACH

1. Managers interested in applying the de-biasing approach should first evaluates its applicability and validity to their particular business context and product or service. The discussion of alternative de-biasing approaches in the paper and Tables 1 and 2 provide a basis for such an evaluation. The de-biasing approach makes most sense when it is impossible or very costly to perform a price experiment with actual customers or use and incentive-aligned approach such as BDM.
2. OPTIONAL: If the BDM approach can be applied among a small sample of consumers (Note: In line with Simmons (2014), we recommend a minimum of 50), it can be used to capture the covariance used for the FULL de-biasing approach and to gauge the price range. The price range can be used to inform the price levels in the DC approach. Note that this step is not required for BASIC and EPSILON de-biasing.
3. Carefully define the range of DC prices and the number of levels in between based on the information gathered in step 2. If you did not perform step 2, use the cheapest and most expensive price levels in your particular market as low and high price levels for DC. When in doubt about the DC price levels, rather use a wider than narrower range. To reduce the sensitivity of the results to misspecifications of the price ranges, make sure to use a parametric approach in calculating the DC mean.
4. Collect OE and DC data using a between subjects design. Randomly allocate half of your participants to the DC question and half to the OE. Make sure that you collect a representative random sample of your population of interest (typically your target market).
5. Apply one of the de-biasing approaches BASIC, EPSILON, or FULL (in case you determined the covariance) as described in the paper.
6. Use de-biased demand functions for your pricing decision-making.

# REFERENCES

Hanemann, Michael (1989), “Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data: Reply,“ American Journal of Agricultural Economics, 71(4), 1057-61.

Simmons, Joe (2014), “MTurk vs. The Lab: Either Way We Need Big Samples,” Available at: <http://datacolada.org/18> [Last Accessed: April 24th, 2020].

1. We used the smartphone as a single incentive to motivate participation in our survey in order to recruit an adequate number of subjects. However, the smartphone was not connected to our stimulus and the incentive-aligned condition under BDM, where proper incentives are offered to the participants so that they are motivated to reveal their true preferences (see Wertenbroch and Skiera 2002 for details). [↑](#footnote-ref-2)
2. It is possible that some consumers may have taken part in the survey just to win the smartphone and may not have been interested in purchasing the hooded sweatshirt. However, there is no reason to believe that they have affected our statistical analyses as subjects were randomly assigned to different experimental groups. [↑](#footnote-ref-3)
3. See e.g., the Stanford Bookstore online: https://www.bkstr.com/stanfordstore/home. [↑](#footnote-ref-4)
4. We acknowledge, however, that some participants may have a reference price from similar products in mind. [↑](#footnote-ref-5)
5. In the sweatshirt study, we removed one outlier with a BDM value of CHF 200 who explicitly stated that she intentionally and falsely over-reported her true WTP. [↑](#footnote-ref-6)
6. Valid AWTP estimation requires that the respondents understand the BDM procedure and the underlying buying process. In our sample, respondents understood the BDM mechanism quite well. As Wertenbroch and Skiera (2002) and Miller et al. (2011) did, we asked the subjects if it was clear to them why it was in their best interest to state exactly the price they were willing to pay. Using a seven-point Likert scale from one (not at all) to seven (very much so), the participants responded with an average of 5.803 in the sweatshirt-study. We used a similar method to determine the understanding of the buying process and found an average of 6.070 for the sweatshirt-study. Finally, we asked respondents if stating their WTP for the product was a task which was easy to understand and complete and participants replied with an average of 6.489 in the sweatshirt-study (see Web Appendix E for the exact wording). [↑](#footnote-ref-7)
7. Note: We did not use our empirical data from Study 1 and Study 2 for the simulation study, since we would lose all observations at the price levels we drop from the analysis. That is, we cannot reassign study participants from discarded DC price levels to other DC price levels still included in the simulation study, since their actual WTP is unkown to us. [↑](#footnote-ref-8)