RUHR ECONOMIC PAPERS

Manuel Frondel
Stephan Sommer
Lukas Tomberg

WTA-WTP Disparity: The Role of Perceived Realism of the Valuation Setting

#832

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #832

Responsible Editor: Roland Döhrn

ISSN 1864-4872 (online) – ISBN 978-3-86788-965-0

Ruhr Economic Papers #832

Manuel Frondel, Stephan Sommer, and Lukas Tomberg

WTA-WTP Disparity: The Role of Perceived Realism of the Valuation Setting

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at **http://dnb.dnb.de**

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/86788965
ISSN 1864-4872 (online)
ISBN 978-3-86788-965-0

Manuel Frondel, Stephan Sommer, and Lukas Tomberg[1]

WTA-WTP Disparity: The Role of Perceived Realism of the Valuation Setting

Abstract

Based on a survey among more than 5,000 German households and a single-binary choice experiment in which we randomly split the respondents into two groups, this paper elicits both households' willingness to pay (WTP) for power supply security and their willingness to accept (WTA) compensations for a reduced security level. In accord with numerous empirical studies, we find that the mean WTA value substantially exceeds the mean WTP bid, in our empirical example by a factor of 3.56. Yet, the WTA-WTP ratio decreases to 2.35 among respondents who believe that the hypothetical valuation setting is likely to become true. Conversely, the WTA-WTP ratio increases to 3.81 among respondents who deem the setting unlikely. On the basis of these results, we conclude that inquiring about respondents' perception of the realism of the valuation setting is an easy-to-implement and promising survey element to mitigate excessive WTA-WTP disparities, particularly if private or quasi-public goods are under scrutiny.

JEL-Code: D12, H41, Q41

Keywords: Willingness-to-pay; willingness-to-accept; stated preferences

December 2019

1 Manuel Frondel, RGS Econ, RUB, and RWI; Stephan Sommer, RWI; Lukas Tomberg, RGS Econ and RWI. *– We are highly grateful for comments and suggestions by Ewa Zawojska and Colin Vance. This work has been supported by the Collaborative Research Center "Statistical Modeling of Nonlinear Dynamic Processes" (SFB 823) of the German Research Foundation (DFG), within Project A3, "Dynamic Technology Modeling". Moreover, we gratefully acknowledge financial support by the Federal Ministry of Education and Research (BMBF) under grant 03SFK4B0 (Kopernikus Project ENavi), grant 01UN1203C (Project AKZEPTANZ within the funding measure "Environmentally and Socially Compatible Transformation of the Energy System") and grant 01LA1823A (Project Eval-Map II within the funding measure "Economics of Climate Change").* *- All correspondence to: Stephan Sommer, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: stephan.sommer@rwi-essen.de*

1 Introduction

To gauge the utility that consumers draw from a non-market good, valuation studies typically elicit either consumers' willingness to pay (WTP) for the good or, alternatively, their willingness to accept (WTA) a compensation to forgo it (Carson et al., 2001). Although economic theory suggests that in the absence of income effects, both measures are equivalent (Willig, 1976), empirical findings indicate a strong disparity, with WTA values often exceeding WTP bids substantially – see, for instance, the review studies by Horowitz and McConnell (2002) and Tunçel and Hammitt (2014).

Various explanations for this disparity have been proposed in the literature. A frequent argument is that its causes are rooted in human preferences. For instance, a prominent explanation for the WTA-WTP disparity is the endowment effect (Kahneman et al., 1990; Thaler, 1980), according to which individuals have a higher valuation for goods that they already own compared to those that they have to purchase yet. Another prominent explanation is provided by Hanemann (1991), who demonstrates from a theoretical perspective that the divergence between the WTA and the WTP depends on the degree of substitutability between goods. According to this explanation, a good with close substitutes would exhibit a relatively small WTA-WTP disparity, whereas large disparities would emerge for goods with imperfect substitutes.

Plott and Zeiler (2005), however, argue that the WTA-WTP disparity is not necessarily a manifestation of human preferences. Instead, these authors suggest that the disparity is due to the misconception of a valuation task, which can be reduced by, for instance, extensive training and practice rounds. In a similar vein, Zhao and Kling (2001) suggest that the particularities of a survey or experiment, e.g. limited information on the good under scrutiny and the impossibility to postpone a purchasing decision in such a setting, induce what they refer to as commitment costs. These costs arise because respondents are forced to commit to the purchase of a good before all relevant information is available,which may increase the stated WTA value, but decrease the stated WTP bid, thereby leading to the WTA-WTP disparity.

Assuming that the WTA-WTP disparity is at least partly caused by peculiarities of a survey that are unrelated to human preferences, it seems natural to explore potential correction mechanisms that address such factors. One approach that recently has gained momentum is called consequentiality, with which the strategically motivated misrepresentation of preferences can be reduced (Vossler et al., 2017; Lloyd-Smith and Adamowicz, 2018). Consequentiality requires that respondents expect political consequences, i.e. that their answers influence political decision making, and/or payment consequences, i.e. that the implied payment is coercive (Carson and Groves, 2007; Vossler et al., 2012; Carson et al., 2014; Johnston et al., 2017). Yet, consequentiality is somewhat restrictive in that it refers exclusively to public goods, but precludes the wide area of private as well as quasi-public goods, with quasi-public goods displaying characteristics of both private and public goods, including (partial) excludability (Carson and Groves, 2007). For private and quasi-public goods, coercive payment via taxes and mandatory levies is less common, whereas for public goods a coercive payment vehicle is indispensable for consequentiality to be effective.

Complementing this strand of the literature, this paper proposes inquiring about respondents' perceived realism of the valuation setting as an alternative correction method to reduce the WTA-WTP disparity, but one that is not limited to public goods. Using a discrete-choice experiment that is embedded in a survey of 5,000 German household heads, we employ this easy-to-implement survey element in the context of the quasi-public good of power supply security and elicit both the individuals' willingness to accept (WTA) compensations for allowing the grid operator to disconnect a household from the grid once a year for a maximum of four hours and their willingness to pay (WTP) to avoid such a power outage.

While numerous studies have analyzed the WTP for power supply security (e.g. Blass et al., 2010; Carlsson and Martinsson, 2008; Carlsson et al., 2011), among which just a few have focused on the WTA-WTP disparity (e.g. Beenstock et al., 1998; Hensher et al., 2014), in line with Praktiknjo (2014), who also investigates German households' WTA and WTP for power supply security, we find that the mean WTA value substantially exceeds the mean WTP bid, by a factor of 3.56. Yet, the WTA-WTP ratio decreases to 2.35 if respondents believe that the hypothetical valuation setting is likely to become reality. Conversely, the WTA-WTP ratio increases to 3.81 among respondents who deem the setting unrealistic. Providing empirical evidence on the correlation between the perceived realism of the valuation setting and the respondents' certainty about their WTA and WTP responses, we hypothesize that those household heads who, for example, deem the grid operators' compensation offers in case of a power outage to be a realistic scenario are less likely to expect unforeseen adverse consequences due to their responses than those who deem this an unrealistic scenario and, hence, the former may be more inclined to reveal their actual preferences.

While our analysis is non-causal in nature, as stated perceptions obtained by follow-up questions do not allow for causal inference (Czajkowski et al., 2017), it bears noting that our key finding of a shrinking WTA-WTP-disparity is not confined to power supply security. Rather, it should apply to all those goods for which public policies play a role, such as in natural monopolies and the subsidization of culture. In fact, the proposed correction mechanism of inquiring about respondents' perceived realism of the valuation setting can be employed for all those goods that are not traded on markets, but whose consumption is – in contrast to public goods – excludable.

The following three sections describe the database, the design of the experiment, and the econometric approach. Thereafter, we present the results of our empirical analysis, not least empirical evidence on the correlation between the perceived realism of the valuation setting and the respondents' certainty about their WTA and WTP responses. The final section summarizes and concludes.

2 Data

To elicit people's preferences for power supply security, we draw on survey data that was gathered in collaboration with the professional survey institute *forsa*. *forsa* maintains a panel of German households that is representative of the German-speaking population aged 14 and above and a state-of-the-art tool that allows panelists to fill out the questionnaire using either a television or the internet. Panelists can interrupt and continue the survey at any time. Between December 23, 2015, and February 19, 2016, we asked 5,640 household heads a variety of energy-related questions, such as their experience with power outages, and about their socioeconomic characteristics (Table 1).

Table 1: Summary Statistics

Variable	Explanation	Full Sample	WTA Group	WTP Group	t Statistics
Age	Age of respondent	55.11	55.31	54.91	-1.13
Female	Dummy: 1 if respondent is a woman	0.377	0.370	0.384	1.10
Children	Dummy: 1 if children live in the household	0.141	0.137	0.145	0.81
East Germany	Dummy: 1 if household resides in East Germany	0.198	0.197	0.200	0.27
College degree	Dummy: 1 if respondent has a college degree	0.331	0.333	0.328	-0.40
Green voter	Dummy: 1 if respondent tends to vote green	0.100	0.102	0.098	-0.53
Income	Monthly household net income in €	2,893	2,908	2,879	-0.75
Outage >30min	Dummy: 1 if respondent has experienced an outage of at least 30 minutes in the last 5 years	0.362	0.368	0.355	-0.94
No. of outages	Number of outages in the last 5 years	1.58	1.62	1.54	-0.95
Duration	Duration of longest outage in minutes in the last 5 years	108	112	105	-0.66
	Share of dropouts:	0.039	0.035	0.043	1.55
	Number of observations:	5,640	2,775	2,865	

Note: Based on the t statistics reported in the last column, we test on differences in variable means across WTP and WTA groups.

About 38% of the participants are women, which is likely due to our decision to ask the household head to participate, as household heads typically make the financial decisions at the household level. Obviously, this percentage does not reflect the share of women in the German population, but it closely fits the percentage of female household heads in Germany (see Table A1 in the appendix). Around a third of the sample holds a college degree, whereas this only holds for about one fifth of the household heads in the German census data. Consequently, well-educated household heads are overrepresented in our sample, which is typical for non-census-based surveys in Germany (Andor et al., 2014), and thus our results are not representative for the German population. To control for environmental attitudes, we construct a binary variable that equals unity if a respondent tends to vote for Germany's green party. This holds true for about 10% of the respondents, which is broadly in line with the result of the most recent national election in 2017.

As experience with power outages might matter for the preferences on power supply security (e.g. Carlsson et al., 2011), we asked several questions pertaining to this issue. For instance, 36.2% of the respondents indicate that, during the last five years, they experienced at least one outage that lasted at least 30 minutes, and about 6% experienced an outage of at least four hours. The low level of experience with outages fits official data: According to the German regulator BNetzA (2019), in the survey year 2016, each customer was disconnected from the grid for only about 12 minutes, on average, as results from the so-called System Average Interruption Duration Index (SAIDI), an indicator for the level of supply security.

3 Experimental Design

At the outset of the experiment, all participants were informed – in German, of course – that for "reasons of power supply security, it might sometimes be necessary to disconnect individual households from the grid". To uncover the value that households ascribe to power supply security, we employed an experimental design in which respondents were randomly split into two experimental groups, the WTA and the WTP group. Following the recommendation of the NOAA-Panel (Arrow et al., 1993), we then asked a single binary-choice question that differs across experimental groups with respect to the underlying preference measure: In the WTA group, we asked participants the following question: "Imagine that your local grid operator offers you a monthly compensation of €x to be allowed to disconnect your household once in a year from the grid for a maximum of four hours without prior notification. Would you accept this offer?" The concrete amount of x was assigned in a random manner, with $x \in \{3, 6, 9\}$.[1]

In the WTP group, participants were asked the following question: "Imagine that your local grid operator asks whether you are willing to pay €x to avoid outages. Otherwise, the grid operator would be allowed to disconnect your household from the grid once a year for a maximum of four hours without prior notification. Would you accept this payment?" The concrete amount of $x \in \{3, 6, 9\}$ was also assigned in a random manner.

Note that the randomly assigned values of €3 to €9 per month, that is, €36 to €108 per year, are within the range of those mean WTP and WTA values that are found in other valuation studies of four-hour power outages for Germany (Praktiknjo, 2014; Schubert et al., 2013). In addition to accepting or rejecting these offers, respondents had the option to indicate a "Don't know" response. With 7.7%, a slightly higher share of respondents made use of this option in the WTP group than in the WTA group (6.2%). While this information is omitted from the analysis presented in Section 5, it will be taken into account in Section 6.

That randomization was successful is indicated by the t statistics reported in the last column of Table 1: There are no statistically significant deviations in the means of socio-economic characteristics across both experimental groups, leading to the conclusion that our empirical results are not driven by systematic differences in the distributions of the explanatory variables.

Subsequent to the valuation question, we asked respondents who answered "Yes" about the certainty in their response. In detail, in the WTA group, this question reads:

[1] In the survey design, we did not differentiate between planned and unplanned power outages. Our intuition is that for an unplanned outage both the WTA and the WTP would be higher than for a planned outage. It remains unclear, though, how the focus on unplanned outages would affect the WTA-WTP ratio. Only if both measures are larger by the same factor for unplanned outages, this would not have any bearing on the ratio of both measures.

”How certain are you that you would actually accept this offer of €x?”. In the WTP group, the corresponding question reads: ”How certain are you that you would actually be willing to pay €x?”. This certainty approach builds upon the work of Champ et al. (1997) and Blumenschein et al. (2007). A large share of 66% of the respondents who accepted the WTA offer indicated that they are ”very certain” about their decision. In the WTP group, the share of very certain ”Yes” responses was slightly lower at 59%.

On a next screen, we elicited the information on our central explanatory variable, by asking the respondents about their perceptions on the likelihood that their electricity supplier will actually make any offer, for example a compensation for reducing power security: “What do you think, how likely is it that your electricity supplier approaches you with such an offer?” The answers to this question are measured on a 5-point Likert scale, ranging from $j = 1$ (very unlikely) to $j = 5$ (very likely).

Exploiting this information on the households’ perception about the realism of the hypothetical setting plays a key role in our empirical analysis. Around 70% of the respondents deem any such offer as very unlikely (Table 2), while only a minority of about 4% of the respondents believes that such an offer is very likely to become true. As indicated by the t statistics reported in the last column of Table 2, there are hardly any differences in the respondents’ perceptions between the WTA and the WTP group. Moreover, the perceived probability that any such offer actually becomes reality does not vary significantly across hypothetical bids and offers (see Table A2 in the appendix).

Beyond dividing our sample into either the WTA or the WTP group, we moreover split the respondents into two additional groups: those who believe that the offer from the electricity supplier is unlikely ($j = 1, 2$) and those who perceive that the offer is likely, at least to some extent ($j = 3, \ldots, 5$). Accordingly, as the central variable employed in our empirical approach, we define a dummy variable π that equals unity if $j = 3, \ldots, 5$ and equals zero otherwise (see Table 2).

Table 2: Responses to the Question: "What do you think: How likely is it that your electricity supplier approaches you with such an offer?"

Response Category	j	π	Full Sample	WTA Group	WTP Group	t Statistics
Very unlikely	$(j = 1)$	0	70.8%	71.9%	69.7%	-1.71
Quite unlikely	$(j = 2)$	0	16.5%	16.2%	16.8%	0.56
Moderately likely	$(j = 3)$	1	6.3%	6.0%	6.6%	0.93
Quite likely	$(j = 4)$	1	2.2%	2.1%	2.3%	0.47
Very likely	$(j = 5)$	1	4.2%	3.9%	4.6%	1.38
Number of observations:			5,284	2,621	2,663	

Note: Based on the t statistics reported in the last column, we test on differences in variable means across WTP and WTA groups.

4 Econometric Approach

The focus of our analysis is on the effect of the perception about the realism of the valuation setting, reflected by variable π, on WTA and WTP. Obviously, a simple comparison of groups for which π either equals zero or unity, as is presented in the subsequent section, will not yield the causal effect of π on WTA and WTP, as π is likely to be endogenous: given that it is unclear whether all confounding factors are captured by our covariates and both groups may differ in unobservable characteristics, potential endogeneity is a caveat of our analysis.[2]

To obtain estimates of the mean WTA and WTP values, we use the approach developed by Bishop and Heberlein (1979) and analyzed by Buckland et al. (1999), Hanemann (1984), and Hanemann and Kanninen (2001). This approach rests on estimating a logistic regression model, such as the following Model (1):

$$Pr(y_i = 1|bid_i, \pi_i, \mathbf{x}_i) = \Lambda[\beta_0 + \beta_1 \ln(bid_i) + \beta_2\pi_i + \boldsymbol{\beta}_{\mathbf{x}}^T \mathbf{x}_i], \tag{1}$$

where y_i is a binary indicator that equals unity if respondent i is willing to accept ei-

[2] A potential causal approach would be to employ an instrumental variable for π that is correlated with π, but uncorrelated with WTA and WTP. A candidate instrument, however, is not available in our case, as well as in most similar studies (e.g. Vossler et al., 2012; Vossler and Watson, 2013; Interis and Petrolia, 2014).

ther €3, €6, or €9 as a compensation for an outage or is willing to pay either of these amounts for avoiding an outage, and equals zero otherwise. Λ is the logistic cumulative distribution function (cdf), where the hypothetical compensation or payment (*bid*) is included in logarithmic terms. By using this logarithmic transformation, it is ruled out that the resulting mean WTA and WTP values become negative, which is a frequent problem when analyzing discrete-choice contingent valuation data (Haab and McConnell, 1997). In fact, this transformation ensures that the probability of accepting a compensation offer approaches zero if the proposed compensation goes to zero. In the WTP case, it ensures that the probability of accepting the offer approaches unity if the offer price goes to zero (Hanemann, 1989). Vector $\mathbf{x}$ subsumes the covariates reported in Table 1, the βs are parameters to be estimated, and π denotes the dummy variable on the perceived realism of the valuation setting described in the previous section.

Using the conditional density function of y given by Model (1), the mean WTA, which renders a respondent indifferent between accepting the offered compensation or declining it, can be estimated on the basis of the following expression (Hanemann, 1984):

$$E[WTA_i|\pi_i, \mathbf{x}_i] = \int_0^{\infty} \Lambda[\hat{\beta}_0 + \hat{\beta}_1 \ln(bid_i) + \hat{\beta}_2 \pi_i + \hat{\boldsymbol{\beta}}_{\boldsymbol{x}}^T \mathbf{x}_i] \, \mathrm{d}bid_i, \tag{2}$$

where $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ and $\hat{\boldsymbol{\beta}}_{\boldsymbol{x}}$ denote the coefficient estimates obtained from the logit regression of Model (1). Building upon Buckland et al. (1999) and Equation (2), we compute the mean WTA for each observation, that is, at individual-specific values of π and $\mathbf{x}$, taking into account that the probability $Pr(y_i = 1)$ of accepting an offer is a function of all possible bids (Parzen, 1960; Hanemann, 1984): $Pr(y_i = 1) = \Lambda_i(bid)$. Integrating the resulting cumulative density function $\Lambda_i(bid)$ over the entire range of bids yields the observation-specific mean WTA: $E[WTA_i]$. Lastly, we arrive at sample means for WTA by averaging $E[WTA_i]$ over all observations.

Because the probability $Pr(y_i = 1)$ of accepting an offer is increasing in the bids in the WTA case, but decreasing in case of WTP, the formula for the mean WTP differs from Equation (2) for the mean WTA – see the theoretical and graphical motivation for this differentiation provided by Hanemann (1984):

$$E[WTP_i|\pi_i, \mathbf{x}_i] = \int_0^\infty \left\{1 - \Lambda[\hat{\beta}_0 + \hat{\beta}_1 \ln(bid_i) + \hat{\beta}_2 \pi_i + \hat{\boldsymbol{\beta}}_3^T \mathbf{x}_i]\right\} \, \mathrm{d}bid_i, \tag{3}$$

but the procedure to arrive at sample means for WTP is identical to that described above.

Following Bishop and Heberlein (1979), to avoid obtaining unreasonably high estimates of mean WTA and WTP values, we only integrate Equations (2) and (3) up to the highest bid of €9, rather than integrating up to infinity, thereby providing conservative estimates of the mean values (Hanemann and Kanninen, 2001). This approach is in line with the NOAA-Panel's recommendation to prefer a conservative estimate over an estimate that risks being too large (Arrow et al., 1993).

5 Results

Focusing first on the descriptive results, in line with economic theory, the willingness to accept compensations for tolerating a power outage increases with the compensation level, yet the difference between compensations of €6 and €3 is indistinguishable in statistical terms (first column of Table 3). Conversely, the willingness to pay for avoiding a power outage decreases with the amounts that respondents have to pay (second column of Table 3).

Turning next to the subsamples of households that deem compensation offers and WTP payments to be realistic scenarios (third and fifth column), we find the rates of accepting any such offer to be notably higher than in the full sample. In contrast, the acceptance rates are lowest among those individuals who believe that any such offer is unlikely (fourth and sixth column). Overall, the pattern that acceptance rates increase

Table 3: Percentage of Yes Responses to Randomly Given Compensation Offers (WTA) and WTP Payments

	Full Sample		WTA Group		WTP Group	
			Perceived Likelihood of an Offer			
	WTA	WTP	Likely ($\pi = 1$)	Unlikely ($\pi = 0$)	Likely ($\pi = 1$)	Unlikely ($\pi = 0$)
€3	16.3%	23.2%	23.7%	15.5%	31.0%	21.6%
	–	–	–	–	–	–
€6	19.2%	15.0%	32.4%	17.7%	21.9%	13.7%
	(1.53)	(-4.30)**	(1.36)	(1.13)	(-1.54)	(-3.96)**
€9	23.7%	11.8%	33.0%	22.6%	15.2%	11.3%
	(3.71)**	(-6.21)**	(1.42)	(3.41)**	(-2.84)**	(-5.29)**
Number of obs.	2,512	2,528	293	2,172	326	2,144

Note: t statistics for testing differences in means relative to the €3 group are in parentheses. ** and * indicate statistical significance at the 1% and 5% level, respectively.

with the compensation level and decrease with the payment amount is observed for both the full sample and the subsamples, indicating the internal validity of our results.

To obtain mean values for WTA and WTP, we employ the approach developed by Bishop and Heberlein (1979) and separately estimate Model (1) for each experimental group. The estimation results reported in Table 4 indicate that if a respondent believes that the hypothetical offer is likely to become reality, both the WTP for avoiding an outage and the WTA for accepting an outage increase. While the latter implies a lower mean WTA, in combination with a lower WTP, this results in a smaller WTA-WTP ratio.

With a few exceptions, socio-economic characteristics do not have any significant bearing on the WTA for tolerating an unplanned outage of four hours. One of these exceptions is household income. In line with economic theory, the WTA decreases with income, whereas the WTP for the avoidance of an outage increases with income. In contrast, the experience with outages does not affect the WTP for power supply security, nor does it impact the WTA for a diminished supply security. This result seems to be at odds with the empirical literature, as for instance Carlsson et al. (2011) find that the WTP for power supply security increases after an outage. However, that study is unique as it was con-

Table 4: Logit Estimation Results for Model (1), separately estimated for the WTA and WTP Group

	WTA Group		WTP Group	
	Coeff.	Std. Error	Coeff.	Std. Error
ln (Bid)	0.405**	(0.130)	-0.772**	(0.132)
Offer deemed likely ($\pi = 1$)	0.654**	(0.158)	0.556**	(0.162)
Age	-0.008	(0.004)	0.009	(0.005)
Female	-0.034	(0.124)	0.359**	(0.128)
East Germany	0.190	(0.139)	-0.472**	(0.173)
Children in household	0.025	(0.183)	0.191	(0.189)
College degree	0.364**	(0.124)	0.240	(0.132)
ln(Income)	-0.444**	(0.116)	0.533**	(0.133)
Experience with outage	-0.072	(0.119)	-0.127	(0.130)
Green voter	0.190	(0.186)	-0.318	(0.217)
Constant	1.531	(0.950)	-5.212**	(1.098)
Number of observations:	2,017		2,077	

Note: ** and * indicate statistical significance at the 1% and 5% level, respectively.

ducted shortly after a severe storm in Sweden that led to power outages of up to three weeks for about 15% of the households, which many of the respondents likely recalled.

Inserting the coefficient estimates reported in Table 4 into Equations (2) and (3) allows us to predict WTA and WTP values for each household observation. Averaging over all individual values yields the mean WTA and WTP for power supply security and a mean WTA-WTP ratio of 3.56 (see first row of Table 5). Our empirical results indicate that this ratio critically depends on whether a respondent believes that the hypothetical setting is likely to become true: Among respondents who deem the offer likely, that is, for $\pi_i = 1$, the mean WTA-WTP ratio shrinks to 2.35. In contrast, among respondents who deem any such offer as unlikely ($\pi_i = 0$), the mean WTA-WTP ratio increases to 3.81. It bears noting that the differences in mean WTA across π values, as well as the differences in mean WTP, are statistically significant, as can be seen from the 95% confidence intervals presented in Table 5.

In sum, our results suggest that the respondents' perception of the realism of our hy-

Table 5: Mean Values for the Willingness to Accept (WTA) a Power Outage of 4 hours once in a year and the Willingness to Pay (WTP) to Avoid siuch a Power Outage

Probability of the offer	Mean WTA	Mean WTP	WTA/WTP ratio
π	7.47	2.10	3.56
	[7.28 – 7.62]	[1.84 – 2.35]	[3.18 – 4.06]
$\pi = 1$	6.65	2.83	2.35
	[6.17 – 7.09]	[2.31 – 3.29]	[1.99 – 2.87]
$\pi = 0$	7.58	1.99	3.81
	[7.39 – 7.73]	[1.73 – 2.26]	[3.34 – 4.40]

Note: 95% Confidence intervals are in brackets and obtained via bootstrapping with 1,000 replications.

pothetical valuation setting has a significant bearing on the WTA-WTP disparity: If respondents deem the hypothetical setting likely to become reality, at least to some extent, the disparity shrinks (see the line with $\pi_i = 1$ in Table 5), but if the setting is perceived as unlikely, the disparity grows (see the line with $\pi_i = 0$ in Table 5). We hypothesize that respondents who deem the setting likely to become reality are less likely to expect unforeseen adverse consequences from their responses than those who deem the setting unrealistic and, hence, may be more inclined to answer according to their actual preferences.

6 Perceived Realism and Uncertainty

In this section, we investigate the relationship between the perceived realism of the valuation setting and a respondent's certainty about her/his answer to the valuation question, thereby exploiting two additional kinds of information on uncertainty: First, while being ignored so far, we include the "Don't know" responses to the valuation question in the subsequent analysis. Second, pursuing the certainty approach described in Section 3, we incorporate the respondents' certainty with respect to their "Yes" answers to the valuation question in a multinomial logit model in which the categorical outcome variable reflects the following four response categories: (1) a certain or (2) an uncertain

”Yes” response, (3) a ”Don’t know” response, with (4) the base category standing for a ”No” response to the valuation question. While the covariates of Model (1) are also included in this multinomial logit model, Table 6 presents the marginal effects of the perceived realism of the valuation setting ($\pi = 1$) on these four response categories. The whole set of coefficient estimates for both the WTA and the WTP group are reported in the appendix, in Table A3 and Table A4, respectively.

Table 6: Marginal Effects of $\pi = 1$ (i. e. the Hypothetical Setting is Deemed Likely), resulting from a Multinomial Logit Model on the Responses to the Valuation Question

Response Options	WTA Group		WTP Group	
No	-0.110**	(0.031)	-0.094**	(0.028)
Certain Yes	0.083**	(0.026)	0.054**	(0.021)
Uncertain Yes	0.021	(0.018)	0.023	(0.017)
Don’t know	0.005	(0.015)	-0.017	(0.017)
Number of observations:	2,111		2,218	

Note: ** and * indicate statistical significance at the 1% and 5% level, respectively.

The marginal effects indicate that in both groups, respondents who deem the hypothetical setting likely are 11.0 (WTA group) and 9.4 percentage points (WTP group) less likely to reject any offer of their grid operator, respectively (see first row of Table 6), whereas for ”Don’t know” and uncertain ”Yes” responses, the marginal effects are statistically indistinguishable from zero. In contrast, the likelihood for a certain ”Yes” response is significantly higher among those respondents who deem any offer of their grid operator on supply security likely. In short, respondents who perceive the hypothetical valuation setting as realistic rather say ”Yes”, instead of ”No”, to any such offer. These results imply the decrease of the WTA-WTP disparity.

The underlying reasons for these results are subject to speculation, but may be rooted in the desire to reduce the risk of regretting a ”Yes” response when ex-ante unknown adverse consequences of this answer would materialize. The perception of this risk may be much weaker for those who believe in the realism of the valuation setting.

7 Conclusion

Using a discrete-choice experiment within a hypothetical setting that is embedded in survey among 5,000 German household heads, this paper has elicited both the willingness to accept (WTA) compensations for allowing the grid operator to disconnect a household from the grid once a year for a maximum of four hours and the willingness to pay (WTP) to avoid such a power outage. In accordance with numerous studies originating from other areas, we find that the mean WTA value substantially exceeds the mean WTP bid, in our empirical example by a factor of 3.56.

Yet, when accounting for respondents' perception about the realism of the hypothetical setting, we find that the WTA-WTP ratio decreases to 2.35 if respondents believe that the hypothetical valuation setting is likely to become true. Conversely, the WTA-WTP ratio increases to 3.81 if respondents deem the setting unlikely. These results indicate that observed WTA values might overestimate genuine WTA values, whereas observed WTP bids might be underestimated if the survey design does not account for the respondents' perception of the realism of the setting.

Based on these results, we conclude that to diminish the WTA-WTP disparity that typically result from stated-preference surveys, inquiring about respondents perception of the realism of the valuation setting is an easy-to-implement and promising element of valuation surveys. We hypothesize that those household heads who, for example, deem the grid operators' compensation offers in case of a power outage to be a realistic scenario are less likely to expect unforeseen adverse consequences due to their responses than those who deem this an unrealistic scenario and, hence, the former may be more inclined to reveal their actual preferences.

Moreover, we interpret our findings as support for the argument by Plott and Zeiler (2005) that the frequently observed WTA-WTP disparity may not be entirely a manifestation of human preferences. Next, given the high level of power supply security in Germany, a hypothetical scenario in which power outages of 4 hours may occur clearly implies a degradation relative to the status quo. In this case, which is denoted by Knetsch (2010) as a domain of losses, this author argues that the value of changes in the provision with a good, such a power supply security, will be more accurately assessed with the WTA, rather than the WTP measure. Given the widely found WTA-WTP disparity, this argument implies that taking the appropriate valuation measure is a critical issue for drawing policy implications. Not least, we also believe that our empirical results are relevant for the evaluation of new technologies and innovative policy measures whose realization appears unlikely: Failing to account for the perceived realism of the valuation setting might result in an excessive WTA-WTP disparity. Hence, we believe it behooves researchers to ask a debriefing question about the realism of the valuation setting.

While this approach to reduce the WTA-WTP disparity can be assumed to be not only valid for power supply security, but for both private and quasi-public goods in general, it bears noting that our analysis is non-causal in nature, because it is unclear whether the stated perceptions on the realism of the valuation setting are subject to unobserved confounding factors that simultaneously influence the answers to both the WTA and WTP questions (Czajkowski et al., 2017). Therefore, further experimental investigations are indispensable to verify the causal nature of the relationship between the perceived realism of the valuation setting and the WTA-WTP disparity.

Appendix

Tables

Table A1: Comparison of the Sample with the Population of Household Heads in Germany

Variable	Sample	Population
Age below 25 years	1.2%	4.6%
Age 25 – 64 years	71.6%	67.0%
Age 65 and above	27.2%	28.4%
Female household head	37.7%	35.5%
East Germany	19.8%	20.9%
Single household	23.9%	41.4%
2 person household	45.7%	34.2%
3 person household	13.6%	12.1%
4 and more person household	16.8%	12.3%
College degree	33.1%	20.4%
High income	11.7%	11.8%

Note: Data for the population are drawn from Destatis (2016) where the main income earner is asked to complete the survey, whereas we ask the household head.

Table A2: Percentage of Respondents who Deem the Offer at least Moderately Likely (i. e. for which $\pi = 1$) across Experimental Groups

Bid/Offer	WTA Group	WTP Group
€3	11.7%	14.2%
	–	–
€6	12.1%	12.9%
	(0.25)	(-0.79)
€9	12.1%	13.5%
	(0.26)	(-0.39)
Number of observations:	2,621	2,663

Note: t statistics for testing differences in means relative to the €3 group are in parentheses.

Table A3: Multinomial Logit Estimations Results for the WTA Group (Base Category: "No" Response to Valuation Question)

Response Options	Certain "Yes"		Uncertain "Yes"		"Don't know"	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
ln (Bid)	0.483**	(0.157)	0.263	(0.211)	0.475*	(0.236)
Offer deemed likely ($\pi = 1$)	0.722**	(0.182)	0.483	(0.260)	0.268	(0.310)
Age	-0.003	(0.005)	-0.013	(0.007)	0.001	(0.008)
Female	-0.019	(0.146)	-0.125	(0.204)	0.275	(0.213)
East Germany	0.088	(0.168)	0.285	(0.218)	0.321	(0.239)
Children in household	0.134	(0.214)	-0.202	(0.309)	0.456	(0.305)
College degree	0.272	(0.149)	0.584**	(0.198)	0.081	(0.229)
ln(Income)	-0.491**	(0.137)	-0.432*	(0.188)	-0.548**	(0.205)
Experience with outage	0.084	(0.140)	-0.328	(0.203)	0.062	(0.211)
Green voter	0.346	(0.211)	-0.261	(0.347)	0.498	(0.302)
Constant	1.064	(1.126)	0.966	(1.519)	0.234	(1.690)
Number of observations:			2,111			

Note: ** and * indicate statistical significance at the 1% and 5% level, respectively.

Table A4: Multinomial Logit Estimations Results for the WTP Group (Base Category: "No" Response to Valuation Question)

Response Options	Certain "Yes"		Uncertain "Yes"		"Don't know"	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
ln (Bid)	-0.785**	(0.168)	-0.721**	(0.194)	-0.059	(0.190)
Offer deemed likely ($\pi = 1$)	0.664**	(0.200)	0.472*	(0.236)	0.368	(0.234)
Age	0.009	(0.007)	0.010	(0.008)	-0.004	(0.007)
Female	0.184	(0.167)	0.520**	(0.185)	0.855**	(0.176)
East Germany	-0.509*	(0.230)	-0.335	(0.241)	-0.417	(0.242)
Children in household	0.130	(0.239)	0.245	(0.280)	-0.026	(0.270)
College degree	0.165	(0.167)	0.256	(0.194)	0.024	(0.194)
ln(Income)	0.841**	(0.178)	0.180	(0.188)	0.165	(0.175)
Experience with outage	-0.269	(0.168)	-0.035	(0.189)	-0.062	(0.182)
Green voter	-0.205	(0.271)	-0.527	(0.345)	-0.090	(0.277)
Constant	-8.141**	(1.489)	-3.590*	(1.546)	-3.736**	(1.429)
Number of observations:			2,218			

Note: ** and * indicate statistical significance at the 1% and 5% level, respectively.

References

Andor, M., Frondel, M., and Vance, C. (2014). Hypothetische Zahlungsbereitschaft für grünen Strom: Bekundete Präferenzen privater Haushalte für das Jahr 2013. *Perspektiven der Wirtschaftspolitik*, 15(4):355–366.

Arrow, K., Solow, R., Portney, P. R., Leamer, E. E., Radner, R., and Schuman, H. (1993). Report of the NOAA panel on contingent valuation. *Federal Register*, 58(10):4601–4614.

Beenstock, M., Goldin, E., and Haitovsky, Y. (1998). Response bias in a conjoint analysis of power outages. *Energy Economics*, 20(2):135–156.

Bishop, R. C. and Heberlein, T. A. (1979). Measuring values of extramarket goods: Are indirect measures biased? *American Journal of Agricultural Economics*, 61(5):926–930.

Blass, A. A., Lach, S., and Manski, C. F. (2010). Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. *International Economic Review*, 51(2):421–440.

Blumenschein, K., Blomquist, G. C., Johannesson, M., Horn, N., and Freeman, P. (2007). Eliciting willingness to pay without bias: Evidence from a field experiment. *The Economic Journal*, 118(525):114–137.

BNetzA (2019). Kennzahlen der Versorgungsunterbrechungen Strom. Bundesnetzagentur, Bonn. `https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Versorgungssicherheit/Versorgungsunterbrechungen/Auswertung_Strom/Versorgungsunterbrech_Strom_node.html`.

Buckland, S. T., MacMillan, D. C., Duff, E. I., and Hanley, N. (1999). Estimating mean willingness to pay from dichotomous choice contingent valuation studies. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 48(1):109–124.

Carlsson, F. and Martinsson, P. (2008). Does it matter when a power outage occurs?–A choice experiment study on the willingness to pay to avoid power outages. *Energy Economics*, 30(3):1232–1245.

Carlsson, F., Martinsson, P., and Akay, A. (2011). The effect of power outages and cheap talk on willingness to pay to reduce outages. *Energy Economics*, 33(5):790–798.

Carson, R. T., Flores, N. E., and Meade, N. F. (2001). Contingent valuation: Controversies and evidence. *Environmental and Resource Economics*, 19(2):173–210.

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

Carson, R. T., Groves, T., and List, J. A. (2014). Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, 1(1/2):171–207.

Champ, P. A., Bishop, R. C., Brown, T. C., and McCollum, D. W. (1997). Using donation mechanisms to value nonuse benefits from public goods. *Journal of Environmental Economics and Management*, 33(2):151–162.

Czajkowski, M., Vossler, C. A., Budziński, W., Wiśniewska, A., and Zawojska, E. (2017). Addressing empirical challenges related to the incentive compatibility of stated preferences methods. *Journal of Economic Behavior & Organization*, 142:47–63.

Destatis (2016). Bevölkerung und Erwerbstätigkeit. Haushalte und Familien. Ergebnisse des Mikrozensus. Artikelnummer: 2010300157004. Statistisches Bundesamt, Wiesbaden.

Haab, T. C. and McConnell, K. E. (1997). Referendum models and negative willingness to pay: Alternative solutions. *Journal of Environmental Economics and Management*, 32(2):251–270.

Hanemann, W. M. (1984). Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics*, 66(3):332–341.

Hanemann, W. M. (1989). Welfare evaluations in contingent valuation experiments with discrete response data: Reply. *American Journal of Agricultural Economics*, 71(4):1057–1061.

Hanemann, W. M. (1991). Willingness to pay and willingness to accept: How much can they differ? *American Economic Review*, 81(3):635–647.

Hanemann, W. M. and Kanninen, B. (2001). The statistical analysis of discrete-response CV data. In Bateman, I. J. and Willis, K. G., editors, *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*, chapter 11, 302–441. Oxford University Press, Oxford.

Hensher, D. A., Shore, N., and Train, K. (2014). Willingness to pay for residential electricity supply quality and reliability. *Applied Energy*, 115:280–292.

Horowitz, J. K. and McConnell, K. E. (2002). A review of WTA/WTP studies. *Journal of Environmental Economics and Management*, 44(3):426–447.

Interis, M. G. and Petrolia, D. R. (2014). The effects of consequentially in binary-and multinomial-choice surveys. *Journal of Agricultural and Resource Economics*, 39(2):201–216.

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., and Vossler, C. A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2):319–405.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98(6):1325–1348.

Knetsch, J. L. (2010). Values of gains and losses: Reference states and choice of measure. *Environmental and Resource Economics*, 46(2):179–188.

Lloyd-Smith, P. and Adamowicz, W. (2018). Can stated measures of willingness-to-accept be valid? Evidence from laboratory experiments. *Journal of Environmental Economics and Management*, 91:133–149.

Parzen, E. (1960). *Modern Probability Theory and Its Applications*. John Wiley & Sons, New York.

Plott, C. R. and Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the "Endowment Effect", subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review*, 95(3):530–545.

Praktiknjo, A. J. (2014). Stated preferences based estimation of power interruption costs in private households: An example from Germany. *Energy*, 76:82–90.

Schubert, D. K. J., von Selasinsky, A., Meyer, T., Schmidt, A., Thuß, S., Erdmann, N., Erndt, M., and Möst, D. (2013). Gefährden Stromausfälle die Energiewende? Einfluss auf Akzeptanz und Zahlungsbereitschaft. *Energiewirtschaftliche Tagesfragen*, 63(10):35–39.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.

Tunçel, T. and Hammitt, J. K. (2014). A new meta-analysis on the WTP/WTA disparity. *Journal of Environmental Economics and Management*, 68(1):175–187.

Vossler, C. A., Bergeron, S., Doyon, M., and Rondeau, D. (2017). Revisiting the gap between the willingness-to-pay and willingness-to-accept for public goods. 23rd Annual Conference of the European Association of Environmental and Resource Economists, Athens.

Vossler, C. A., Doyon, M., and Rondeau, D. (2012). Truth in consequentiality: Theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, 4(4):145–171.

Vossler, C. A. and Watson, S. B. (2013). Understanding the consequences of consequentiality: Testing the validity of stated preferences in the field. *Journal of Economic Behavior & Organization*, 86:137–147.

Willig, R. D. (1976). Consumer's surplus without apology. *American Economic Review*, 66(4):589–597.

Zhao, J. and Kling, C. L. (2001). A new explanation for the WTP/WTA disparity. *Economics Letters*, 73:293–300.